Identification-robust analysis of DSGE and structural macroeconomic models

Jean-Marie Dufour[a], Lynda Khalaf[b], Maral Kichian[c] *†

[a] McGill University; [b] Carleton University; [b] University of Ottawa.

Abstract

Full- and limited-information identification-robust methods are proposed for structural systems, notably DSGE models, which are valid whether identification is weak or strong, theory-intrinsic or data-specific. The proposed methods are applied to a standard New Keynesian system for the U.S. Single- and multi-equation estimation and fit are also compared. When a unique rational-expectation stable equilibrium is imposed, the model is rejected. In contrast, limited-information inference produces informative results regarding forward-looking behavior in the NKPC and precise conclusions on feedback coefficients in the reaction function, which cannot be reached via single-equation methods.

Keywords: Identification-Robust Inference; DSGE; New Keynesian model; Full-information; Limited-information.

JEL classification: C52, C53, E37

---

*The authors thank Carlos de Resende, Frank Kleibergen, Vadim Marmer, Sophocles Mavroeidis, Ulrich Müller, Frank Schorfheide, two anonymous referees, and the Editor Robert G. King for several useful comments. Timothy Grieder and Erick Moyneur provided valuable research assistance. This work was supported by the Bank of Canada (Research Fellowship), the William Dow Chair in Political Economy (McGill University), the Canada Research Chair Program (Econometrics, Université de Montréal, and Environment, Université Laval), the Institut de Finance Mathématique de Montréal (IFM2), the Alexander-von-Humboldt Foundation (Germany), the Canadian Network of Centres of Excellence (program on Mathematics of Information Technology and Complex Systems [MITACS]), the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Fonds de Recherche sur la Société et la Culture (Québec), the Fonds de Recherche sur la Nature et les Technologies (Québec). The views in this paper are our own and do not necessarily reflect those of the Bank of Canada.

†[a] Department of Economics, McGill University, Leacock Building, Room 519, 855 Sherbrooke Street West, Montréal, Québec H3A 2T7, Canada. TEL: (1) 514 398 4400 ext. 09156; FAX: (1) 514 398 4938; Email: jean-marie.dufour@mcgill.ca; Homepage: http:\\www.jeanmariedufour.com. [b] Economics Department, Carleton University, Loeb Building 1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6 Canada. Tel (1) 613 520 2600 ext. 8697; Email: Lynda_Khalaf@carleton.ca. [c] Corresponding author. Economics Department and Graduate School of Public and International Affairs, University of Ottawa, 120 University, Ottawa, Canada. TEL: (1) 613 562 5800 ext. 1872; Email: mkichian@uottawa.ca.

1. Introduction

Optimization-based macroeconomic models, in particular structures derived from dynamic stochastic general equilibrium (DSGE) assumptions, are routinely used for analyzing macroeconomic issues. In this respect, the solutions of log-linearized versions of these models are frequently taken to the data in order to obtain realistic quantitative answers to the questions studied. Classical and Bayesian estimations have both been used for this purpose, including methods that consider jointly all model restrictions (full-information [**FI**] approaches), and methods that focus on matching only some aspects of the data (limited-information [**LI**] approaches). However, finding reliable estimates for the parameters of such models is a challenging problem, regardless of the estimation strategy. In a recent survey, Schorfheide (2010) discusses, among others, two important (and related) reasons for this: weak identification and assumptions which are auxiliary to the theory yet necessary to complete a model, such as restrictions on disturbance distributions and information sets. This paper studies both problems, proposes econometric tools designed to overcome their consequences, and applies these tools to the New Keynesian model.

A number of studies have documented identification problems in well-known estimated models such as the New Keynesian Phillips Curve (NKPC) [see, for example, Dufour, Khalaf and Kichian (2006, 2010), Ma (2002), Mavroeidis (2004, 2005), Nason and Smith (2008), Kleibergen and Mavroeidis (2009)]; Taylor-type monetary policy rules [Mavroeidis (2010), Inoue and Rossi (2011)]; and the Euler equation for output [Fuhrer and Rudebusch (2004), Magnusson and Mavroeidis (2010)]. For multi-equation models, several studies have explored identification difficulties, the proper recovery of macroeconomic dynamics from structural VARs, and the role of added measurement errors; see Kim (2003), Beyer and Farmer (2007), Fernandez-Villaverde, Rubio-Ramirez, Sargent and Watson (2007), Ruge-Murcia (2007), Canova and Sala (2009), Chari, Kehoe and McGrattan (2009), Consolo, Favero and Paccagnini (2009), Chevillon, Massmann and Mavroeidis (2010), Iskrev (2010), Magnusson and Mavroeidis (2010), Moon and Shorfheide (2010), Cochrane (2011), Komunjer and Ng (2011), Andrews and Mikusheva (2011), and Granziera, Lee, Moon and Schorfheide (2011).

Macroeconomists are rarely dogmatic in favour of a fully-specified model as an end in itself. Rather, models are viewed mainly as quantitative benchmarks for the evaluation of substantive economic issues. While there is a consensus that certain models are useful for this purpose, there is less agreement on how such models should be parameterized when taken to the data. Ideally, one would like to focus on implications of interest conforming with micro-founded structures while allowing the data to *speak freely* on the dimensions along which these may lack fit. In particular, the following features can affect identification and inference validity. First, an important challenge consists in minimizing the effects of auxiliary assumptions. For instance, innovations arising from measurement errors are usually non-fundamental. Alternatively, the existence of a unique rational expectation solution may challenge theory [see Cochrane (2011)]. Second, DSGE-VAR methods broadly assess the structural form against an unrestricted VAR where the included variables are determined by the DSGE. The literature is witnessing a growing awareness on the possibility of misspecifying the benchmark and its consequences. Variable omission is a third recognized difficulty, since by construction and because of their specificity, DSGE models may exclude empirically relevant data. For all these reasons, the consequences of spuriously completing models should be taken into account.

This paper proposes *identification-robust* inference methods, *i.e.* methods which are valid whether identification is weak or strong, for DSGE setups. For definitions and surveys of the relevant econometric literature, see, for example, Stock, Wright and Yogo (2002), Dufour (2003), and Kleibergen and Mavroeidis (2009). Despite the considerable associated econometric literature, identification-robust methods for multi-equation systems are still scarce [see Moon and Shorfheide (2010), Granziera, Lee, Moon and Schorfheide (2011), Guerron-Quintana, Inoue and Kilian (2009), Magnusson and Mavroeidis (2010) and Andrews and Mikusheva (2011)]. We introduce two system-based identification-robust methods which can address either all of the restrictions implied by the model ["full-information" inference], or only some of those restrictions ["limited-information" inference]. So the latter approach (implicitly) considers a more general setup, though it retains basic features of the original model. We argue these approaches should be viewed as complementary, rather than mutually exclusive. Comparing LI with FI inference provides a useful specification check, and our incomplete-model alternative allows the researcher to draw inferences which are more robust to auxiliary model assumptions (such as the information used by economic agents to form their expectations). Both methods rely on estimation and test procedures whose statistical validity is not affected by identification issues and questionable auxiliary assumptions.

We apply these tools to an illustrative three-equation New Keynesian model, estimated from U.S. data. This fundamental structure has been extensively studied and forms the building block of many other more complex models; see Clarida, Gali and Gertler (1999), Woodford (2003), Christiano, Eichenbaum and Evans (2005), Linde (2005), Benati (2008), Del Negro, Schorfheide, Smets and Wouters (2007), to mention a few. Three features of the New Keynesian model are addressed. First, inflation persistence is studied within the NKPC, given the on-going debate in this regard [see the survey by Schorfheide (2008)]. Second, the output gap coefficient in the NKPC and the real interest parameter in the output equation are analyzed, as currently available results lead to conflicting conclusions on the impact of these variables [see Schorfheide (2010)]. For clarity, these are called the *forcing variables* of the corresponding equations. Third, the implications of imposing a unique rational expectation solution on the feedback coefficients in the Taylor rule are revisited, in light of serious issues arising from determinacy underscored, for example, by Mavroeidis (2010) and Cochrane (2011). Comparisons between our FI and LI assessments of these questions are discussed.

Our findings can be summarized as follows. When a stable and unique equilibrium is imposed to complete the model, it is rejected by the data. This is an important sense in which our analysis can be seen as an exploration of the pervasiveness of auxiliary FI assumptions. In contrast, although insignificant forcing variables in the NKPC and the output curve cannot be ruled out, our LI multi-equation results provide realistic conclusions on the nature of the NKPC, and yield precise estimates of feedback coefficients which appear consistent with the Taylor principle. It is shown that such conclusions cannot be reached via single-equation methods. These results indicate that a multi-equation estimation of the considered model can still utilize the information in the contemporaneous relationship between output, inflation, and interest rates, which positively affects identification and inference.

In section 2, our framework and empirical model are described. Section 3 presents the methodology. Empirical results are provided in section 4. Section 5 offers some conclusions.

2. Framework

Consider the general structural form

$$\Gamma_0 X_t = \Gamma_1 X_{t-1} + C + \omega \nu_t + \psi \eta_t \tag{1}$$

where X_t is vector of m^* variables, C is a vector of constants, ν_t is an exogenous shock, and η_t is a vector of expectation errors such that $E_t(\eta_{t+1}) = 0$. Collect all the parameters of (1) in the vector ϑ. Typically, only a vector [denoted Y_t] of n^* components of X_t is observable. Time-t expectations for some of the variables may also be included in X_t. Using standard techniques [see Anderson and Moore (1985), King and Watson (1998), Sims (2002), Anderson (2008)] and appropriate restrictions on ϑ [denoted as $\vartheta \subset \Theta$], (1) can be solved into $X_t = C_0 + C_1 X_{t-1} + G\nu_t$, where C_0, C_1 and G are functions of ϑ. If the model has no solution for a given parameter value $\vartheta = \vartheta_0$, this means ϑ_0 is not consistent with a rational-expectation model where the information set is restricted to information variables included in the model; so we can determine whether any given value ϑ_0 is admissible (for such a model) by checking whether a solution does exist. Focusing on a unique stable rational expectation solution, Y_t is an infinite VAR that can be approximated [see Fernandez-Villaverde, Rubio-Ramirez, Sargent and Watson (2007) and Ravenna (2007)] by the restricted form

$$\mathrm{Y}_t = B_0(\vartheta) + B_1(\vartheta)\,\mathrm{Y}_{t-1} + \cdots + B_p(\vartheta)\,\mathrm{Y}_{t-p} + \Sigma(\vartheta)\,u_t, \quad u_t \sim N(0,\, I_{n^*}) \,, \tag{2}$$

where $B_0(\vartheta), \ldots, B_p(\vartheta)$ are constructed by truncation or by population regression as in Del Negro, Schorfheide, Smets and Wouters (2007). Most standard DSGE models are covered by (1). Special cases may also admit finite-order VAR representations for which (2) holds.

The proposed partial specification is analogous to GMM. Define $\epsilon_{it}(\mathrm{Y}, \theta)$, $i = 1, \ldots, n$, where Y denotes observable data on endogenous and exogenous variables and θ [which may be equal to ϑ, a subset of ϑ, or some transformation of the latter] the parameters of interest, such that if (1) holds then $\epsilon_{it}(\mathrm{Y}, \theta)$ is orthogonal to a vector of k_i instruments Z_{it} at the true θ. Collecting all different variables from each of the Z_{it} into a k-dimensional vector Z_t so that $Z_{it} = A_i Z_t$ where A_i is a $k_i \times k$ selection matrix, we propose to map the n orthogonality conditions into estimating and testing the multivariate regression of $\epsilon_{it}(\mathrm{Y}, \theta)$ on Z_t with *i.i.d.* or serially dependent V_t :

$$\epsilon_t(\mathrm{Y}, \theta) = \Pi Z_t + \mathrm{V}_t, \quad \epsilon_t(\mathrm{Y}, \theta) = (\epsilon_{1t}(\mathrm{Y}, \theta), \ldots, \epsilon_{nt}(\mathrm{Y}, \theta))' . \tag{3}$$

That is, at the true θ and letting Π_i' refer to the ith row of Π, we have:

$$A_i \Pi_i = 0. \tag{4}$$

Our empirical analysis focuses on a prototypical New Keynesian model [see Clarida, Gali and Gertler (1999) and Linde (2005)] where, for $t = 1, \ ..., \ T$:

$$\pi_t = \omega_f E_t \pi_{t+1} + (1 - \omega_f)\,\pi_{t-1} + \gamma y_t + \varepsilon_{\pi t} \,, \tag{5}$$

$$y_t = \beta_f E_t y_{t+1} + \sum_{j=1}^{4} \left(1 - \beta_f\right) \beta_{yj} y_{t-j} - \beta_r^{-1} (R_t - E_t \pi_{t+1}) + \varepsilon_{yt} \,, \tag{6}$$

$$R_t = \gamma_\pi \left(1 - \sum_{j=1}^{3} \rho_j\right) \pi_t + \gamma_y \left(1 - \sum_{j=1}^{3} \rho_j\right) y_t + \sum_{j=1}^{3} \rho_j R_{t-j} + \varepsilon_{Rt} \,, \tag{7}$$

π_t is aggregate inflation, y_t is the output gap, R_t is the nominal interest rate, and the disturbance vectors $\varepsilon_t = (\varepsilon_{\pi t}, \varepsilon_{yt}, \varepsilon_{Rt})'$ have zero-mean with covariance matrix Ω. Parameters γ and β_r^{-1} are referred to as the coefficients on the forcing variable in the NKPC and the output equation, respectively. Let

$$\theta = \left(\omega_f,\, \gamma,\, \beta_f,\, \beta_r,\, \gamma_\pi,\, \gamma_y, \rho_1,\, \rho_2,\, \rho_3\right)', \quad \phi = \left(\theta',\, \beta_{y1},\, \beta_{y2},\, \beta_{y3},\, \beta_{y4}\right)' \tag{8}$$

refer to model "deep" parameters, and let Θ and Φ denote the associated parameter spaces.

Our FI method assumes $(\varepsilon_{\pi t}, \varepsilon_{yt},\, \varepsilon_{Rt})' \overset{iid}{\sim} N(0, \Omega)$ with Ω invertible. Model (5)-(7) may be represented as in (1) by replacing expectation variables with actual values plus errors, and then solving forward into (2) with $\mathrm{Y}_t = (\pi_t, y_t,\, R_t)'$, $p = 4$, $B_1(\vartheta) = B_1(\phi), \ldots, B_p(\vartheta) = B_p(\phi)$ and $\Sigma(\vartheta) = \Sigma(\phi,\, \Omega)$.

From a LI perspective, (5)-(7) imply that $\epsilon_t(\mathrm{Y},\, \theta) = (\epsilon_{\pi t}(\mathrm{Y},\, \theta), \epsilon_{yt}(\mathrm{Y},\, \theta), \epsilon_{Rt}(\mathrm{Y},\, \theta))'$ is uncorrelated with available instruments for the true θ, where Y stacks Y_t, $t = 1, \ldots, T$ and

$$\epsilon_{\pi t}(\mathrm{Y},\, \theta) = \pi_t - \omega_f \pi_{t+1} - (1 - \omega_f)\pi_{t-1} - \gamma y_t\,, \tag{9}$$

$$\epsilon_{yt}(\mathrm{Y},\, \theta) = y_t - \beta_f y_{t+1} + \beta_r^{-1}\left(R_t - \pi_{t+1}\right)\,, \tag{10}$$

$$\epsilon_{Rt}(\mathrm{Y},\, \theta) = R_t - \left(1 - \sum_{j=1}^{3} \rho_j\right)\left(\gamma_\pi \pi_t + \gamma_y y_t\right) - \sum_{j=1}^{3} \rho_j R_{t-j}\,. \tag{11}$$

The predetermined variables in the system, denoted the "intra-model" instruments are:

$$\dot{Z}_t = (\pi_{t-1},\, R_{t-1},\, R_{t-2},\, R_{t-3})',\ \ddot{Z}_t = (y_{t-1},\, y_{t-2},\, y_{t-3},\, y_{t-4})'\,. \tag{12}$$

The information set can also be expanded using (lags of) non-modelled variables as extra instruments. These are grouped in a vector denoted $\tilde{Z}_t$.

3. Methodology

Whereas traditional estimation methodology involves first finding a point estimate and its distribution, from which confidence intervals and tests are then built, we proceed here in the reverse order: we start from a test procedure for different parameter values, build confidence regions from these, and finally get a point estimate. A confidence region with level $1-\alpha$ is obtained by "inverting" an LI or FI identification-robust test with level α, *i.e.* a test whose level α is controlled without identification conditions. A test is "inverted" by finding the set of parameter values which are not rejected by the test procedure. Depending on circumstances, this can be done either analytically or numerically. A point estimate may then be obtained by picking from the confidence region the parameter value associated with the largest test p-value. The confidence set covers the parameter of interest with the stated probability (at least $1-\alpha$). Under identification failure (or weak identification), the set will be noninformative (possibly unbounded), as it should be [see Dufour (1997)]. The set can also be empty, which implies that the structural model is rejected at the considered level.

3.1. Structural limited-information method

In the context of (1)-(4), consider $H_{01}: \theta = \theta_0$ (θ_0 known) and the artificial regression

$$\epsilon_t(\mathrm{Y}, \theta_0) = \Pi Z_t + V_t \Leftrightarrow \epsilon(\theta_0) = zb + v, \quad z = (I_n \otimes Z), \tag{13}$$

where Z is the $T \times k$ matrix of instruments with t-th row equal to Z_t', v is the nT-dimensional vector that stacks V_t, $t = 1, \ldots, T$, $b = vec(\Pi')$ and $\epsilon(\theta_0)$ is the nT-dimensional vector of structural errors evaluated at θ_0. We propose to assess $H_{01}^*: A_i\Pi_i = 0 \Leftrightarrow Ab = 0$ where $\epsilon_t(\mathrm{Y}, \theta)$, Π, A_i and A conform with (3)-(4). A is the selection matrix with rank m (the total number of tested coefficients) that imposes (4). Indeed, if H_{01} is true, H_{01}^* should hold; in other words, if θ_0 represents the true parameter vector, then additional information from predetermined variables should be irrelevant.

Because the right-hand-side regressors in (13) are not 'endogenous', standard statistics can be applied. A multivariate statistic is used whose approximate [assuming homoskedasticity] null distribution is $F(m, n(T-k))$, even with weak instruments:

$$\mathcal{W}(\theta_0) = \frac{n(T-k)}{m} \frac{\left(A\hat{b}(\theta_0)\right)' \left[A\left(\hat{\Sigma}_V(\theta_0) \otimes (Z'Z)^{-1}\right)A'\right]^{-1} \left(A\hat{b}(\theta_0)\right)}{\left(\epsilon(\theta_0) - z\hat{b}(\theta_0)\right)' \left(\hat{\Sigma}_V^{-1}(\theta_0) \otimes I_n\right) \left(\epsilon(\theta_0) - z\hat{b}(\theta_0)\right)} \tag{14}$$

where $\hat{b}(\theta_0)$ and $\hat{\Sigma}_V(\theta_0)$ are the unrestricted OLS and covariance estimators from (13).

The test inversion itself must be conducted numerically. One can sweep, for example, economically meaningful choices for θ_0, and for each choice considered, compute $\mathcal{W}(\theta_0)$ and its p-value. The parameter vectors for which the p-values are greater than the level α constitute the identification-robust confidence region with level $1-\alpha$. Confidence sets for each individual component of θ may then obtained by *projection*, i.e. by finding all values of the relevant component for which at least one vector in the confidence set has this component value. Projection-based confidence intervals (which may be wider than the previous sets) can also be built by finding the smallest and largest values of each parameter in the confidence region. More generally, for any (scalar) function $g(\theta)$, a valid confidence interval (with level $1-\alpha$) can be obtained by minimizing and maximizing $g(\theta)$ over the θ values included in the joint confidence region. Each component of θ is defined as a linear combination of θ, *i.e.* $g(\theta) = a'\theta$, where a is a conformable selection vector (consisting of zeros and ones); for example, $\omega_f = (1, 0, \ldots, 0)\,\theta$. The associated $a'\theta$ function is then optimized [using Simulated Annealing] over θ such that $\mathcal{W}(\theta) < F_\alpha(m, n(T-k))$, where $F_\alpha(.)$ denotes the α-level cut-off point. For further discussion of the projection method, see Abdelkhalek and Dufour (1998) and Dufour and Taamouti (2005).

Observe that $\left[\overline{\mathcal{W}} = \min_{\theta_0} \mathcal{W}(\theta_0)\right] \geq F_\alpha(m, n(T-k)) \Leftrightarrow \mathcal{W}(\theta_0) \geq F_\alpha(m, n(T-k)), \forall\theta_0$. So comparing $\overline{\mathcal{W}}$ to $F_\alpha(m, n(T-k))$ provides an identification-robust specification (J-type) test at level α. If $\overline{\mathcal{W}} < F_\alpha(m, n(T-k))$, the associated confidence set is not empty.

Expectation errors may entail serial dependence for V_t in (13) [Mavroeidis (2004)]. Using MacKinlay and Richardson (1991), a HAC criterion is thus considered:

$$\mathcal{J}(\theta_0) = T\,\hat{d}'D'\left[D\left(\left(\frac{Z'Z}{T}\right)^{-1} \otimes I_n\right) S_T \left(\left(\frac{Z'Z}{T}\right)^{-1} \otimes I_n\right) D'\right]^{-1} D\hat{d}\,, \tag{15}$$

$$S_T = \Psi_{0T} + \sum_{j=1}^{l} \left(\frac{l-j}{l} \right) \left[\Psi_{jT} + \Psi_{jT}^{'} \right], \quad \Psi_{jT} = \frac{1}{T} \sum_{t=j+1}^{T} \left(Z_t \otimes \hat{v}_t \right) \left(Z_{t-j} \otimes \hat{v}_{t-j} \right)', \quad (16)$$

where $\hat{d}$ corresponds to $\hat{b}$ reshaped so that $d = vec(\Pi)$, D corresponds to a conformable reshaping of A (so that $Ab = 0$ if and only if $Dd = 0$), and $\hat{v}_t$ is the unrestricted OLS residual from (13). We report results for $l = 4$. Its approximate null distribution is $\chi^2(m)$, even with weak instruments. The confidence region based on inverting $\mathcal{J}(\theta_0)$ can also be empty [when $\bar{\mathcal{J}} = \min_{\theta_0} \mathcal{J}(\theta_0) \geq \chi^2_\alpha(m)$] or unbounded. $\mathcal{J}(\theta_0)$ and the continuously updated GMM-type objective function [Stock and Wright (2000)] are asymptotically equivalent given certain regularity conditions. Kleibergen and Mavroeidis (2009) recommend, under specific assumptions, a χ^2 null distribution with degrees-of-freedoms reduced to m minus the number of parameters tested (here, the dimension of θ). However, recent studies [see Ray and Savin (2008), Gungor and Luger (2009) and Beaulieu, Dufour and Khalaf (2010)] show that system-based HAC criteria may severely over-reject in finite samples.

For model (5)-(7), with $Z_t = (\dot{Z}_t', \ddot{Z}_t')'$ as in (12), then A_1 and A_3 should select all coefficients of the first and third equation of (13), whereas A_2 should select the coefficients in the second equation associated with $\dot{Z}_t$. Indeed, the coefficients on the output gap lags are free in the output equation, and thus the exclusion of these coefficients is not tested within the second equation. If $Z_t = (\dot{Z}_t', \ddot{Z}_t', \tilde{Z}_t')'$, the A_i matrices should also select the coefficients of $\tilde{Z}_t$ for all the equations.

3.2. Full-information method

The method described above does not use all the restrictions entailed by a "closed" rational-expectation model where the information set is restricted to variables included in the model, so that an explicit solution can be derived. Clearly, if we accept such restrictions, tighter inference can be achieved. We call the method which takes into account these restrictions the FI method, as opposed to the previous LI method. Given (1)-(2), $H_{02} : \vartheta = \vartheta_0$, where ϑ_0 is known but restricted so that an associated rational expectation solution exists, can be tested by assessing $H_{02}^{*} : \Pi = 0$ within the artificial VAR:

$$U_t\left(\mathrm{Y}, \vartheta_0\right) = \Pi Z_t + W_t, \tag{17}$$

where $U_t\left(\mathrm{Y}, \vartheta_0\right) = \mathrm{Y}_t - B_0\left(\vartheta_0\right) - B_1\left(\vartheta_0\right)\mathrm{Y}_{t-1} - \cdots - B_p\left(\vartheta_0\right)\mathrm{Y}_{t-p}$, and Z_t includes as many lags of each component of Y_t. When (2) holds exactly and ϑ can be partitioned as $\vartheta = (\phi', \bar{\phi}')'$ so that $B_1(.), \ldots, B_p(.)$ depend on ϕ but not on $\bar{\phi}$, (2) can be written $\mathrm{Y}_t = B_0(\phi) + B_1(\phi)\mathrm{Y}_{t-1} + \cdots + B_p(\phi)\mathrm{Y}_{t-p} + \Sigma\left(\phi, \bar{\phi}\right) u_t$, and we can focus on the partial hypothesis $H_{02} : \phi = \phi_0$. This leads to the artificial VAR:

$$U_t\left(\mathrm{Y}, \phi_0\right) = \Pi Z_t + W_t\,, \tag{18}$$

where $U_t\left(\mathrm{Y}, \phi_0\right) = \mathrm{Y}_t - B_0\left(\phi_0\right) - B_1\left(\phi_0\right)\mathrm{Y}_{t-1} - \cdots - B_p\left(\phi_0\right)\mathrm{Y}_{t-p}$. This applies to model (5)-(7), with ϕ as in (8), and $Z_t = (\dot{Z}_t', \ddot{Z}_t')'$ as in (12): so long as $U_t\left(\mathrm{Y}, \phi_0\right)$ exists, testing for H_{02}^{*} within (18) provides a test of H_{02}. For this purpose, the LR-type multivariate statistic [Dufour and Khalaf (2002)] is used:

$$\mathcal{L}\left(\phi_0\right) = \left(\frac{\mu\tau - 2\lambda}{Kn}\right) \frac{1 - \left(|\hat{\Sigma}_W\left(\phi_0\right)|/|\hat{\Sigma}_W^0\left(\phi_0\right)|\right)^{1/\tau}}{\left(|\hat{\Sigma}_W\left(\phi_0\right)|/|\hat{\Sigma}_W^0\left(\phi_0\right)|\right)^{1/\tau}}\,, \tag{19}$$

$$\mu = (T - K) - \frac{(n^{*} - K + 1)}{2}, \quad \lambda = \frac{n^{*}K - 2}{4}\,, \tag{20}$$

$$\tau = \begin{cases} [(K^2 n^{*2} - 4)/(K^2 + n^{*2} - 5)]^{1/2} & ,\ if \quad K^2 + n^{*2} - 5 > 0\,, \\ 1 & ,\ \text{otherwise}, \end{cases} \tag{21}$$

where n^{*} is the dimension of Y_t, K is the dimension of Z_t, $\hat{\Sigma}_W^0\left(\phi_0\right)$ and $\hat{\Sigma}_W\left(\phi_0\right)$ give the constrained [imposing $\Pi = 0$] and unconstrained sum of squared errors matrices from (18). $\mathcal{L}\left(\phi_0\right)$ has an approximate [imposing homoskedasticity] $F(Kn^{*}, \mu\tau - 2\lambda)$ null distribution even if identification is weak. In model (5)-(7), $n^{*} = 3$, $Z_t = (\dot{Z}_t', \ddot{Z}_t')'$ so $K = 8$.

The test inversion procedure is similar to that presented in the previous section, using, for example, a grid search over the economically meaningful values for ϕ, or by projection-based methods. Choices for ϕ are restricted to ensure that the above-defined $U_t\,(\mathrm{Y}, \phi_0)$ exist. With model (5)-(7), for every candidate ϕ_0 value, the usual existence conditions are checked using the Anderson and Moore algorithm. The FI confidence region thus admits the possibility of being unbounded or empty [when $\bar{\mathcal{L}} = \min_{\phi_0} \mathcal{L}\,(\phi_0) \geq F_\alpha(Kn\,,\, \mu\tau - 2\lambda)$].

When (2) is an approximation, the VAR residuals may not be *i.i.d.* The magnitude of the discrepancy decreases with large p, though a HAC version of the test statistic could be used for each parameter value tested for building the confidence set.

4. Empirical analysis

We study model (5)-(7) as a well-known example of general structures consistent with the literature. This analysis is viewed as illustrative in various respects. First, (5)-(7) includes lags in the output and interest rate equations that are not strictly derived from New Keynesian foundations. Completing a New Keynesian model requires non-theory-based choices, for example, the inclusion of auxiliary shocks or measurement errors, and assumptions about the law of motion of the shocks. Several reasonable options can be considered for this purpose, but none emerges as the ideal choice. Adding lags to justify an *i.i.d.* assumption on ε_t follows Linde (2005). Second, model (5)-(7) imposes no cross-equation restrictions on regression parameters. Since existing work provides no consensus view in this regard, our specification suggests a minimal set of assumptions for estimation purposes.

Third, model (5)-(7) is a special case of (1) in which the number of structural shocks is equal to the number of endogenous variables. Its solution has the form (2) with $B_1\,(\vartheta) = B_1\,(\phi)\,, \ldots,\; B_p\,(\vartheta) = B_p\,(\phi)$, *i.e.* $B_1\,(.)\,, \ldots,\; B_p\,(.)$ depends on ϕ, the model deep parameters defined in (8), but does not depend on Ω. This allows one to conveniently partial Ω out in estimation. The solution also imposes exclusion restrictions on $B_1\,(.)\,, \ldots,\; B_p\,(.)$ so although $p = 4$, the solved model in fact includes four lags of y_t, three lags of R_t and only one lag of π_t. The same exclusion restrictions are imposed on the unrestricted benchmark VAR considered.

Fourth, again conforming with the above-cited literature, the solution that is empirically maintained rules out sunspot equilibria. Our closed-model approach follows the usual practice of restricting parameter values, so that a unique rational expectation solution exists. This can be quite restrictive [see, for example, King (2000) and Cochrane (2011)] and needs to be pointed out as it may suggest an important interpretation to an eventual model rejection.

Finally, one of the criticisms routinely advanced against the considered model is that its parsimony implies a limited information set that may lack credibility. The intervening variables are the output gap, inflation, and a short-term interest rate, which implies that lags of these variables should suffice to adequately capture monetary policy. For modern economies, this is counterfactual. A more flexible setup would allow additional information, reflecting the data-rich environment within which policy makers operate. One way to link equilibrium founded structures with relevant aggregates that are not explicitly modeled is to consider additional instruments, referred to as "extra-model instruments". As an illustrative example of external instruments, lags 2 and 3 of both wage and commodity price inflation, are considered, conformable with the literature.

On balance, our illustrative framework does not depart from common practice: although reasonable and substantiated in published empirical works, our assumptions remain strict and will serve to illustrate the ability of our proposed methods to reject false models.

Applications are conducted using U.S. quarterly data for the sample extending from 1962Q1 to 2005Q3. The GDP deflator is used for the price level (P_t) and the Federal Funds rate is used as the short-run interest rate. For the output gap, two measures are considered: one is a real-time measure of the output gap, in the sense that the gap value at time t does not use information beyond that date. This ensures that the lags of the output gap are valid for use as instruments. Thus, as in Dufour, Khalaf and Kichian (2006, 2010), one proceeds iteratively: to obtain the value of the gap at time t, GDP is detrended with data ending in

t. The sample is then extended by one observation and the trend is re-estimated. The latter is used to detrend GDP, yielding a value for the gap at time $t+1$. This process is repeated until the end of the sample. A quadratic trend is used for this purpose. A second measure is the standard quadratically-detrended output gap (that uses the full sample) as in Linde (2005), and which is included for comparison purposes. The log of both these output gap series are taken. Finally, as in Linde (2005), all the data is demeaned prior to estimation.

4.1. Substantive questions

Three features of the New Keynesian model are assessed using (5)-(7). First, intrinsic inflation persistence within the NKPC is studied. Formally, we test whether values of ω_f less than 0.5 can be ruled out, *i.e.* whether the NKPC is conclusively more forward-looking than backward-looking. The purely forward-looking case [*i.e.* $\omega_f = 1$] is also studied, to see if it can be refuted. For insights and perspectives on the importance of lagged inflation, see Linde (2005), Benati (2008), Fair (2008), Nason and Smith (2008), Schorfheide (2008) and the references therein. The question is asked whether our system approach can sharpen our inference on the nature of the NKPC relative to single-equation methods.

Second, the estimated coefficients on the forcing variables in the NKPC and the output equation are examined. Formally, tests are conducted to see if the hypotheses $\gamma = 0$ and , $\beta_r^{-1} = 0$ can be conclusively refuted. As emphasized in Schorfheide (2010), reported estimates of forcing variables coefficients [specifically of the NKPC] are "fragile" across available studies and cover [among others] values near zero, implying that changes in demand pressures have no impact on inflation. In contrast with single-equation models, system-based estimation utilizes the information in the contemporaneous relationship between output, inflation, and interest rates, which may better capture the parameters describing transmission mechanisms. We thus ask whether more realistic predictions are achieved by our system approach relative to single-equation methods.

Third, the extent to which a system approach can recover any useful information on the feedback coefficients in the Taylor rule (γ_π and γ_y) is verified. Mavroeidis (2010) reports identification problems for these parameters from a single-equation perspective. Fundamental issues with such rules - arising from imposing unique rational expectation solutions when New Keynesian type models are brought to data - have recently been pointed out by Cochrane (2011). Although a sole reliance on γ_π and γ_y to interpret such issues can be misleading, Cochrane (2011) provides a motivation for assessing the worth of system-based inference on the Taylor rule, which suggests to check whether imposing stability on the considered system has any empirical support.

4.2. Results and discussion

In what follows and unless otherwise indicated, a 5% significance level is applied in the discussion of inference outcomes. Our system inference produces a striking result. With both gap measures, the model is rejected using the FI method. The model is also rejected for both gap measures using the multi-equation HAC statistic, with and without external instruments. With our multi-equation *i.i.d.* LI method, an empty confidence set is obtained when the standard quadratically-detrended output gap is used. In contrast, the model is not rejected with the real-time output measure of the gap using this same statistic.

It is worth comparing these results with those of Linde (2005). Using Monte Carlo experiments, Linde argues that FIML methods are superior to GMM-type approaches for inference on the structural parameters, and his estimations show that the NKPC is preponderantly backward-looking. Using either one of two different measures for y_t yields qualitatively similar results. Our finding is that FI actually leads us to reject the model, and furthermore, that the proxies used for the gap have a strong impact on LI estimations. This last conclusion also contrasts with Kleibergen and Mavroeidis (2009) who report that their (single-equation) estimates of the NKPC are empirically invariant to the gap measure.

One possible reason for why we obtain conflicting outcomes with the different gap measures using the system LI statistic is the instrument validity problem discussed in Doko-Tchatoka and Dufour (2008) and Dufour, Khalaf and Kichian (2010): given that the standard output gap measure is obtained using all of the sample observations, its lags may be correlated with time t errors. More subtle arguments can be raised on the validity of lags as instruments. For example, in the context of the New Keynesian model, Cochrane (2011) argues that the interaction of assumptions on disturbances with assumptions for determinacy may make lags of endogenous variables inappropriate for use as predetermined regressors. With regard to assumptions on disturbances, our model passes the LI test imposing *i.i.d.* regression errors and fails when serial dependence is allowed. This observation should be qualified: (i) spurious rejections may easily occur, for it is well known that asymptotic system HAC-based tests tend to be oversized; (ii) the lag structure adopted to justify *i.i.d.* errors is quite restrictive. Results without HAC remain more restrictive in the sense that they rule out MA errors and heteroskedasticity, so there is a trade-off between robustness and finite-sample accuracy. Perhaps more to the point is our model rejection with the FI statistic, because FI is based on a solution that imposes more than just model consistency: it imposes determinacy as well, and that may be an important factor driving the rejection.

We do not claim that we formally test determinacy. Our FI rejection may also be linked to the usual culprits, *i.e.* it, may have more to do with unsuitable exogenous driving processes than with the credibility of the New Keynesian model itself. Although related with regard to their econometric implications on regression errors, the problems arising from ill-fitted shock processes and determinacy are fundamentally different. One can always add lag-length restrictions as approximations, yet a unique and stable rational expectation solution may require stronger assumptions. The LI method is applied maintaining the same lag-length restrictions on disturbances as the FI one; in contrast with FI, this gives the model a chance of passing the tests, which is interesting to notice.

One may object at any further analysis based on the considered structure when its underlying equilibrium restrictions are empirically unsubstantiated. There is an active debate on the right specification of the New Keynesian model, so despite a rejection with FI, we proceed with our interpretation of (5)-(7) as an incomplete structure. Table 1 reports parameter estimates and associated identification-robust projections for the elements of θ, for the cases where the model is not rejected, *i.e.* with the real-time gap, and using our LI method. Tables 2 and 3 report confidence intervals based on single-equation identification-robust methods [used in Dufour, Khalaf and Kichian (2006, 2010) and Mavroeidis (2010)] which impose the structural constraints of each equation [Table 2], as well as the completely unconstrained method proposed by Dufour and Taamouti (2005) [Table 3].

The point estimates in Table 1 appear compatible with the literature on models estimated using the real-time gap. In particular, the coefficient on the expected inflation term of the NKPC is high, indicating forward-looking behavior. This conclusion was also reached by Gali, Gertler and Lopez-Salido (2005), Sbordone (2005), and Smets and Wouters (2007). Similarly, the coefficients of the Taylor Rule are not far from the numbers that Taylor (1993) had suggested and what other studies have found for the post-Volcker era (see, for instance, Clarida, Gali and Gertler (2000)).

Point estimates do not change much whether the full instrument set or the model-consistent instrument subset are used. However, outcomes are subtly different when the sensitivity of the confidence intervals is assessed relative to the information set. In particular, it can be ascertained that the NKPC is forward-looking irrespective of the instrument set [values of ω_f below 0.5 are rejected], whereas values near one for ω_f and less than one for γ_π are ruled out by the full set of instruments but cannot be rejected by the model-consistent instrument subset. This observation also holds when multi-equation estimates are contrasted with single-equation ones, especially for ω_f.

The confidence intervals in Table 2 suggest that, when cross-equation information is not accounted for, the model-consistent instruments are weakly informative on the NKPC relative to the expanded instrument set. While the confidence intervals for ω_f become much tighter when the instruments set is expanded and the *i.i.d.* assumption is relaxed, in contrast with our multi-equation based results, the pure forward-looking case [$\omega_f = 1$] cannot be rejected. It is also worth noting from Table 3 that the unrestricted confidence intervals for the forward-looking coefficient in the NKPC, treated as a reduced form [*i.e.* when the restriction that the forward and backward-looking terms sum up to one is relaxed], covers values exceeding one. Values less than 0.5 cannot be ruled out by the single-equation results, except with the HAC statistic and the standard gap measure (with which we rejected the model from a system-based LI perspective). Aside from this exception, single-equation confidence intervals on the NKPC are much more sensitive to changes in the information set than to changes in the gap measure.

Again from Table 2, single-equation estimation of the output equation produces empty sets whether structural restrictions are imposed or not, whether the *i.i.d.* assumption on errors is imposed or not, and with both gap measures. The Taylor rule is rejected under all single-equation assumptions with the standard gap measure. With the real-time gap measure, support for the rule seems more fragile, in the sense that results vary dramatically with different instruments and assumptions. For example, with model consistent instruments, confidence intervals for γ_π and γ_y imposing and relaxing *i.i.d.* disturbances are wide suggesting the same identification difficulties documented by Mavroeidis (2010). In contrast, expanding the instrument sets leads to rejecting the equation except with an *i.i.d.* disturbance, in which case we again find wide confidence sets revealing weak identification.

Focusing on the results with the real-time output measure, two points deserve notice when single-equation evidence is contrasted with our LI multi-equation inference. First, despite their imperfections when considered on their own as single equations, including both output and interest rate equations in the system sharpens our inference on the NKPC. In contrast with single-equation methods, system-based estimation reveals useful information on ω_f: the NKPC is conclusively more forward than backward-looking, and the pure forward-looking case can be ruled out. Such a conclusion cannot be reached with a single-equation approach.

Second, our LI system inference is quite informative for the interest rate equation. In contrast with the high estimate uncertainty that was found with our single-equation approaches, LI system-based confidence intervals on γ_π and γ_y are tightly centered around values compatible with Taylor (1993), especially particularly when an expanded instrument set is used. We do not claim that ruling out estimation uncertainty on γ_π and γ_y evacuates the deep interpretation issues [see King (2000) and Cochrane (2011)] associated with these parameters within a New Keynesian reaction function. Nevertheless, our LI method allows cross-equation variables to interact contemporaneously with minimal assumptions on the underlying dynamics, which delivers precise estimates of feedback coefficients which appear compatible with the Taylor principle. Such a conclusion, again, cannot be reached with a single-equation approach.

Another important observation is the insignificance [when the model is not rejected] at the 5% level of the parameter on the output gap coefficient in the NKPC. The value of the parameter on the real interest rate in the output equation is also found to be quite small, often hitting the lower bound of 0.03 (more precisely, the elasticity of intertemporal substitution hits the maximal value of 30.00 allowed in the estimation). This confirms the findings of Rudd and Whelan (2006) and Benati (2008).

While not uncommon, insignificant forcing variables in the NKPC and IS are an empirical puzzle. So far, available identification-robust evidence on this issue is restricted to the NKPC. Kleibergen and Mavroeidis (2009) apply partialled-out single-equation tests, which under specific conditions [for example, assuming that all relevant instruments are used] may provide more powerful tests than projection-based methods, and yet cannot rule out a flat NKPC even with such statistics. We are however not sure of the appropriateness of these statistics given our sample size; it can be verified that the simulation study reported by Kleibergen and Mavroeidis (2009) uses a sample size of 1000 observations. Magnusson and Mavroeidis (2010) also confirm this finding using the labor share, and an identification-robust minimum distance estimation method based on a reduced-form VAR process for π_t and y_t. This result is noteworthy because the authors document, through an empirically relevant simulation study, that their reliance on an underlying VAR provides more powerful inference than standard single-equation GMM, which still does not address the puzzle. Our study adds credible structure to such a multi-equation analysis and yields a similar outcome. The same puzzle is also found to plague the IS equation.

While all issues raised by Schorfheide (2010) can drive such findings, two possible interpretations may be suggested. First, it is indeed the case that the NKPC and the IS equations are flat with respect to the forcing variables, which is a dilemma that challenges theory. Second, the model transmission mechanism is incomplete or misspecified, so forward and backward-looking terms in the NKPC and the IS curve still absorb all information in the data, even when the modeled variables are allowed to interact contemporaneously across equations. Using single-equation methods, no empirical support was found for the output equation and there was very weak support of the interest rate equation, which lends credibility to the second interpretation. Our FI test suggests that the overall empirical model lacks support, which may be a plausible - although radical - resolution of this dilemma.

Our results can be summarized as follows. Results with FI are negative, establishing that one popular empirical specification lacks support. In contrast, as our LI results suggest, there is still sufficient statistical information in the sample to learn something useful on the nature of the NKPC, as well as the feedback terms in the Taylor rule regression. The model fares better when stability restrictions are relaxed, yet one important puzzle remains with the insignificance of forcing variables in the NKPC and IS curves. This [along with our rejection with FI] may be interpreted as a challenge to a popular theory. Since our specification is illustrative in various dimensions, we prefer to interpret our results as a motivation for further methodological improvements.

5. Conclusion

One can always add assumptions to complete models, as often occurs when models including popular DSGE specifications are taken to data. The existence of a unique and stable rational expectation solution is one key ingredient in this literature. Choices - that can have a substantial impact on subsequent inference - regarding underlying shock processes and observables are other examples of enduring concerns. But it must be asked whether such assumptions are unduly strict, for the case can often be made that some are way more restrictive than economic theory requires. We contribute, via a concrete prototypical example based on the New Keynesian model, to this debate.

On the methodology side, this paper proposed econometric tools that can control statistical error whether the model is complete or not, whether all or a subset of model equations are involved, and whether the latter are statistically identified or not. Our FI methods are not restricted to the model studied here and are sufficiently general to cover any structure that can be solved into an approximated VAR in observables. Our LI methods are even more general, requiring orthogonality conditions akin to GMM.

The approaches proposed in this paper also contribute to the literature on the New Keynesian model. A standard three-equation model encompassing an NKPC, an IS curve and a Taylor rule, is estimated for the United States from 1962Q1 to 2005Q3. We impose and relax the assumption of closed rational expectation model, maintaining similar lag-restrictions on regression disturbances in both cases. In the latter case, single- and multi-equation estimation and fit are compared. When a unique equilibrium is imposed to complete the model, it is rejected by the data. In contrast, our LI method helps recover important information on structural parameters that cannot be reached via single-equation methods.

A key puzzling ingredient remains regarding the forcing variables in the NKPC and the IS curve. Nevertheless, the LI method generates realistic conclusions on the nature of the NKPC, and yields precise estimates of feedback coefficients, which are consistent with the Taylor principle. These results suggest that the unique rational expectation assumption is unduly restrictive for the model studied.

More broadly, two possible uses are envisioned for our proposed procedures. First, our FI method is useful in that it provides a built-in check for whether complete modeling assumptions are counterfactual. While FI approaches may be preferred by adept model builders, complete statistical assumptions can be easier to reject, which may be unwarranted. Again, one can learn from our FI checks on how to overcome deficiencies in structures that lack fit. Second, our LI method is useful in that it can utilize cross-equation information on the variables with as few restrictions as possible, which may have much more to tell about a model than its single-equation counterparts when FI assumptions must be relaxed.

References

Abdelkhalek, T., Dufour, J.-M., 1998. Statistical inference for computable general equilibrium models, with application to a model of the Moroccan economy. Review of Economics and Statistics LXXX, 520–534.

Anderson, G., 2008. Solving linear rational expectations models: A horse race. Computational Economics 31, 95–113.

Anderson, G., Moore, G., 1985. A linear algebraic procedure for solving linear perfect foresight models. Economics Letters 17, 247–252.

Andrews, I., Mikusheva, A., 2011. Maximum likelihood inference in weakly identified DSGE models. Tech. rep., Massachusetts Institute of Technology.

Beaulieu, M.-C., Dufour, J.-M., Khalaf, L., 2010. Identification-robust estimation and testing of the Zero-Beta CAPM. Tech. rep., Mc Gill University, Université Laval and Carleton University.

Benati, L., 2008. Investigating inflation persistence across monetary regimes. Quarterly Journal of Economics 123, 1005–60.

Beyer, A., Farmer, R., 2007. Testing for indeterminacy: An application to U.S. monetary policy. American Economic Review 97, 524–29.

Canova, F., Sala, L., 2009. Back to square one: Identification issues in DSGE models. Journal of Monetary Economics 54, 431–49.

Chari, V., Kehoe, T., McGrattan, E., 2009. New Keynesian models: Not yet useful for policy analysis. American Economic Journal: Macroeconomics 1, 242–66.

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

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

Clarida, R., Galí, J., Gertler, M., 1999. The science of monetary policy: A New Keynesian perspective. Journal of Economic Literature 37, 1661–1707.

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

Cochrane, J. H., 2011. Determinacy and identification with Taylor rules. Journal of Political Economy 119, 565–615.

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

Del Negro, M., Schorfheide, F., Smets, F., Wouters, R., 2007. On the fit of New Keynesian models. Journal of Business and Economic Statistics 25, 123–43.

Doko-Tchatoka, F., Dufour, J.-M., 2008. Instrument endogeneity and identification-robust tests: some analytical results. Journal of Statistical Planning and Inference 138, 2649–2661.

Dufour, J.-M., 1997. Some impossibility theorems in econometrics, with applications to structural and dynamic models. Econometrica 65, 1365–1389.

Dufour, J.-M., 2003. Identification, weak instruments and statistical inference in econometrics. Canadian Journal of Economics 36, 767–808.

Dufour, J.-M., Khalaf, L., 2002. Simulation based finite and large sample tests in multivariate regressions. Journal of Econometrics 111, 303–322.

Dufour, J.-M., Khalaf, L., Kichian, M., 2006. Inflation dynamics and the New Keynesian Phillips Curve: an identification robust econometric analysis. Journal of Economic Dynamics and Control 30, 1707–1728.

Dufour, J.-M., Khalaf, L., Kichian, M., 2010. On the precision of Calvo parameter estimates in structural NKPC models. Journal of Economic Dynamics and Control 34, 1582–1595.

Dufour, J.-M., Taamouti, M., 2005. Projection-based statistical inference in linear structural models with possibly weak instruments. Econometrica 73, 1351–1365.

Fair, R., 2008. Testing price equations. European Economic Review 52, 1424–1437.

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

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

Galí, J., Gertler, M., Lopez-Salido, J. D., 2005. Robustness of the estimates of the hybrid New Keynesian Phillips Curve. Journal of Monetary Economics 52, 1107–18.

Granziera, E., Lee, M., Moon, H. R., Schorfheide, F., 2011. Inference for VARs identified with sign restrictions. Tech. rep., University of Pennsylvania.

Guerron-Quintana, P., Inoue, A., Kilian, L., 2009. Inference in weakly identified DSGE models. Tech. rep., North Carolina State University and University of Michigan.

Gungor, S., Luger, R., 2009. Exact distribution-free tests of mean-variance efficiency. Journal of Empirical Finance 16, 816–829.

Inoue, A., Rossi, B., 2011. Testing for weak identification in possibly nonlinear models. Journal of Econometricsl 161, 246–261.

Iskrev, N., 2010. Local identification in DSGE models. Journal of Monetary Economics 57, 189–202.

Kim, J., 2003. Functional equivalence between intertemporal and multisectoral investment adjustment costs. Journal of Economic Dynamics and Control 27, 533–49.

King, R. G., 2000. The new IS-LM model: Language, logic, and limits. Federal Reserve Bank of Richmond Economic Quarterly 86, 45–103.

King, R. G., Watson, M. W., 1998. The solution of singular linear difference systems under rational expectations. International Economic Review, 39, 1015–1026.

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 DSGE models. Econometrica forthcoming.

Linde, J., 2005. Estimating New Keynesian Phillips Curves: A full information maximum likelihood approach. Journal of Monetary Economics 52, 1135–49.

Ma, A., 2002. GMM estimation of the new Phillips curve. Economic Letters 76, 411–417.

MacKinlay, A. C., Richardson, M. P., 1991. Using generalized method of moments to test mean-variance efficiency. The Journal of Finance 46, 511–527.

Magnusson, L., Mavroeidis, S., 2010. Identification-robust minimum distance estimation of the new keynesian phillips curve. Journal of Money Credit and Banking 42, 1582–1595.

Mavroeidis, S., 2004. Weak identification of forward-looking models in monetary economics. Oxford Bulletin of Economics and Statistics 66, 609–635.

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, 421–449.

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

Moon, H. R., Shorfheide, F., 2010. Bayesian and frequentist inference in partially identified models. Tech. rep., University of Pennsylvania.

Nason, J., Smith, G., 2008. Identifying the New Keynesian Phillips Curve. Journal of Applied Econometrics 23, 525–51.

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

Ray, S., Savin, E., 2008. The performance of heteroskedasticity and autocorrelation robust tests: a Monte Carlo study with an application to the three-factor Fama-French asset-pricing model. Journal of Applied Econometrics 23, 91–109.

Rudd, J., Whelan, K., 2006. Can rational expectations sticky price models explain inflation dynamics? American Economic Review 96, 303–20.

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

Sbordone, A., 2005. Do expected marginal costs drive inflation dynamics. Journal of Monetary Economics 52, 1183–97.

Schorfheide, F., 2008. DSGE model-based estimation of the New Keynesian Phillips Curve. Economic Quarterly 94, 397–433.

Schorfheide, F., 2010. Estimation and evaluation of DSGE models: Progress and challenges. Tech. rep., Invited Lecture at 2010 Econometric Society World Congress in Shanghai.

Sims, C. A., 2002. Solving linear rational expectations models. Computational Economics 20, 1–20.

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

Stock, J. H., Wright, J. H., 2000. GMM with weak identification. Econometrica 68, 1097–1126.

Stock, J. H., Wright, J. H., Yogo, M., 2002. A survey of weak instruments and weak identification in generalized method of moments. Journal of Business and Economic Statistics 20, 518–529.

Taylor, J., 1993. Discretion versus policy rule in practice. Tech. rep., Carnegie-Rochester Conference Series on Public Policy 39.

Woodford, M., 2003. Interest and Prices. Princeton University Press, Princeton, New Jersey.

Table 1. Multi-equation inference - Real-time output gap

Equation	Coefficient	Model-Consistent Instruments	All Instruments
NKPC	ω_f	0.781	0.748
		$[0.577, 0.951]$	$[0.657, 0.848]$
	γ	0.002	-0.011
		$[-0.016, 0.015]$	$[-0.028, 0.005]$
Output	β_f	0.373	0.471
		$[0.233, 0.456]$	$[0.385, 0.556]$
	β_r	28.57	30.0
		$[25.91, 30.0]$	$[24.591, 30.0]$
Taylor Rule	γ_π	1.296	1.326
		$[0.957, 1.578]$	$[1.126, 1.560]$
	γ_y	0.417	0.417
		$[0.281, 0.512]$	$[0.313, 0.544]$
	ρ_1	1.042	1.064
		$[1.009, 1.154]$	$[1.002, 1.125]$
	ρ_2	-0.357	-0.424
		$[-0.533, -0.312]$	$[-0.511, -0.337]$
	ρ_3	0.207	0.248
		$[0.168, 0.312]$	$[0.190, 0.305]$
$\min_{\theta_0} \mathcal{W}(\theta_0)$		1.537	1.445
p-value		0.064	0.057

Note: The estimated model is (5)-(7), with the real-time output gap measure. Estimation applies the limited information method presented in section 3.1.

Table 2: Single equation structure-restricted confidence sets

Coefficient	Standard gap		Real-time gap	
Inflation equation; intra-model instruments				
	iid-GAR	GAR-HAC	*iid*-GAR	GAR-HAC
ω_f	[0.200, 1.0]	[0.510, 1.0]	[0.045, 1.0]	[0.470, 1.0]
γ	[−0.100, 0.050]	[−0.070, 0.010]	[−0.095, 0.055]	[−0.050, 0.015]
Inflation equation; all instruments				
	Standard gap		Real-time gap	
	iid-GAR	GAR-HAC	*iid*-GAR	GAR-HAC
ω_f	[0.315, 1.0]	[0.440, 1.0]	[0.310, 1.0]	[0.455, 1.0]
γ	[−0.10, 0.055]	[−0.055, 0.010]	[−0.09, 0.060]	[−0.040, 0.015]
Output equation; intra-model instruments				
	Standard gap		Real-time gap	
	iid-GAR	GAR-HAC	*iid*-GAR	GAR-HAC
β_f	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
β_r	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
Output equation; all instruments				
	Standard gap		Real-time gap	
	iid-GAR	GAR-HAC	*iid*-GAR	GAR-HAC
β_f	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
β_r	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
Taylor rule; intra-model instruments				
	Standard gap		Real-time gap	
	iid-GAR	GAR-HAC	*iid*-GAR	GAR-HAC
γ_π	$\emptyset$	$\emptyset$	[0.700, 1.950]	[0.700, 1.950]
γ_y	$\emptyset$	$\emptyset$	[0.050, 0.950]	[0.000, 0.950]
Taylor rule; all instruments				
	Standard gap		Real-time gap	
	iid-GAR	GAR-HAC	*iid*-GAR	GAR-HAC
γ_π	$\emptyset$	$\emptyset$	[0.700, 1.950]	$\emptyset$
γ_y	$\emptyset$	$\emptyset$	[0.050, 0.950]	$\emptyset$

Note: The model is (5)-(7), estimated equation by equation, ignoring contemporaneous correlation of disturbances. GAR refers to single-equation generalized Anderson-Rubin method, which applies, equation by equation, the same inference approach as the limited information presented in section 3.1.

Table 3: Single equation reduced-form confidence sets

Inflation equation				
	Intra-model instruments		All instruments	
Coefficient of	Standard Gap	Real-time Gap	Standard Gap	Real-Time gap
$E_t\pi_{t+1}$	[0.892, 1.379]	[0.866, 1.440]	[0.891, 1.230]	[0.865, 1.191]
y_t	[−0.137, 0.026]	[−0.095, 0.082]	[−0.115, 0.026]	[−0.090, 0.054]
Output Equation				
	Intra-model instruments		All instruments	
	Standard Gap	Real-time Gap	Standard Gap	Real-Time gap
$E_t y_{t+1}$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
$R_t - E_t\pi_{t+1}$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
Taylor rule				
	Intra-model instruments		All instruments	
	Standard gap	Real-time gap	Standard gap	Real-time gap
π_t	$\emptyset$	[0.062, 0.234]	$\emptyset$	[0.100, 0.197]
R_t	$\emptyset$	[0.016, 0.079]	$\emptyset$	[0.028, 0.065]

Note: The model is (5)-(7), estimated equation by equation, ignoring contemporaneous correlation of disturbances and relaxing within-equation restriction. Confidence intervals apply the unrestricted projection method from Dufour and Taamouti (2005).