Structural multi-equation macroeconomic models: complete versus limited-information identification-robust estimation and fit*

Jean-Marie Dufour† McGill University

Lynda Khalaf‡ Carleton University

Maral Kichian§ Bank of Canada

May 28, 2012

* The authors thank Carlos de Resende, Frank Kleibergen, Vadim Marmer, Sophocles Mavroeidis, Ulrich Müller, 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 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 Canada Council for the Arts (Killam Fellowship), the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, 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.

† William Dow Professor of Economics, McGill University, Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Centre interuniversitaire de recherche en économie quantitative (CIREQ). Mailing address: Department of Economics, McGill University, Leacock Building, Room 519, 855 Sherbrooke Street West, Montréal, Québec H3A 2T7, Canada. TEL: (1) 514 398 8879; FAX: (1) 514 398 4938; e-mail: jean-marie.dufour@mcgill.ca . Web page: http://www.jeanmariedufour.com

‡ Economics Department, Carleton University, Centre interuniversitaire de recherche en économie quantitative (CIREQ), and Groupe de recherche en économie de l'énergie, de l'environnement et des ressources naturelles (GREEN), Université Laval. Mailing address: Economics Department, Carleton University, Loeb Building 1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6 Canada. Tel (613) 520-2600-8697; FAX: (613)-520-3906. email: Lynda_Khalaf@carleton.ca.

§ Canadian Economic Analysis Department, Bank of Canada, Ottawa, Canada. TEL: (613) 782 7460; FAX: (613) 782 7163; email: mkichian@bankofcanada.ca.

ABSTRACT

We propose two system-based identification-robust methods for structural models including DSGEs that are valid whether identification is weak or strong, and whether identification is theory-intrinsic and/or data specific. The first one is a full-information method, which relies on restrictions strong enough to allow the existence of a rational-expectations solution, while the second one is a limited-information approach that relies on weaker assumptions even though it remains system-based. We apply the proposed methods to a standard New Keynesian model for the U.S. We impose and relax a unique rational expectation solution, maintaining similar lag-restrictions on regression disturbances in both cases. In the latter case, we also compare single-equation to multi-equation estimation and fit. We find that when a unique stable equilibrium is imposed to complete the model, it is rejected by the data. In contrast, limited-information multi-equation inference produces informative results - that cannot be reached via single-equation methods - regarding the importance of forward-looking behavior in the NKPC, and precise conclusions on the feedback coefficients in the reaction function which are not at odds with the Taylor principle.

JEL classification: C52, C53, E37

1 Introduction

Optimization-based macroeconomic models, and, in particular, dynamic stochastic general equilibrium (DSGE) setups, are routinely used for analyzing a multitude of macroeconomic issues, such as studying the effects of alternative policies and conducting welfare analysis. In this respect, the solutions of the 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, it is becoming increasingly clear from the literature that 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 the above: (i) the lack of identification or weak identification, and (ii) assumptions - for example on disturbances, or arising from the model solution - that are subsidiary to the theory yet necessary to complete a model. This paper studies both problems, proposes econometric tools designed to overcome their consequences, and applies these tools to the New Keynesian model.

Identification, which is a long-known econometric concept, relates to the ability of making inferences on theoretical model parameters from observed data.[1] Sometimes, we may face situations where it is theoretically difficult (if not impossible) to determine structural parameters of economic interest from observed data. Identification failure occurs, for example, when the objective function does not respond to some structural parameters, and weak identification arises when the objective function is multi-modal or does not display sufficient curvature in certain parameter regions. Both problems have profound negative implications for usual asymptotic estimators and for the conduct of meaningful inference. Moreover, as emphasized by Canova and Sala (2009), calibrating troublesome parameters or conducting Bayesian estimation and inference do not necessarily overcome these difficulties, and may indeed exacerbate them.[2] In addition, rank and order identification conditions for linear simultaneous equations that were developed in past decades do not apply in the DSGE context; see Komunjer and Ng (2011).

A number of studies document identification problems or failures in well-known estimated macroeconomic models. These include building block equations derived from DSGE setups. For instance, the New Keynesian Phillips Curve (NKPC) has been examined among others by Dufour, Khalaf, and Kichian (2006, 2010a, 2010b), Ma (2002), Mavroeidis (2004, 2005), Nason

[1] For early references see, for example, Koopmans (1950), Hannan (1971) and Zellner and Palm (1974).

[2] See also Guerron-Quintana, Inoue, and Kilian (2009) and Rios Rull, Schorfheide, Fuentes-Albero, Kryshk, and SantaEulalia-Llopis (2011) with regards to Bayesian method.

and Smith (2008), and Kleibergen and Mavroeidis (2009). The results indicate that the NKPC is weakly identified. Similarly, Mavroeidis (2010) and Inoue and Rossi (2011) report identification concerns in Taylor-type monetary policy rules such as the one used by Clarida, Galí, and Gertler (2000). Cochrane (2011) raises further concerns with such rules related to determinacy. The Euler equation for output has also been examined and identification problems documented, notably by Fuhrer and Rudebusch (2004) and Magnusson and Mavroeidis (2010).

As for multi-equation macroeconomic models, studies reporting identification concerns include Ruge-Murcia (2007) who examines a one-sector real business cycle model, Canova and Sala (2009) who consider a DSGE model based on the Smets and Wouters (2003) and Christiano, Eichenbaum, and Evans (2005) setups, Iskrev (2010) who finds issues with the Smets and Wouters (2007) model, Komunjer and Ng (2011) who examine a model considered by An and Schorfheide (2007) and Magnusson and Mavroeidis (2010) who document identification difficulties in a two-equation model that is derived from the fundamental three-equation New Keynesian system. Concerns have also been raised by for example Beyer and Farmer (2007), Cochrane (2011), Kim (2003), Chari, Kehoe, and McGrattan (2009), Fernandez-Villaverde, Rubio-Ramirez, Sargent, and Watson (2007) and Consolo, Favero, and Paccagnini (2009), in multi-equation models with regard to observational equivalence, proper recovery of macroeconomic dynamics from structural VARs, and the role of added ad-hoc measurement errors in DSGE setups.

Broadly speaking, 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 various substantive objects of interest. While there is some consensus that certain models[3] are in principal useful for this purpose, there is less agreement on how such models should be parametrized 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. Examples of the latter dimensions that have been pointed out in the literature particularly with regard to DSGE models include the following.[4] First, an important challenge in DSGE modeling is to minimize the effects of subsidiary assumptions required to complete a model. For instance innovations arising from measurement errors are usually non-fundamental. Alternatively, in some models, the existence of a unique rational expectation solution challenges theory [see Cochrane (2011) with regard to the New Keynesian model]. Second, a typical DSGE-VAR approach raises truncation problems. The

[3]These include for example the New Keynesian model we consider in this paper for concreteness.

[4]See *e.g.* Del Negro, Schorfheide, Smets, and Wouters (2007), Ruge-Murcia (2007), Guerron-Quintana (2010), Schorfheide (2010), Rios Rull, Schorfheide, Fuentes-Albero, Kryshk, and SantaEulalia-Llopis (2011) and the references therein for a general discussion with concrete examples.

finite order VAR representation of the solved model will only be exact if the endogenous state variables are observable and included in the VAR [see, for example, Ravenna (2007)]. Third, DSGE-VAR methods broadly assess the structural form against an unrestricted VAR where - regardless of its statistical fit - the variables that are included in the latter are exactly determined by those that enter the former. The literature is witnessing a growing awareness among applied researchers about the possibility of misspeficiying the benchmark and its consequences. More general specifications for benchmarks are thus gaining popularity, including for example dynamic factor models and VARs augmented by factors.[5] Variable omission is a recognized difficulty, since by construction and because of their specificity, DSGE models may exclude empirically relevant data.[6] For all these reasons, the consequences of spuriously completing models are of obvious concern.

The above shows that identification problems are likely to be prevalent in DSGE models. Indeed, they seem to be almost inherent to the DSGE structure since additionally they partly stem from the presence of forward-looking terms in the model and from the presence of multiple parameter nonlinearities. These can affect both the "population and sample identification" (using the terminology of Canova and Sala (2009)).[7] More importantly, and as will be explained later, identification problems are not necessarily detectable when traditional estimation and test methods are applied.

Authors, such as Canova and Sala (2009), Iskrev (2010) and Komunjer and Ng (2011), propose different approaches to checking identification *ex ante*, suggesting ways to improve the objective function with respect to identification that is intrinsic to the theory. Komunjer and Ng (2011) and Iskrev (2010) develop, under alternative assumptions, formal conditions for verifying the identification of DSGE models. Canova and Sala (2009) focus on estimation using impulse-response matching, and recommend examining the sensitivity of impulse responses to different parameter values through graphic and numerical means. These works constitute important steps in DSGE modeling since they guide model builders away from parameter subspaces

[5]See *e.g.* Del Negro, Schorfheide, Smets, and Wouters (2007) and discussions by Christiano L. J., Gallant A. R., Sims C. A., Faust J. and L. Kilian. See also Boivin and Giannoni (2006), Consolo, Favero, and Paccagnini (2009), and Paccagnini (2011).

[6]Convincing examples are discussed in *e.g.* Consolo, Favero, and Paccagnini (2009). These include fiscal policy variables in a DSGE model focusing on monetary policy, or foreign sector variables in a DSGE model for a closed economy, or model irrelevant financial indicators, term structure variables or variables such as the commodity price index that nevertheless affect the decision of policy makers.

[7]Population identification concerns the mapping between the DSGE and its solution, which - given that it is intrinsic to the theory - is relevant even if we had infinite amounts of observed data. Sample identification concerns the informational link between the structural parameters and the objective function, which is specific to a particular dataset and sample size.

for which objective functions may be non-identified or weakly-identified. In this manner, DSGE solutions are given the best 'fighting chance' before conducting structural estimation and inference. However, there still remains the issue of sampling variability, *i.e.*, the particular dataset and sample size that the applied macroeconomist is confronted with, and where sample identification problems intervene. Given the sample sizes and data qualities that are typically available to macroeconomists, the latter problems can be quite severe, distorting parameter estimates and significantly hampering meaningful inference. For example, Canova and Sala (2009) using a simulation exercise and moment-matching, show that estimation biases are large even with sample sizes of 5000 observations, and conclude that confidence bands obtained around impulse-responses are not informative in the presence of weak-identification.[8]

In this paper, we propose so-called *identification-robust* estimation and test methods, both full-information-based and limited-information-based, for multi-equation DSGE setups. The methods are just as valid and just as reliable whether identification is weak or strong, and whether identification problems arise from issues intrinsic to the theory and/or from sample variability.[9] Therefore, applying these, rather than traditional estimation and inference methods, provides the assurance to the empirical researcher that the obtained quantitative answers from the considered models are always statistically meaningful. In addition, our proposed methods show 'automatically' the extent to which each structural parameter is identified, and, in the event of weak or non-identification, quantify properly the amount of estimate uncertainty (compared to ones obtained from traditional inference methods). Additional advantages include, for example, built-in identification-robust specification tests and, in the limited-information method case, robustness of the applied tests to missing instruments. The latter are variables that contain useful information for identifying one or more structural parameters but that are neither considered in the structure of the DSGE, nor in its econometric counterpart. Finally, the methods are relatively easy to implement.

The concept of identification-robustness will be discussed in some detail in the methodology section, but some intuition can be gleaned about it from the following: when some structural parameters of a given model are non-identifiable or only identifiable on a subset of the parameter space (identification here can be either in the population- and/or sample-identification sense), for confidence sets of estimates to be valid, they should allow for the possibility of being unbounded. That is, if objective functions are flat (or almost flat), all (or practically all) parameter values in the parameter space should be equally admissible as estimates and should therefore show

[8]Useful simulation results on finite sample statistical problems can also be found in Ruge-Murcia (2007) and Jondeau and Le Bihan (2008); see also Guerron-Quintana (2010) on data related difficulties.

[9]For surveys regarding identification difficulties and identification-robust methods in limited-information contexts see, for example, Dufour (2003), Stock, Wright, and Yogo (2002) and Kleibergen and Mavroeidis (2009).

up in the confidence set, thus reflecting the identification status of the examined parameters. Identification-robust methods feature this desirable property. In contrast, any method that, by construction, leads to a confidence *interval* with *bounded limits*, will necessarily have poor coverage under weak-identification (Dufour (1997)). Therefore, intervals of the form {estimate $\pm$ (asymptotic standard error) $\times$ (asymptotic critical point)}, including the delta-method, are fundamentally wrong and cannot be size-corrected. Furthermore, identification difficulty does not necessarily imply that asymptotic standard errors are excessively large. Indeed, the opposite may occur, with tight confidence intervals concentrated on wrong parameter values.

To a certain extent, identification-robust procedures have been gaining credibility in macroeconomics within the context of single-equation models.[10] Yet, despite the considerable associated econometric literature, identification-robust methods for multi-equation systems are still scarce.[11] We propose two system-based identification-robust methods which can address either all of the restrictions implied by the considered model [that is, provide FI inference], or only some of those restrictions [that is, provide LI inference]. Our identification-robust (yet system-based) LI method is predicated on orthogonality conditions that are implied by the model, and is, in some respects, related to the S-sets from Stock and Wright (2000). This paper is among the first to propose a FI method, based on the econometric specification defined by the closed form rational-expectations-consistent solution of the DSGE model.[12] Thus, as well as being identification-robust, it provides the usual features model builders typically seek from full-information contexts. Both our methods rely on transforming the task of estimation and testing from a world where identification difficulties will distort and invalidate the latter to a fully standard context where there is no need to worry about identification issues. Notably, in the full-information case, this is done while maintaining all of the DSGE constraints reflected in the VAR solution. Both methods also rely on 'inverting' identification-robust tests, discussed in the next sections.

We apply these tools, using U.S. data, to an illustrative three-equation New Keynesian model. This fundamental structure has been extensively studied in the literature and forms the

[10]Studies using such approaches include Ma (2002), Dufour, Khalaf, and Kichian (2006, 2010a, 2010b), Mavroeidis (2004, 2005, 2010), Nason and Smith (2008), Kleibergen and Mavroeidis (2009) and Chevillon, Massmann, and Mavroeidis (2010).

[11]Work focusing on identification (from other perspectives than the methods presented here) in multi-equation set-ups include 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).

[12]Guerron-Quintana, Inoue, and Kilian (2009) propose, in addition to a Bayesian method, an alternative likelihood-based method that requires identifying some though not all deep parameters in a model. Andrews and Mikusheva (2011) also examine weak identification in maximum likelihood, with specific focus on adequate estimation of Fisher Information.

building block of many other more complex models.[13] On the substantive side, we address three features of the New Keynesian model. First, we study inflation persistence within the NKPC, given the on-going debate in this regard [see, for example, the survey by Schorfheide (2008)]. Second, we analyze the coefficient of the output gap in the NKPC and of the real interest rate in the output equation, since [as argued by Schorfheide (2010)] available results on their estimates suggest conflicting conclusions about their impact. For clarity, we refer to these coefficients throughout the paper as the coefficients on each equation's *forcing variable*. Third, we revisit the implications of imposing a unique rational expectations solution on the feedback coefficients in the Taylor rule, in light of serious issues arising from determinacy reported by, for example, Mavroeidis (2010) and Cochrane (2011). Comparisons between our full-information and limited information assessments of these questions are discussed. Further comparisons between these and the application of existing univariate identification-robust methods are also discussed, where each method integrates and assesses, to different degrees, the model's structural restrictions.

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. Assumptions underlying a unique solution are restrictive enough to make the complete statistical model easier to reject. This is an important sense in which our analysis can be seen as an exploration of the pervasiveness of subsidiary FI assumptions. In contrast, and although insignificant forcing variables in the NKPC and the output curve cannot be ruled out, our LI multi-equation results allow us to formulate realistic conclusions on the nature of the NKPC, and to obtain precise predictions for feedback coefficients that are not at odds with the Taylor principle. We show that such conclusions cannot be reached via single-equation methods. Taken collectively, results suggest that a multi-equation estimation of the considered model (even when FI assumptions are - and must be - relaxed) can still utilize the information in the contemporaneous relationship between output, inflation, and interest rates, which positively affects identification and inference.

The paper is organized as follows. In section 2, we introduce our general estimation frameworks as well as our representative empirical model. Our methodology is discussed in section 3. Data and empirical results are presented in section 4. We conclude in section 5. A technical Appendix complements the methodology section.

[13] Refer for example to Clarida, Galí, and Gertler (1999), Woodford (2003), Christiano, Eichenbaum, and Evans (2005), Linde (2005), Fair (2008), Benati (2008), Del Negro, Schorfheide, Smets, and Wouters (2007), to mention a few.

2 Framework

This section presents and motivates our econometric set-up. We first discuss a general case that covers standard (linearized) DSGE models. For concreteness, we next consider a prototypical application based on a variant of the New Keynesian model. The latter, extensively studied by Clarida, Galí, and Gertler (1999), is still used in policy circles, and its reported early successes and failures have been the engine behind the development of the various rich DSGE structures that are currently available in the literature. The considered application is also tractable enough to easily illustrate our estimation and test approaches.

2.1 The general set-up

Consider the general structural form

$$\Gamma_0 X_t = \Gamma_1 X_{t-1} + C + \omega \nu_t + \psi \eta_t \tag{2.1}$$

where X_t is vector of m^* variables, C is a vector of constants, ν_t is an exogenous random shock, η_t is a vector of expectations errors (not given exogenously) such that $E_t(\eta_{t+1}) = 0$. Typically, only a subset [denoted Y_t] n^* of the m^* variables included in X_t is observable. The model may be forward looking, in which case time-t expectations for some of the variables would also be included in X_t. For further reference, collect all of the parameters of (2.1) in the vector ϑ. Full Information estimation requires a complete model, which in turn requires specifying the dynamic structure and the distribution of the model's exogenous disturbances, for example as *i.i.d.* or auto-regressive Gaussian processes.

Most standard DSGE models can fit - or can be completed to fit - within this framework. Specifically, available theoretical models can be log-linearized around steady states leading to the (2.1) structure, often with subsidiary assumptions about some of its components. As a matter of fact, general equilibrium theory rarely dictates complete probabilistic structures. So in practice, theoretical structures are completed into the (2.1) form, imposing auxiliary assumptions mainly on exogenous shocks.[14]

2.1.1 Fully specified models

Viewed as a complete structure, (2.1) can be solved forward and its solution has a state-space form that can be expressed as a VARMA model in the observables. The latter can be approximated by a finite-order VAR model whose coefficients are nonlinear functions of the parameters of interest. Specifically, using standard numerical techniques and conventional restrictions on

[14] For a recent survey, see *e.g.* Schorfheide (2010).

ϑ [denoted as $\vartheta \subset \Theta$], (2.1) can be solved into $X_t = C_1 X_{t-1} + C_0 + G\nu_t$, where C_0, C_1 and G are convolutions of ϑ.[15] The solved model admits a restricted VAR approximation in which the number of shocks is equal to the length of the vector of observable variables Y_t:

$$\mathrm{Y}_t = B_0(\vartheta) + B_1(\vartheta)\mathrm{Y}_{t-1} + \quad ... \quad + B_p(\vartheta)\mathrm{Y}_{t-p} + \Sigma(\vartheta)u_t \tag{2.2}$$

with $u_t \sim Normal\ (0, I_{n^*})$. Indeed, focusing on parameters that lead to a unique stable rational expectations solution,Y_t is an infinite VAR that can be approximated via (2.2). The coefficients $B_0(\vartheta), ..., B_p(\vartheta)$, viewed as a function of ϑ, may be constructed by truncation of the infinite VAR or by population regression as in *e.g.* Del Negro, Schorfheide, Smets, and Wouters (2007).[16] Special cases of (2.1) may also admit finite-order VAR representations for which (2.2) holds exactly. We express the associated unrestricted VAR as

$$\mathrm{Y}_t = \mathrm{B}_0 + \mathrm{B}_1\mathrm{Y}_{t-1} + \quad ... \quad + \mathrm{B}_p\mathrm{Y}_{t-p} + \Sigma_u u_t. \tag{2.3}$$

2.1.2 Limited Information Representations

If one trusts (2.1) as a fully specified structure, FIML is typically recommended. Conformably, we first propose an identification-robust full information method. We also propose an alternative approach that overcomes three limitations of FIML: (i) the need to solve the model and the associated constraints; (ii) restrictive and in some dimensions non-model-based dynamic structures; (iii) limited information sets. To be clear, we are not proposing these two methods as necessarily mutually exclusive. Rather, we see our full-information method as providing a useful specification check, and our incomplete-model alternative as allowing a researcher to robustify inferences against auxiliary model assumptions.

The proposed partial specification approach is analogous to generalized method of moments. We thus adopt the following representation. Select n orthogonality conditions of interest compatible with the theoretical model of interest, or with (2.1). Formally, define $\epsilon_{ti}(\mathrm{Y}, \theta)$, $i = 1, ..., n$, where Y denotes observable data on endogenous and exogenous variables and θ the parameters of interest, such that if (2.1) holds then $\epsilon_{ti}(\mathrm{Y}, \theta)$ is orthogonal to a vector of k_i instruments Z_{ti} at the true parameter vector.[17] Collecting all different variables from each of the Z_{ti} into a

[15] See, for example, Anderson and Moore (1985), King and Watson (1998), Sims (2002), Anderson (2008), Komunjer and Ng (2011) and the references therein.

[16] On underlying conditions and truncation costs, see *e.g.* Fernandez-Villaverde, Rubio-Ramirez, Sargent, and Watson (2007) and Ravenna (2007). See also Ruge-Murcia (2007) for simulation evidence on problems associated - among others - with a model's solution. Unique solutions typically rule out unstable equilibria; see Cochrane (2011) for a critical assessment of the implications of this practice on the Taylor rule.

[17] θ may be equal to ϑ, a subset of ϑ or some transformation of the latter.

k-dimensional vector Z_t so that $Z_{ti} = A_i Z_t$ 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 $\epsilon_{ti}(\mathrm{Y}, \theta)$ on Z_t. Expressed differently, the orthogonality restrictions considered imply that at the true parameter vector

$$A_i \Pi_i = 0 \tag{2.4}$$

where $\Pi_i{}'$ is the ith row of Π, the coefficient of the instruments in the following regression

$$\epsilon_t(\mathrm{Y}, \theta) = \Pi Z_t + \mathrm{V}_t, \quad \epsilon_t(\mathrm{Y}, \theta) = (\epsilon_{t1}(\mathrm{Y}, \theta), ..., \epsilon_{tn}(\mathrm{Y}, \theta))', \quad \Pi = \begin{bmatrix} \Pi_1' \\ \vdots \\ \Pi_n' \end{bmatrix}. \tag{2.5}$$

The latter auxiliary n-equation multivariate regression, imposing (2.4), provides a convenient expression for most standard multi-equation and DSGE based structural limited information macroeconomic models. Our notation emphasizes the dependence on θ. The error vector V_t may be *i.i.d.* with possibly non-diagonal variance/covariance matrix. Serial dependence of unknown form may also be considered. Instruments may or may not be lags of the endogenous variables which intervene in $\epsilon_t(.)$. In practice, instruments most often include lagged endogenous variables, as well as a few other variables that are not part of the model and are assumed predetermined or exogenously evolving. The auxiliary system (2.5) is thus most often a VAR, possibly with unequal lags and augmented with non-model explanatory variables. Instruments may also be common to all equations, in which case the Seemingly Unrelated Regression [SURE] type exclusion restrictions [in the sense that exclusion restriction differ across the equations of the system] from (2.4) simplify to $\Pi = 0$. Although limited-information-based, our strategy nonetheless treats the n equations as a system.

2.2 A prototype empirical model

Whereas our methodology is general, we focus on a prototypical New Keynesian application with three equations, based on Linde (2005): a hybrid NKPC equation, an aggregate demand equation and an interest rate rule. Specifically, the model is

$$\begin{aligned} \pi_t &= \omega_f E_t \pi_{t+1} + (1 - \omega_f)\pi_{t-1} + \gamma y_t + \varepsilon_{\pi,t}\,, \\ y_t &= \beta_f E_t y_{t+1} + \sum_{j=1}^{4} \left(1 - \beta_f\right) \beta_{y,j} y_{t-j} - \beta_r^{-1} (R_t - E_t \pi_{t+1}) + \varepsilon_{y,t}\,, \\ R_t &= \gamma_\pi \left(1 - \sum_{j=1}^{3} \rho_j\right) \pi_t + \gamma_y \left(1 - \sum_{j=1}^{3} \rho_j\right) y_t + \sum_{j=1}^{3} \rho_j R_{t-j} + \varepsilon_{R,t}\,, \end{aligned} \tag{2.6}$$

where, for $t = 1, \ ..., \ T$, π_t is aggregate inflation, y_t is the output gap, R_t is the nominal interest rate, $(\varepsilon_{\pi,t}, \varepsilon_{y,t}, \varepsilon_{R,t})' = \Omega \varepsilon_t$ and ε_t is a zero-mean disturbance with identity variance-covariance matrix. For further reference, let

$$\theta = \left(\omega_f, \ \gamma, \ \beta_f, \ \beta_r, \ \gamma_\pi, \ \gamma_y, \rho_1, \ \rho_2, \ \rho_3\right)', \tag{2.7}$$

$$\phi = \left(\theta', \ \beta_{y,1}, \ \beta_{y,2}, \ \beta_{y,3}, \ \beta_{y,4}\right)', \tag{2.8}$$

refer to vectors of the model's "deep" parameters, and let Θ and Φ denote the associated parameter spaces. We next present further modeling assumptions (all viewed as illustrative) comparing our FI case to its LI counterpart. For clarity, and as defined above, we refer to γ and β_r^{-1}as the coefficients on the forcing variable in the NKPC and the output equation respectively.

2.2.1 Full Information assumptions

Our FI method assumes $(\varepsilon_{\pi,t}, \varepsilon_{y,t}, \varepsilon_{R,t})' \overset{iid}{\sim} N(0, \Omega)$ with Ω invertible but not necessarily diagonal. Model (2.6) may be represented as in (2.1) for example by replacing expectations of variables with actual values of the same variables, while adding expectation error terms to the equation. From there on, the model can be solved into the form (2.2) with $\mathrm{Y}_t = (\pi_t, y_t, R_t)'$, $p = 4$, $B_1(\vartheta) = B_1(\phi), ..., B_p(\vartheta) = B_p(\phi)$ and $\Sigma(\vartheta) = \Sigma(\phi, \Omega)$ where the error terms are three-dimensional *i.i.d.* multivariate standard normal.

Despite a broad consensus on a common theoretical basis, there is less consensus in this literature on how to complete the model's probabilistic structure for estimation purposes. Our assumptions on (2.6) are considered prototypical. Specific dimensions along which model (2.6) is viewed as illustrative [with reference, for example, to the above cited works on the New Keynesian model] are discussed in Section 4.

2.2.2 Limited information assumptions

Model (2.6) implies that the vector $(\epsilon_{\pi,t}(\mathrm{Y}, \theta), \epsilon_{y,t}(\mathrm{Y}, \theta), \epsilon_{R,t}(\mathrm{Y}, \theta))'$ with elements defined as

$$\begin{aligned} \epsilon_{\pi,t}(\mathrm{Y}, \theta) &= \pi_t - \omega_f \pi_{t+1} - (1 - \omega_f)\pi_{t-1} - \gamma y_t, \\ \epsilon_{y,t}(\mathrm{Y}, \theta) &= y_t - \beta_f y_{t+1} + \beta_r^{-1}(R_t - \pi_{t+1}), \\ \epsilon_{R,t}(\mathrm{Y}, \theta) &= R_t - \left(1 - \sum_{j=1}^{3} \rho_j\right)(\gamma_\pi \pi_t + \gamma_y y_t) - \sum_{j=1}^{3} \rho_j R_{t-j}, \end{aligned} \tag{2.9}$$

and reflecting expectational error - coming from the replacement of expected terms with their observed future values plus an error term - which we refer to as a three-dimensional 'Structural

Residual' (akin to Euler errors in the context of GMM) is uncorrelated with available instruments for the true value of $\theta \in \Theta$. In our notation,Y refers to the observed data on $\mathrm{Y}_t = (\pi_t, y_t, R_t)'$, $t = 1, .., T$, conforming with our general set-up. We thus have testable orthogonality conditions which can be taken to the data. For later reference, we also define

$$\dot{Z}_t = (\pi_{t-1}, R_{t-1}, R_{t-2}, R_{t-3})', \ \ddot{Z}_t = (y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4})'. \tag{2.10}$$

The elements of $\dot{Z}_t$ and $\ddot{Z}_t$ are the predetermined variables in the system and which we refer to hereafter as "intra-model" instruments.

The assumption that $\dot{Z}_t$ and $\ddot{Z}_t$ are the only relevant instruments for the consider model seem implausible as descriptions of policy-maker behavior. We thus expand the information set using (lags of) non-modelled variables as extra instruments, and group them in a vector denoted $\tilde{Z}_t$.[18] Specific dimensions along which our considered information set is viewed as illustrative are discussed in Section 4.

3 Methodology

Both estimation strategies have advantages and disadvantages, none restricted to the models under consideration in this paper.[19] However, both methods share the following difficulty. If the confidence intervals and hypothesis tests that result from these estimation strategies are, as is typically the case, validated through the use of standard asymptotic arguments, they can easily become unreliable when there are identification difficulties.[20] Instead, the methods that we propose are identification-robust. Thus, they are valid whether identification is weak or strong, and whether identification problems arise from issues intrinsic to the theory and/or from sample variability. This section provides a mostly descriptive discussion of our approaches; complete formulae and further references are relegated to the Appendix.

[18]For supportive evidence on the worth on extra-model instruments in identification-robust instrumental regressions, refer *e.g.* to Dufour, Khalaf, and Kichian (2010b). In estimating four structural inflation equations [two specifications based on Eichenbaum and Fisher (2007) and the Blanchard and Gali (2007, 2010) specifications], Dufour, Khalaf, and Kichian (2010b) find that inference on the Calvo parameter sharpens importantly when rather than restricting instruments to the lags of each model's endogenous variables, the lags of the endogenous variables from all considered models are used as instruments for each model.

[19]For a general discussion, the reader may refer to Canova (2007, Chapters 4-6).

[20]The econometric literature is extensive on the topic. Refer, for example, to Dufour (1997, 2003), Staiger and Stock (1997), Wang and Zivot (1998), Zivot, Startz, and Nelson (1998), Dufour and Jasiak (2001), Kleibergen (2002, 2005), Stock, Wright, and Yogo (2002), Moreira (2003), Dufour and Taamouti (2005, 2007), Andrews, Moreira, and Stock (2006), Hoogerheide, Kaashoek, and van Dijk (2007), Bolduc, Khalaf, and Moyneur (2008), Bolduc, Khalaf, and Yelou (2010), Chaudhuri, Richardson, Robins, and Zivot (2010) and Kleibergen and Mavroeidis (2010).

3.1 Test Inversion and identification robustness

What makes our methods identification-robust can be explained by first noting that both our FI and LI approaches share the following fundamental premise: whereas in traditional estimation methodology a point estimate is found first and confidence intervals are then constructed, we proceed in reverse. That is, first we build a confidence region, then we obtain a point estimate from it. The confidence region, at level $(1-\alpha)$ (say, 95%), is obtained by "inverting" a test specifically designed so that, whatever the identification situation of the considered model, its significance level remains at α (in this case, 5%). A test that satisfies this property and its associated confidence region are referred to as being "identification-robust".

Inverting a test means assembling, analytically or numerically, the set of parameter values that are not rejected by this test. A point estimate can be obtained by picking the least-rejected parameters from within the confidence region, that is, by choosing those parameter values that are associated with the largest test p-value.[21] A built-in specification check, providing an overall assessment of the structural model restrictions, is also available within such test-inversion-based procedures. In particular, if the generated confidence region is empty, the model is rejected at the considered test level.

While the above-cited econometric literature has documented the superiority of such methods over traditional estimation and inference approaches, except for the few above-cited works, multi-equation models have not been directly addressed. We propose two identification-robust tests for inversion that are suited within the above described general and specific models [(2.1) and (2.6)]. For convenience, we first discuss our LI method. This will help present the basic identification-robustness principles we follow for both LI and FI methods.

3.2 The structural Limited Information method

In the context of (2.1)-(2.4), consider the null hypothesis

$$H_{01}:\ \theta=\theta_0 \tag{3.11}$$

where θ_0 is a known value. We introduce an n-dimensional system of artificial regressions

$$\epsilon_t\left(\mathrm{Y},\theta_0\right)=\Pi Z_t+V_t, \tag{3.12}$$

in which context we propose to assess

$$H_{01}^{*}:A_i\Pi_i=0 \tag{3.13}$$

[21] These are the so-called Hodges-Lehmann point estimates; see Hodges and Lehmann (1963, 1983), and Dufour, Khalaf, and Kichian (2006, 2010a, 2010b).

where $\epsilon_t(\mathrm{Y},\theta)$, Π and A_i conform with (2.5)-(2.4). Indeed, if the null (3.11) is true, then (3.13) should hold. Therefore, once can simply test for H_{01}^* within (3.12) to assess the hypothesis (3.11). That is, if θ_0 represents the true parameters, then additional information from predetermined variables should be irrelevant. This is very convenient because the system represented by (3.12) does not require statistical identification (the right-hand side regressors are not 'endogenous') so that usual statistics for testing the exclusion of regressors can be applied in a straightforward manner. In addition, since the left-hand-side of (3.12) is the stacked n-dimensional structural residual that conforms with the structural model, the exclusion restrictions in H_{01}^* represent orthogonality conditions as in GMM. Moreover, structural information from the model is also captured by the contemporaneously-correlated disturbances within (3.12) since they embed the correlation structure of the considered structural residuals. With model (2.6), $\theta_0 = (\omega_f^0, \gamma^0, \beta_f^0, \beta_r^0, \gamma_\pi^0, \gamma_y^0, \rho_1^0, \rho_2^0, \rho_3^0)'$, $\epsilon_t(\mathrm{Y},\theta_0) = (\epsilon_{\pi,t}(\theta_0), \epsilon_{y,t}(\theta_0), \epsilon_{R,t}(\theta_0))'$ where $\epsilon_{\pi,t}(\mathrm{Y},\theta)$, $\epsilon_{y,t}(\mathrm{Y},\theta)$ and $\epsilon_{R,t}(\mathrm{Y},\theta)$ are as in (2.9), $n = 3$, and if $Z_t = \left(\dot{Z}_t', \ddot{Z}_t'\right)'$ with $\dot{Z}_t$ and $\ddot{Z}_t$ as in (2.10), then A_1 and A_3 should select all coefficients of the first and third equation of (3.12), whereas A_2 should select the coefficients in the second equation associated with $\dot{Z}_t$.[22] If $Z_t = (\dot{Z}_t', \ddot{Z}_t', \tilde{Z}_t')'$ then for all equations the A_i matrices should also select the coefficients of the extra-model instruments $\tilde{Z}_t$.

3.2.1 Test procedure

To test the H_{01}^* hypothesis, traditional criteria may be used in the framework of the artificial regression (3.12), on recalling that the latter does not suffer from the endogenous regressor problem. Indeed, this is the intuition exploited by Stock and Wright (2000) leading to inverting the GMM objective function; related arguments also underlie the procedures analyzed in Kleibergen and Mavroeidis (2009) in the context of the single-equation NKPC. A weighting matrix, treated as a function of θ, matters importantly here. An optimal weighting matrix and continuous updating is required for some of the efficient methods proposed by Kleibergen and Mavroeidis (2009).

Our test criterion differs from traditional GMM practices in the following ways: (i) we use a weighting function that, while depending explicitly on the tested θ_0 value of the parameter for efficiency and identification-robustness purposes, avoids the iterative continuous updating optimal GMM,[23] and (ii) we embed the cross-equation restrictions from (2.6) into the test, via

[22] Note that, in this example, the coefficients on the output gap lags are free in the output equation, and thus the exclusion of their coefficients is not tested within the second equation of (3.12).

[23] Our method does not require iterating the GMM objective function because we rely on the SURE artificial regression framework. In this case it is well-known [see Dufour and Khalaf (2002a, 2002b, 2003) and the references

analytically-tractable estimates of variance/covariance matrix of (3.12).

We use one of the most popular F-type Wald statistics in SURE analysis (see the Appendix for details). The test, denoted $\mathcal{W}(\theta_0)$, has an approximate [imposing homoskedasticity] null distribution given by $F(m, n(T-k))$, where n is the dimension of the structural residual vector, k is the number of regressors per equation [the dimension of Z_t] and m is the total number of tested coefficients in (3.12). This null distribution is standard and does not depend on unknown parameters even if instruments used are weak. In our empirical analysis, $m = 2k + (4 + q)$ where q is the number of external instruments in $\tilde{Z}_t$.

3.2.2 Test inversion

The test inversion itself must be conducted numerically. One can use, for example, a grid search over the economically-meaningful set of values for θ. Thus, one can sweep, in turn, the choices for θ_0, and for each choice considered, compute the relevant test statistic, $\mathcal{W}(\theta_0)$, and its associated p-value. The parameter vectors for which the p-values are greater than the level α collected together constitute the identification-robust confidence region with level $1 - \alpha$.

Moving from the joint confidence region to individual confidence sets for each component of θ is achieved by projecting this region, i.e. by computing, in turn, the smallest and largest values for each parameter included in this region. A point estimate can also be obtained from the joint confidence set. This corresponds to the model that is most compatible with the data, or, alternatively, that is least-rejected, and is given by the vector of parameter values with the highest p-value in the set.

Projection-based confidence sets are obtained numerically as follows. By definition, a set can be obtained for any function $g(\theta)$ by minimizing and maximizing the function $g(\theta)$ over the θ values included in the joint confidence region. We thus define each component of θ as a linear combination of θ, of the form $g(\theta) = a'\theta$, where a is a conformable selection vector (consisting of zeros and ones); for example, $\omega_f = (1, 0, \ldots, 0)\,\theta$. We then obtain the projection set by numerical optimization of the associated $a'\theta$ function over θ such that $\mathcal{W}(\theta) < F_\alpha(m, n(T-k))$. We use Simulated Annealing (see Goffe, Ferrier, and Rogers (1994) for this purpose).

therein] that iterative GLS does not improve efficiency in practice. Moving from the GMM to the SURE-GLS context via an artificial regression that nonetheless captures the same orthogonality restrictions thus greatly decreases numerical burdens without efficiency costs.

3.2.3 Specification Checks, dynamics and the information set

Note that the cut-off point for the $\mathcal{W}(\theta_0)$ statistic is the same for any value θ_0 under test. Define $\overline{\mathcal{W}} = \min_{\theta_0} \mathcal{W}(\theta_0)$. Since it is the case that

$$\min_{\theta_0} \mathcal{W}(\theta_0) \geq F_\alpha(m, n(T-k)) \Leftrightarrow \mathcal{W}(\theta_0) \geq F_\alpha(m, n(T-k)), \quad \forall \theta_0 \tag{3.14}$$

where $F_\alpha(.)$ denotes the α-level cut-off point under consideration, then referring the latter statistic to an $F(m, n(T-k))$ cut-off point (say at level α) provides an identification-robust specification test. Indeed, this is a sort of identification-robust J-test.

Such a specification check can be carried out before the test inversion step to save computation time; if the outcome is not significant [*i.e.* if $\min_{\theta_0} \mathcal{W}(\theta_0) < F_\alpha(m, n(T-k))$], then we can be sure that the associated confidence sets for θ will not be empty. In view of the underlying nonlinearity, the latter minimizations must be performed numerically. We again recommend a global optimization procedure such as Simulated Annealing because there is no reason to expect that $\mathcal{W}(\theta_0)$ is a smooth function of θ_0.

Expectation errors resulting from the replacement of expected terms with observed future values plus an error, may lead to MA effects when a LI method is considered; see Mavroeidis (2004) and Kleibergen and Mavroeidis (2009). Serial dependence may also be required to fit the data. We thus introduce a HAC-type version of our statistic to account for potential serial dependence problems. Formally, we extend the popular multivariate method introduced by MacKinlay and Richardson (1991) and analyzed by Ravikumar, Ray, and Savin (2000) to the macroeconomic model under consideration (details are found in the Appendix). The statistic we use, denoted $\mathcal{J}(\theta_0)$, has an approximate null distribution given by $\chi^2(m)$ where m is, as defined above, the total number of coefficients tested out within (3.12). This null distribution is again standard, does not depend on unknown parameters even if instruments used are weak, and leads to a cut-off point for the $\mathcal{J}(\theta_0)$ statistic that is, again, the same for any value θ_0 under test. The associated confidence region based on inverting the HAC statistic thus also admits the possibility of being both empty [when $\min_{\theta_0} \mathcal{J}(\theta_0) \geq \chi^2_\alpha(m)$] and unbounded, with the former indicating model misspecification, and the latter, lack of identification.

The statistic that we use and the continuously updated GMM-type objective function (as in, for example, Stock and Wright (2000)) are asymptotically equivalent given certain regularity conditions. Kleibergen and Mavroeidis (2009, 2010) propose an adjusted cut-off point that, under specific assumptions, can lead to rejecting the model more liberally.[24] For example, instead of the $\chi^2(m)$, these authors recommend using the χ^2 approximate distribution with

[24] These can often be more restrictive than, for example, the assumptions of Stock and Wright (2000).

degrees-of-freedoms reduced to m minus the number of parameters tested (here, the dimension of θ). However, recent econometric studies show that system-based HAC criteria suffer from the curse of dimensionality as much as (and perhaps even more than) their *i.i.d.* counterparts, and may thus perform poorly in finite samples.[25] This suggests that one must interpret the results of HAC tests [with and without the correction from Kleibergen and Mavroeidis (2009, 2010)] with caution.

The applied LI tests do not suffer (in the sense that size is not affected) from a further complication that frequently arises in practice: the case of "missing instruments". The latter are variables (such as lags of endogenous variables or predetermined variables) that contain useful information for identifying one or more structural parameters but that are neither considered in the theoretical framework of the DSGE, nor in its econometric version. Details regarding validity of our LI tests when such instruments are missing are discussed in the Appendix, for concreteness, in the context of our empirical model.

3.3 Full-information method

Given (2.1)-(2.2), consider the null hypothesis given by

$$H_{02} : \vartheta = \vartheta_0 \tag{3.15}$$

where the parameter values with the zero superscript are assumed to be known, but unlike in our LI method, are restricted so that an associated rational expectation solution exists. Following the logic set out for the LI case, we introduce an n^*-dimensional artificial VAR

$$U_t\left(\mathrm{Y},\vartheta_0\right) = \Pi Z_t + W_t \tag{3.16}$$

$$U_t\left(\mathrm{Y},\vartheta_0\right) = \mathrm{Y}_t - B_0\left(\vartheta_0\right) - B_1\left(\vartheta_0\right)\mathrm{Y}_{t-1} - \quad \ldots \quad - B_p\left(\vartheta_0\right)\mathrm{Y}_{t-p} \tag{3.17}$$

in which Z_t includes as many lags of each component of Y_t as dictated by the structure. In this context we propose to assess

$$H_{02}^* : \Pi = 0. \tag{3.18}$$

In addition, consider the special case where for which (2.2) holds exactly and ϑ can be partitioned as $\vartheta = (\phi', \bar{\phi}')'$ so that the coefficients $B_1(.), ..., B_p(.)$ depend on ϕ but not on $\bar{\phi}$. In this case, (2.2) simplifies to

$$\mathrm{Y}_t = B_0\left(\phi\right) + B_1\left(\phi\right)\mathrm{Y}_{t-1} + \quad \ldots \quad + B_p\left(\phi\right)\mathrm{Y}_{t-p} + \Sigma\left(\phi, \bar{\phi}\right)u_t. \tag{3.19}$$

[25] For recent simulation evidence, see Ray and Savin (2008), Gungor and Luger (2009) and Beaulieu, Dufour, and Khalaf (2010); these studies use simulation designs based on financial models where sample sizes may be much larger than what is typically available in usual macroeconomic contexts.

so we can focus on the partialled-out hypothesis

$$H_{02} : \phi = \phi_0, \tag{3.20}$$

leading to the artificial VAR

$$U_t\left(\mathrm{Y},\phi_0\right) = \Pi Z_t + W_t, \tag{3.21}$$

$$U_t\left(\mathrm{Y},\phi_0\right) = \mathrm{Y}_t - B_0\left(\phi_0\right) - B_1\left(\phi_0\right)\mathrm{Y}_{t-1} - \quad ... \quad - B_p\left(\phi_0\right)\mathrm{Y}_{t-p}. \tag{3.22}$$

This holds in the case of (2.6), with ϕ as in (2.8), and $Z_t = \left(\dot{Z}_t', \ddot{Z}_t'\right)'$ with $\dot{Z}_t$ and $\ddot{Z}_t$ as in (2.10).

Testing for H_{02}^* within the artificial VAR [(3.16) or (3.21)] provides a test of H_{02} so long as $U_t\left(\mathrm{Y},\phi_0\right)$ exists. That is, we transform the task of estimation and testing from a world where identification difficulties will distort and invalidate the latter to a fully standard context where there is no need to worry about identification issues. This is done while maintaining all of the DSGE constraints reflected in the VAR solution, and thus the method is full-information-based.

When (2.2) is an approximation, then the VAR residuals are not *i.i.d.* The magnitude of the discrepacny decreases with large p, though a HAC version of the considered test statistic could be used for each parameter value to be tested in the construction of the confidence set.

3.3.1 Test procedure

To test the exclusion restrictions in H_{02}^* within the subspace restricted by the existence of $U_t\left(\mathrm{Y},\phi_0\right)$, we use one of the most popular likelihood-based test statistic in multivariate regression analysis, which is a monotonic transformation of the likelihood ratio criterion (details can be found in the Appendix). The statistic, denoted $\mathcal{L}\left(\vartheta_0\right)$ or $\mathcal{L}\left(\phi_0\right)$ in the partialled-out case, and has an approximate [imposing homoskedasticity] null distribution of $F(Kn^*, \mu\tau - 2\lambda)$, where n^* is the dimension of Y_t, K is the dimension of Z_t, and where μ, τ and λ are given in (C.32)-(C.34) and depend only on n, T and K.

In our empirical application, $n^* = n = 3$, $Z_t = \left(\dot{Z}_t', \ddot{Z}_t'\right)'$ so $K = 8$. As with the LI case, this null distribution is standard and does not depend on unknown parameters even if identification is weak.

3.3.2 Test inversion

The test inversion procedure is similar to that presented in the previous section. Using, for example, a grid search over the economically-meaningful set of values for ϕ, we sweep, in turn, the choices for ϕ_0. Choices for the latter are of course restricted to ensure that the above-defined $U_t\left(\mathrm{Y},\phi_0\right)$ exist. To do so, we check for the usual existence conditions for every candidate ϕ_0

value, using the Anderson and Moore algorithm. For each acceptable choice, we compute the relevant test statistic, $\mathcal{L}(\phi_0)$ and its associated p-value. The parameter vectors for which the p-values are greater than the level α constitute the identification-robust confidence set with level $1-\alpha$.

Again, it is possible to construct the confidence region using projection-based methods. They can be obtained for any function $h(\phi)$ by minimizing and maximizing the functions $h(\phi)$ over ϕ such that $\mathcal{L}(\phi_0) < F_\alpha(Kn\,,\,\mu\tau-2\lambda)$. One can then obtain point estimates and confidence intervals as explained before.

3.3.3 Specification Checks

The cut-off points for the $\mathcal{L}(\phi_0)$ statistic are the same for any value ϕ_0 under test. So if we define $\bar{\mathcal{L}} = \min_{\phi_0} \mathcal{L}(\phi_0)$, referring the latter to a F cut-off point (say at level α) with degrees-of-freedom as in (C.32)-(C.34), provides an identification robust specification test. This follows from the fact that

$$\min_{\phi_0} \mathcal{L}(\phi_0) \geq F_\alpha(Kn\,,\,\mu\tau-2\lambda) \Leftrightarrow \mathcal{L}(\phi_0) \geq F_\alpha(Kn\,,\,\mu\tau-2\lambda), \quad \forall\phi_0 \tag{3.23}$$

where $F_\alpha(.)$ denotes the α-level cut-off point under consideration.

$\bar{\mathcal{L}}$ assesses the model as a complete specification, where assumptions are sufficiently strong to ensure the existence of a unique model solution. It is important to stress that while such a feature is valued by many model builders, complete model assumptions are restrictive. We expect that this will make statistical relationships easier to reject and will constitute an important limitation for empirical work. On the positive side, one can learn from a specification check - when as with $\bar{\mathcal{L}}$, it is hardwired into FI estimation - on how to improve structures that are at odds with the data.

4 Empirical analysis

We study model (2.6) as a concrete and well-known example of general structures consistent with the literature. This analysis is viewed as illustrative in various respects. First, (2.6) includes lags in the output and interest rate equations that are not strictly derived from New Keynesian foundations. Typically, completing a New Keynesian model requires non-theory based choices regarding, for example, the inclusion of auxiliary shocks or measurement errors, and assumptions about the law of motion of the considered shocks. Various - reasonable although typically *ad hoc* - options are considered for this purpose, and on balance, none emerges as a best choice.

The assumption we make by adding lags to justify an *i.i.d.* Gaussian assumption on ε_t follows Linde (2005).[26]

Second, model (2.6) imposes no cross-equation restrictions on regression parameters. Since existing works provide no consensus view in this regard, our specification suggests a minimal set of assumptions for estimation purposes.

Third, model (2.6) is a special case of (2.1) in which the number of structural shocks is exactly equal to the number of endogenous variables.[27] Furthermore, its solution [with reference to (2.2)] satisfies $B_1(\vartheta) = B_1(\phi), ..., B_p(\vartheta) = B_p(\phi)$, that is $B_1(.), ..., B_p(.)$ depends on ϕ, the model's deep parameters of interest defined in (2.8) but does not depend on Ω. This allows one to conveniently partial Ω out in estimation.[28] The solution of (2.6) also imposes exclusion restrictions on $B_1(.), ..., B_p(.)$ 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. Conformably, these same exclusion restrictions are imposed on the unrestricted benchmark VAR (2.3) considered. The natural question here is how well does this benchmark represent the data.

Fourth, again conforming with the above cited literature, the solution we empirically maintain rules out sunspot equilibria. In other words, our closed model approach follows the usual practice of restricting parameter values so that a unique rational expectations 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. Specifically, 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 macroeconomies, this is counterfactual. Rather than maintaining this restriction, 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, which we refer to as "extra-model instruments". Again as an illustrative example, as external instruments we consider lags 2 and 3 of both wage and commodity price inflation, conformable with the literature. Using these

[26] It is worth noting that Linde (2005) assumes a diagonal covariance matrix; see Curdia and Reis (2011) for supportive evidence in favour of relaxing the usual uncorrelated shocks assumption.

[27] Recall that for general structures of the (2.1) form, lag truncation affects the accuracy of the VAR(p) solution and the associated benchmark. The statistical solved form is a VAR exactly only when all the endogenous variables are observable.

[28] Although details are not provided, this observation conforms with Linde (2005, footnote (20)) which suggests that variance parameters have been partialled out in FIML.

additional variables in the estimation may potentially sharpen inference about some structural parameters of interest, in an econometrically convenient manner.

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.

4.1 The data

We conduct our applications using U.S. data for the sample extending from 1962Q1 to 2005Q3. We use the GDP deflator for the price level, P_t, and the Fed Funds rate as the short-run interest rate.

Our LI estimations can be conducted using either intra-model instruments, or intra-model instruments supplemented with external ones. As external instruments, we consider lags 2 and 3 of both wage and commodity price inflation conformable with the literature. Specifically, commodity price inflation data may capture global factors.[29] Finally, as in Linde (2005), all our data is demeaned prior to estimation. Demeaning all variables, including instruments when used, corresponds to allowing for unrestricted constants in the model studied (that is, it allows to express explained variables in deviation with respect to (potentially non-zero) unknown equilibrium values). Constraints on equation constants are evacuated by demeaning and thus are not accounted for.[30] The demeaning also allows for a fair comparison between our results and Linde's.

For the output gap, we consider two measures. The first 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, 2010a, 2010b), we proceed iteratively: to obtain the value of the gap at time t, we detrend GDP 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, and yields 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. The 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. We then take the log of both these output gap series.

[29] It has been argued [see e.g. Boivin and Giannoni (2009)] that international factors may be relevant in the determination of the U.S. macroeconomy.

[30] Canova and Sala (2009) provide an illustrative example that discusses the consequences of such an empirical practice. See also Cochrane (2011) for other issues with model intercepts.

4.2 Substantive questions

On the substantive side, we use (2.6) to assess three features of the New Keynesian model. First, we study intrinsic inflation persistence within the NKPC. Formally, we test whether values of ω_f less than 0.5 can be ruled out, that is whether the NKPC is conclusively more forward than backward looking. We also check whether the pure forward looking case [that is $\omega_f = 1$] can be refuted. For insights and perspectives on the importance of lagged inflation, refer to Linde (2005), Benati (2008), Fair (2008), Nason and Smith (2008), Schorfheide (2008) and the references therein. The surveys by Schorfheide (2008, 2010) confirm enduring disagreements in this regard. We thus ask whether our systems approach can sharpen our inference on the nature of the NKPC, relative to single equation methods.

Second, we check whether insignificant forcing variables in the NKPC and the output equation can be ruled out. Formally, we test whether 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. The consequences of the researcher getting the slopes of the NKPC and IS curve wrong are of obvious concern. In contrast to single-equation models, systems 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 captured by our systems approach relative to single equation methods.

Third, we ask whether a systems approach can recover any useful information on the feedback coefficients in the Taylor rule (γ_π and γ_y). Mavroeidis (2010) reports identification problems for these parameters from a single-equation perspective. Fundamental issues with such rules - arising from imposing unique rational expectations 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 systems-based inference on the Taylor rule, which suggests to check whether imposing stability on the considered system has any empirical support.

4.3 Results and discussion

Our systems inference produces a striking result. With both gap measures, the model is rejected at the 5% level using our FI method. Again, for both gap measures, the model is also rejected at the 5% level using the multi-equation HAC statistic, with and without external instruments. Formally, with both gap measures and using our above defined notation we have: $\min_{\phi_0} \mathcal{L}(\phi_0) \geq F_{5\%}(Kn\,,\,\mu\tau - 2\lambda)$, and $\min_{\theta_0} \mathcal{J}(\theta_0) \geq \chi^2_{5\%}(m)$ with and without external instruments [degrees-of-freedom are given in subsections (3.2.1) and (3.3.1)]. Note that in what follows and unless otherwise indicated, a 5% level is implied when we discuss test rejections. With our multi-equation LI method and an *i.i.d.* errors assumption, we obtain an empty confidence set [$\min_{\theta_0} \mathcal{W}(\theta_0) < F_{5\%}(m, n(T-k))$] 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 so far to those obtained by Linde (2005), since these are quite different (despite our reliance on similar specifications). Linde argues, using Monte Carlo experiments, that FIML methods are superior to GMM-type approaches at uncovering the true values of the structural parameters, and his estimations on US data show that the NKPC is preponderantly backward-looking, and moreover, that using either different measures for the y_t variable will yield qualitatively similar results. We find that FI actually leads us to reject the model, and furthermore, that the proxies used for the gap have profound implications for LI estimations. This last conclusion also contrasts with Kleibergen and Mavroeidis (2009) who report using single equation weak-instruments robust methods that their estimations on 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 systems LI statistic is the instrument validity problem discussed in Doko-Tchatoka and Dufour (2008) (in the context of general IV-based inference) and Dufour, Khalaf, and Kichian (2010b) (in the case of various empirical NKPC specifications): when some of the instruments or lagged endogenous variables of the model are not truly predetermined, tests [including the identification-robust procedures] that rely on them will yield spurious results. Given that the standard output gap measure is obtained using all of the sample observations, its lags $(t - \tau)$ may actually be correlated with the time t error terms, making them inappropriate for use as predetermined variables or legitimate instruments.

More subtle arguments can be raised that question 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 regards to assumptions on disturbances, note that our model passes the LI test imposing *i.i.d.* regression errors and fails when serial dependence is allowed. This observation has to be qualified given: (i) the risk of spurious rejections since the systems HAC statistics are known to be oversized, and (ii) the

Table 1. Multi-equation Inference - Real-Time Output Gap

Equation	Coefficient	Model-Consistent Instruments	Intra and Extra-model 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 (2.6), with the real-time output gap measure [refer to section 4.1]. Estimation applies the limited information method presented in section 3.2.

specific lag structure adopted to justify *i.i.d.* errors.[31] 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 here. Our FI rejection may also be linked to the usual culprits, that is, 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 more forceful assumptions. Our LI method is applied maintaining the same lag-length restrictions on disturbances as the FI one, and in contrast to FI, the former method gives the model a chance, which is interesting to point out.

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 (2.6) 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, that is with the real-time gap, and using our LI method. Tables 2 and 3 report our confidence intervals with single equation identification-robust methods. These include the Generalized Anderson-Rubin (GAR) methods [see Stock and Wright (2000)] used by Kleibergen and Mavroeidis (2009), Mavroeidis (2010) and Dufour, Khalaf, and Kichian (2006, 2010a, 2010b) that 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 are not at odds with the literature with regard to the 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. Such a conclusion was arrived at among others by Galí, Gertler, and Lopez-Salido (2005) and Sbordone (2005), who examine the closed-form version of the NKPC in a single-equation context, and by Smets and Wouters (2007) who estimate a medium scale DSGE model using Bayesian methods; for a more global perspective, see Schorfheide (2008). 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, Galí, and Gertler (2000)).

[31] 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.

Point estimates do not change much whether the full instrument set or the model-consistent instrument subset are used. However, we find that outcomes are subtly different when we assess the sensitivity of the confidence intervals to the considered information set. In particular, while it is possible to ascertain that the NKPC is a forward-looking variable based on both instruments sets [values of ω_f below 0.5 are rejected], values near one for ω_f and less than one for γ_π are ruled out with our full set of instruments, and are not rejected with its model-consistent subset. This result, particularly in the case of ω_f, is even more noteworthy when our multi-equation estimates are contrasted with single equation ones.

The confidence intervals from 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 tighten up importantly when the instruments set is expanded and when the *i.i.d.* assumption is relaxed, in contrast with our multi-equation based results, the pure forward looking case [that is $\omega_f = 1$] cannot be refuted. 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 [that is when the restriction that the forward and backward looking terms sum up to one is relaxed], covers values exceeding one. Furthermore, values less than 0.5 cannot be refuted with a single equation method except with the HAC statistic and the standard gap measure (with which we rejected the model from a LI yet multi-equation 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 our single equation assumptions with the standard gap measure. With the real-time gap measure, we find quite fragile support for the rule, in the sense that results vary dramatically with the different considered 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 as 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 to single-

Table 2: Single Equation Structure-Restricted Confidence Sets

Inflation Equation; Intra-Model Instruments				
Coefficient	Standard Gap		Real-time Gap	
	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 estimated model is (2.6), equation by equation, ignoring contemporaneous correlation of disturbances. GAR refers to the single-equation generalized Anderson-Rubin method [refer to Stock and Wright (2000), Kleibergen and Mavroeidis (2009) and Dufour, Khalaf, and Kichian (2006, 2010a, 2010b)], which applies the same inference approach as the limited information presented in section 3.2, equation by equation. HAC refers to the Newey-West serial dependence correction.

equation methods, systems based estimation reveals useful information regarding ω_f, implying that the NKPC is conclusively more forward than backward looking, and that the pure forward looking case can be refuted. Such a conclusion cannot be reached with a single-equation approach.

Second, despite its fragility as a single equation, our LI system inference on the interest rate equation is quite informative. In contrast to the high estimate uncertainty we found with our single equation approaches, our LI systems-based confidence intervals on γ_π and γ_y are tightly centered around values compatible with Taylor (1993), particularly when an expanded instrument set is used. Recall that although systems-based, our LI method does not impose a unique equilibrium, and the latter assumption when imposed leads to rejecting the structure. 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 about underlying dynamics, which delivers precise estimates of feedback coefficients that are not at odds with the Taylor principle. Such a conclusion, again, cannot be reached with a single-equation approach.

Another point that is important to raise is the insignificance [using confidence intervals based on the statistics that have not led to model rejections] at the 5% level, of the parameter on the output gap term in the NKPC. We also find [again, using confidence intervals based on the statistics that have not led to model rejections] that the value of the parameter on the real interest rate in the output equation is 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 is not at odds with findings by, for example, Rudd and Whelan (2006) and Benati (2008). Broadly speaking, the survey of Schorfheide (2010) also suggests that insignificant coefficients are not unusual for forcing variables relative to available empirical studies.

While not uncommon, insignificant forcing variables in the NKPC and IS are an empirical puzzle. So far, available identification-robust evidence is this regard is restricted to the NKPC. Kleibergen and Mavroeidis (2009) apply partialled-out single equation statistics that under specific conditions [for example, imposing that one has accounted for all relevant instruments] may provide more powerful tests than projection-based methods, and yet cannot rule out a flat NKPC even with such statistics.[32] Magnusson and Mavroeidis (2010) also confirm this same 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 particularly

[32] We are 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.

because they document, via 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 finds a similar outcome. Perhaps equally importantly, we also find that the same puzzle plagues the IS equation.

While all issues raised by Schorfheide (2010) can be driving such findings, two possible interpretations can be suggested. First, it is indeed the case that the NKPC and the IS equations are flat, which is a dilemma that challenges New Keynesian theory. Second, the transmission mechanisms in the considered model are incomplete or misspecified so the 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, we find no empirical support for the considered output equation and very fragile evidence supportive of the considered interest rate equation, which lends credibility to the second interpretation. Our FI test suggests that the overall empirical model lacks support, which may also be viewed as a plausible - although radical - escape from this dilemma.

On balance, our results can be summarized as follows. Recall that we are evaluating the New Keynesian theory in conjunction with non-fundamental assumptions about underlying dynamics. 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 that allows us to learn something useful about the nature of the NKPC, and about the feedback terms in the Taylor rule regression.

With FI methods, assumptions exogenous to the theory must be taken as given, so it would be desirable to impose as few subsidiary restrictions as possible. These include restrictions on the regression disturbances, on the uniqueness of a rational expectation solution, and on the underlying information set.[33] Counterintuitively, complete-model based analysis of the New Keynesian theory rely on very strict such assumptions. Our FI method embeds a specification check, which formally rejects our FI assumptions, including the existence of a unique rational expectation solution. The model fares better when stability restrictions are relaxed, yet one important puzzle remains with the insignificance of forcing variables in the NKPC and the IS curve. 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, with focus on LI.

[33] See Dufour and Taamouti (2007) on problems resulting from closing the information set on identification-robust econometric methods. See also Stock (2010) for a recent perspective regarding "whimsical assumptions" in econometrics, defined as "assumptions subsidiary to the empirical purpose at hand, but which affect inference about the causal effect of interest".

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 estimated model is (2.6), equation by equation, ignoring contemporaneous correlation of disturbances and relaxing within-equation restriction. Estimation applies the Anderson-Rubin method from Dufour and Taamouti (2005).

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 we must ask 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, we propose econometric tools that can control statistical error whether the considered 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 considered model 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 set forth in this paper also contribute to the literature on the New Keynesian model. We estimate a standard three-equation model for the United States encompassing an NKPC, an IS curve and a Taylor rule, from 1962Q1 to 2005Q3. We impose and relax a unique rational expectation solution, maintaining similar lag-restrictions on regression disturbances in both cases. In the latter case, we also compare single-equation to multi-equation estimation and fit. We find that 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, our LI method generates realistic conclusions on the nature of the NKPC, and yields precise predictions for feedback coefficients that are not at odds with the Taylor principle. These results suggest that the unique rational expectations assumption is unduly restrictive for the considered model.

More broadly, we envision two possible uses for our proposed procedures. (1) 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. Then again, one can learn from our FI checks on how to overcome deficiencies in structures that lack fit. (2) Our LI method is useful in that it can utilize cross-equation information on the modeled 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 are - or must be - relaxed.

Appendix

A Structural LI statistics

Consider the multivariate regression (3.12) rewritten in stacked form as

$$\epsilon(\theta_0) = (I_n \otimes Z)b + v \tag{A.24}$$

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, $t = 1, ..., T$, $b = vec(\Pi')$ and

$$\epsilon(\theta_0) = (\epsilon_{\pi,1}(\mathrm{Y},\theta_0), \ \ldots, \epsilon_{\pi,T}(\mathrm{Y},\theta_0), \epsilon_{y,1}(\mathrm{Y},\theta_0), \ \ldots, \epsilon_{y,T}(\mathrm{Y},\theta_0), \ \epsilon_{R,1}(\mathrm{Y},\theta_0), \ \ldots, \epsilon_{R,T}(\mathrm{Y},\theta_0))'$$

is the nT-dimensional vector of structural errors evaluated at θ_0. In this context, the transformed null hypothesis H_{01}^* may be tested using the usual SURE-type F tests. We consider the statistic:

$$\mathcal{W}(\theta_0) = \left(\frac{n(T-k)}{m}\right) \frac{\left(A\hat{b}\right)' \left[A\left(\hat{\Sigma}_v \otimes (X'X)^{-1}\right)A'\right]^{-1}\left(A\hat{b}\right)}{(\epsilon(\theta_0) - (I_n \otimes X)\hat{b})'\left(\hat{\Sigma}_v^{-1} \otimes I_n\right)(\epsilon(\theta_0) - (I_n \otimes X)\hat{b})} \tag{A.25}$$

where $\hat{b}$ and $\hat{\Sigma}_v$ denote the unrestricted OLS coefficient and variance/covariance estimators from (A.24) and A is the selection matrix that captures (2.4). Observe that $\hat{b}$ and $\hat{\Sigma}_v$ depend on θ_0.[34] In our empirical model, $n = 3$, $k = 8 + q$ when q extra-model instruments are used, and A is the $m \times 3k$ selection matrix with $m = 2k + 4 + q$

$$A = \begin{bmatrix} A_\pi \\ A_y \\ A_R \end{bmatrix}, \quad \begin{array}{rc} A_\pi = & \begin{bmatrix} I_{(k)} & zeros(k,2k) \end{bmatrix} \\ A_y = & \begin{bmatrix} zeros(4+q,k) & \bar{A} & zeros(4+q,k) \end{bmatrix} \\ A_R = & \begin{bmatrix} zeros(k,2k) & I_{(k)} \end{bmatrix} \end{array}, \quad \bar{A} = \begin{bmatrix} I_{(2)} & 0 & 0 \\ 0 & 0 & I_{(q+3)} \end{bmatrix}.$$

The HAC statistic [see MacKinlay and Richardson (1991) and Ravikumar, Ray, and Savin (2000)] we use is

$$\begin{aligned} \mathcal{J}(\theta_0) &= T\,\hat{d}'D'\left[D\left(\left(\frac{X'X}{T}\right)^{-1} \otimes I_n\right)S_T\left(\left(\frac{X'X}{T}\right)^{-1} \otimes I_n\right)D'\right]^{-1} D\hat{d}, \\ S_T &= \Psi_{0,t} + \sum_{j=1}^{l}\left(\frac{l-j}{l}\right)\left[\Psi_{j,T} + \Psi_{j,T}'\right], \quad \Psi_{j,T} = \frac{1}{T}\sum_{t=j+1}^{T}(X_t \otimes \hat{v}_t)(X_{t-j} \otimes \hat{v}_{t-j})' \end{aligned} \tag{A.26}$$

[34] The statistic $\mathcal{W}$ corresponds to equation (10.11) of Srivastava and Giles (1987, Chapter 10) and to equation (49) in Dufour and Khalaf (2003, equation (49)).

where for convenience, $\hat{d}$ corresponds to $\hat{b}$ reshaped such that $d = vec(\Pi)$, D corresponds to a conformable reshaping of the selection vector A, and $\hat{v}_t$ is the three dimensional unrestricted OLS residual from (A.24). Results reported use $l = 4$. The asymptotic distribution of this statistic is χ^2 with degrees-of-freedom equal to the number of restrictions in D (here, $m = 2k + 4 + q$ as above).

B Invariance of LI method to missing instruments

The econometric model underlying our LI estimation of (2.6) can be rewritten as

$$\begin{aligned} \pi_t &= [\omega_f \pi_{t+1} + (1-\omega_f)\,\pi_{t-1} + \gamma y_t] + \epsilon_{\pi,t} \\ y_t &= [\beta_f y_{t+1} + \textstyle\sum_{j=1}^{4} \left(1-\beta_f\right) \beta_{y,j} y_{t-j} - \beta_r \left(R_t - \pi_{t+1}\right)] + \epsilon_{y,t} \\ R_t &= [\gamma_\pi \left(1 - \textstyle\sum_{j=1}^{3} \rho_j\right) \pi_t + \gamma_y \left(1 - \textstyle\sum_{j=1}^{3} \rho_j\right) y_t + \textstyle\sum_{j=1}^{3} \rho_j R_{t-j}] + \epsilon_{R,t} \end{aligned} \tag{B.27}$$

where $\epsilon_t = (\epsilon_{\pi,t}, \epsilon_{y,t}, \epsilon_{R,t})'$ is a zero-mean contemporaneously correlated disturbance vector that integrates rational expectation error. By conducting the test of (3.13) in the context of (3.12) as a test of (3.11), as described, we obtain a p-value that it is robust to the specification of the fundamental DGPs under consideration, to measurement errors and excluded instruments. Indeed, the test conducted in this framework only requires that the (unrestricted) reduced form for the system is given, up to an error term, by some function of: (i) the predetermined variables in the system [here, the considered lags of π_t, y_t and R_t], (ii) any extra instruments $\tilde{Z}_t$ used in the test, and (iii) possibly a set of further explanatory variables, denoted $\tilde{Q}_t$, which were not used in the test; these may include further lags of the endogenous variables, and/or further predetermined or exogenous variables that are omitted from the test, that is, are missing from the multivariate regression (3.12). To see this, suppose that the reduced form takes the unrestricted VAR specification

$$\begin{aligned} \pi_t &= a_\pi \pi_{t-1} + \textstyle\sum_{j=1}^{3} b_{\pi,j} R_{t-j} + \sum_{j=1}^{4} c_{\pi,j} y_{t-j} + \varpi_\pi' Q_t + \nu_{\pi,t}, \\ y_t &= a_y \pi_{t-1} + \textstyle\sum_{j=1}^{3} b_{y,j} R_{t-j} + \sum_{j=1}^{4} c_{y,j} y_{t-j} + \varpi_y' Q_t + \nu_{y,t}, \\ R_t &= a_R \pi_{t-1} + \textstyle\sum_{j=1}^{3} b_{R,j} R_{t-j} + \sum_{j=1}^{4} c_{R,j} y_{t-j} + \varpi_R' Q_t + \nu_{R,t}, \end{aligned} \tag{B.28}$$

where $Q_t = \left(\tilde{Z}_t', \tilde{Q}_t'\right)'$. It is straightforward to check that substituting the right-hand of (B.28) into the right-hand-side of (3.12) still leads, under the null hypothesis (3.11), to

$$\epsilon_{\pi,t}\left(\theta_0\right) = \epsilon_{\pi,t}, \quad \epsilon_{y,t}\left(\theta_0\right) = \sum_{j=1}^{4} \beta_{y,j}(1-\beta_f) y_{t-j} + \epsilon_{y,t}, \quad \epsilon_{R,t}\left(\theta_0\right) = \epsilon_{R,t}, \tag{B.29}$$

which justifies the tests we apply.

C Full-information statistic

Consider the multivariate regression (3.21) rewritten in matrix form as

$$U(\vartheta_0) = \boldsymbol{Z}\Pi + W \tag{C.30}$$

where $U(\vartheta_0)$ is the $T \times n$ matrix with row $\boldsymbol{U}_t(\mathrm{Y},\vartheta_0)'$, $\boldsymbol{Z}$ is the $T \times K$ matrix of predetermined variables in the system and W is the $T \times n$ matrix with row W_t'. Let

$$\begin{aligned} \hat{W} &= U(\vartheta_0) - Z\hat{\Pi}, \quad \hat{\Pi} = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'U(\vartheta_0), \\ \hat{W}_0 &= U(\vartheta_0), \end{aligned}$$

so $\hat{W}_0'\hat{W}_0$ and $\hat{W}'\hat{W}$ give the constrained [imposing $\Pi = 0$] and unconstrained sum of squared errors matrices from (C.30). The statistic we use is

$$\mathcal{L}(\phi_0) = \left(\frac{\mu\tau - 2\lambda}{Kn}\right) \frac{1 - \left(|\hat{W}'\hat{W}|/|\hat{W}_0'\hat{W}_0|\right)^{1/\tau}}{\left(|\hat{W}'\hat{W}|/|\hat{W}_0'\hat{W}_0|\right)^{1/\tau}}. \tag{C.31}$$

This statistic has been shown to perform well in the multivariate regression literature, with the following approximate null distribution

$$\mathcal{L}(\phi_0) \sim F(Kn\,,\, \mu\tau - 2\lambda) \tag{C.32}$$

$$\mu = T - K - \frac{(n - K + 1)}{2}\,,\ \lambda = \frac{nK - 2}{4} \tag{C.33}$$

$$\tau = \begin{cases} [(K^2n^2 - 4)/(K^2 + n^2 - 5)]^{1/2} & ,\ if\ \ K^2 + n^2 - 5 > 0 \\ 1 & ,\ otherwise. \end{cases} \tag{C.34}$$

$|\hat{W}'\hat{W}|/|\hat{W}_0'\hat{W}_0|$ is the well known Wilks statistic; see Dufour and Khalaf (2002a, 2002b, 2003). Observe that $\hat{W}$ and $\hat{W}_0$ depend on ϕ_0. In our empirical analysis, $n = 3$ and $K = 8$.

References

AN, S., AND F. SCHORFHEIDE (2007): "Bayesian Analysis of DSGE Models," *Econometric Reviews*, 26, 113–72.

ANDERSON, G. (2008): "Solving Linear Rational Expectations Models: A Horse Race," *Computational Economics*, 31, 95–113.

ANDERSON, G., AND G. MOORE (1985): "A linear algebraic procedure for solving linear perfect foresight models," *Economics Letters*, 17, 247–252.

ANDREWS, D. W. K., M. J. MOREIRA, AND J. H. STOCK (2006): "Optimal Two-Sided Invariant Similar Tests for Instrumental Variables Regression," *Econometrica*, 74, 715–752.

ANDREWS, I., AND A. MIKUSHEVA (2011): "Maximum Likelihood Inference in Weakly Identified DSGE Models," Discussion paper, Massachusetts Institute of Technology.

BEAULIEU, M.-C., J.-M. DUFOUR, AND L. KHALAF (2010): "Identification-Robust Estimation and Testing of the Zero-Beta CAPM," Discussion paper, 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., AND R. FARMER (2007): "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *American Economic Review*, 97, 524–29.

BLANCHARD, O., AND J. GALI (2007): "Real Wage Rigidities and the New Keynesian Model," *Journal of Money, Credit and Banking*, 39, 35–65.

——— (2010): "Labor Markets and Monetary Policy: A New Keynesian Model with Unemployment," *American Economic Journal - Macroeconomics*, 2, 1–30.

BOIVIN, J., AND M. GIANNONI (2006): "DSGE Models in a Data-Rich Environment," Discussion paper, NBER Working Paper No. 12772.

——— (2009): "Global Forces and Monetary Policy Effectiveness," in *International Dimensions of Monetary Policy*, ed. by J. Galí, and M. Gertler, chap. 8, pp. 429–478. University of Chicago Press., Chicago, U.S.

BOLDUC, D., L. KHALAF, AND E. MOYNEUR (2008): "Identification-Robust Simulation-Based Inference in Joint Discrete/Continuous Models for Energy Markets," *Computational Statistics and Data Analysis*, 52, 3148–3161.

BOLDUC, D., L. KHALAF, AND C. YELOU (2010): "Identification Robust Confidence Sets Methods for Inference on Parameter Ratios with Application to Discrete Choice Models," *Journal of Econometrics*, 157, 317–327.

CANOVA, F. (2007): *Methods for Applied Macroeconomic Research.* Princeton University press, New Jersey.

CANOVA, F., AND L. SALA (2009): "Back to Square One: Identification Issues in DSGE Models," *Journal of Monetary Economics*, 54, 431–49.

CHARI, V., T. KEHOE, AND E. MCGRATTAN (2009): "New Keynesian Models: Not Yet Useful for Policy Analysis," *American Economic Journal: Macroeconomics*, 1, 242–66.

CHAUDHURI, S., T. RICHARDSON, J. ROBINS, AND E. ZIVOT (2010): "A New Split-Sample Score Test in Linear Instrumental Variables Regression," *Econometric Theory*, 26, 1820–1837.

CHEVILLON, G., M. MASSMANN, AND S. MAVROEIDIS (2010): "Inference in models with adaptive learning," *Journal of Monetary Economics*, 57, 341–351.

CHRISTIANO, L., M. EICHENBAUM, AND M. EVANS (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economics*, 113, 1–45.

CLARIDA, R., J. GALÍ, AND M. GERTLER (1999): "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature*, 37, 1661–1707.

——— (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., C. FAVERO, AND A. PACCAGNINI (2009): "On the Statistical Identification of DSGE Models," *Journal of Econometrics*, 150, 99–115.

CURDIA, V., AND R. REIS (2011): "Correlated Disturbances and U.S. Business Cycles," Discussion paper, Columbia University.

Del Negro, M., F. Schorfheide, F. Smets, and R. Wouters (2007): "On the Fit of New Keynesian Models," *Journal of Business and Economic Statistics*, 25, 123–43.

Doko-Tchatoka, F., and J.-M. Dufour (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.

——— (2003): "Identification, Weak Instruments and Statistical Inference in Econometrics," *Canadian Journal of Economics*, 36, 767–808.

Dufour, J.-M., and J. Jasiak (2001): "Finite Sample Limited Information Inference Methods for Structural Equations and Models with Generated Regressors," *International Economic Review*, 42, 815–843.

Dufour, J.-M., and L. Khalaf (2002a): "Exact Tests for Contemporaneous Correlation of Disturbances in Seemingly Unrelated Regressions," *Journal of Econometrics*, 106, 143–170.

——— (2002b): "Simulation Based Finite and Large Sample Tests in Multivariate Regressions," *Journal of Econometrics*, 111, 303–322.

——— (2003): "Finite Sample Tests in Seemingly Unrelated Regressions," in *Computer-Aided Econometrics*, ed. by D. E. A. Giles, chap. 2, pp. 11–35. Marcel Dekker, New York.

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

——— (2010a): "Estimation Uncertainty in Structural Inflation Models with Real Wage Rigidities," *Computational Statistics and Data Analysis*, 54, 2554–2561.

——— (2010b): "On the Precision of Calvo Parameter Estimates in Structural NKPC Models," *Journal of Economic Dynamics and Control*, 34, 1582–1595.

Dufour, J.-M., and M. Taamouti (2005): "Projection-Based Statistical Inference in Linear Structural Models with Possibly Weak Instruments," *Econometrica*, 73, 1351–1365.

——— (2007): "Further Results on Projection-Based Inference in IV Regressions with Weak, Collinear or Missing Instruments," *Journal of Econometrics*, 139, 133–153.

Eichenbaum, M., and J. Fisher (2007): "Estimating the Frequency of Price Re-optimization in Calvo-Style Models," *Journal of Monetary Economics*, 54, 2032–47.

Fair, R. (2008): "Testing Price Equations," *European Economic Review*, 52, 1424–1437.

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

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

Galí, J., M. Gertler, and J. D. Lopez-Salido (2005): "Robustness of the Estimates of the Hybrid New Keynesian Phillips Curve," *Journal of Monetary Economics*, 52, 1107–18.

Goffe, W. L., G. D. Ferrier, and J. Rogers (1994): "Global Optimization of Statistical Functions with Simulated Annealing," *Journal of Econometrics*, 60, 65–99.

Granziera, E., M. Lee, H. R. Moon, and F. Schorfheide (2011): "Inference for VARs Identified with Sign Restrictions," Discussion paper, University of Pennsylvania.

Guerron-Quintana, P. (2010): "What you match does matter: the effect of data on DGSE estimation," *Journal of Applied Econometrics*, 25, 774–804.

Guerron-Quintana, P., A. Inoue, and L. Kilian (2009): "Inference in Weakly Identified DSGE Models," Discussion paper, North Carolina State University and University of Michigan.

Gungor, S., and R. Luger (2009): "Exact distribution-free tests of mean-variance efficiency," *Journal of Empirical Finance*, 16, 816–829.

Hannan, E. (1971): "The Identification Problem for Multiple Equation Systems with Moving Average Errors," *Econometrica*, 39, 751–65.

Hodges, Jr, J. L., and E. L. Lehmann (1963): "Estimates of Location Based on Rank Tests," *The Annals of Mathematical Statistics*, 34, 598–611.

Hodges, Jr, J. L., and E. L. Lehmann (1983): "Hodges-Lehmann Estimators," in *Encyclopedia of Statistical Sciences, Volume 3*, ed. by N. L. Johnson, S. Kotz, and C. Read, pp. 642–645. John Wiley & Sons, New York.

Hoogerheide, L. F., J. F. Kaashoek, and H. K. van Dijk (2007): "On the shape of posterior densities and credible sets in instrumental variable regression models with reduced rank: An application of flexible sampling methods using neural networks," *Journal of Econometrics*, 139, 154–180.

Inoue, A., and B. Rossi (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.

Jondeau, E., and H. Le Bihan (2008): "Examining bias in estimators of linear rational expectations models under misspecification," *Journal of Econometrics*, 143, 375–395.

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., and M. W. Watson (1998): "The solution of singular linear difference systems under rational expectations," *International Economic Review,*, 39, 1015–1026.

Kleibergen, F. (2002): "Pivotal Statistics for Testing Structural Parameters in Instrumental Variables Regression," *Econometrica*, 70, 1781–1803.

——— (2005): "Testing Parameters in GMM Without Assuming that They Are Identified," *Econometrica*, 73, 1103–1123.

Kleibergen, F., and S. Mavroeidis (2009): "Weak Instrument Robust Tests in GMM and the New Keynesian Phillips Curve," *Journal of Business and Economic Statistics*, 27, 293–311.

Kleibergen, F., and S. Mavroeidis (2010): "Inference on subsets of parameters in GMM without assuming identification," Discussion paper, Working paper, Brown University.

Komunjer, I., and S. Ng (2011): "Dynamic Identification of DSGE Models," *Econometrica*, forthcoming.

Koopmans, T. C. (1950): *Statistical Inference in Dynamic Economic Models, Cowles Comission Monograph, Volume 10.* John Wiley & Sons, New York.

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., and M. P. Richardson (1991): "Using Generalized Method of Moments to Test Mean-Variance Efficiency," *The Journal of Finance*, 46, 511–527.

Magnusson, L., and S. Mavroeidis (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.

——— (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., and F. Shorfheide (2010): "Bayesian and Frequentist Inference in Partially Identified Models," Discussion paper, University of Pennsylvania.

Moreira, M. J. (2003): "A Conditional Likelihood Ratio Test for Structural Models," *Econometrica*, 71, 1027–1048.

Nason, J., and G. Smith (2008): "Identifying the New Keynesian Phillips Curve," *Journal of Applied Econometrics*, 23, 525–51.

Paccagnini, A. (2011): "Model Validation in the DSGE Approach: A Survey," Working paper, Bocconi University.

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

Ravikumar, B., S. Ray, and E. Savin (2000): "Robust Wald Tests in SUR Systems with Adding-up Restrictions," *Econometrica*, 68, 715–720.

Ray, S., and E. Savin (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.

Rios Rull, J. V., F. Schorfheide, C. Fuentes-Albero, M. Kryshk, and R. SantaEulalia-Llopis (2011): "Methods versus Substance: Measuring the Effects of Technology Shocks," Discussion paper, University of Pennsylvania.

Rudd, J., and K. Whelan (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, 397433.

——— (2010): "Estimation and Evaluation of DSGE Models: Progress and Challenges," Discussion paper, 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., and R. Wouters (2003): "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, 1, 1123–75.

——— (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97, 586–606.

Srivastava, V. K., and D. E. Giles (1987): *Seemingly Unrelated Regression Equations Models. Estimation and inference*. Marcel Dekker, New York.

Staiger, D., and J. H. Stock (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65, 557–586.

Stock, J. H. (2010): "The Other Transformation in Econometric Practice: Robust Tools for Inference," *The Journal of Economic Perspectives*, 24, 83–94.

Stock, J. H., and J. H. Wright (2000): "GMM with Weak Identification," *Econometrica*, 68, 1097–1126.

Stock, J. H., J. H. Wright, and M. Yogo (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," Discussion paper, Carnegie-Rochester Conference Series on Public Policy 39.

Wang, J., and E. Zivot (1998): "Inference on Structural Parameters in Instrumental Variable Regression with Weak Instruments," *Econometrica*, 66, 1389–1404.

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

Zellner, A., and F. Palm (1974): "Time Series Analysis and Simultaneous Equation Econometric Models," *Journal of Econometrics*, 2, 17–54.

Zivot, E., R. Startz, and C. R. Nelson (1998): "Valid Confidence Intervals and Inference in the Presence of Weak Instruments," *International Economic Review*, 39, 1119–1144.