BANK OF CANADA
BANQUE DU CANADA

Working Paper/Document de travail
2009-19

Structural Multi-Equation Macroeconomic Models: Identification-Robust Estimation and Fit

by Jean-Marie Dufour, Lynda Khalaf, and Maral Kichian

www.bank-banque-canada.ca

Bank of Canada Working Paper 2009-19

June 2009

Structural Multi-Equation Macroeconomic Models: Identification-Robust Estimation and Fit

by

Jean-Marie Dufour,[1] Lynda Khalaf,[2] and Maral Kichian[3]

[1]Department of Economics
McGill University
Montréal, Quebec, Canada H3A 2T7
and CIRANO, CIREQ
jean.marie.dufour@mcgill.ca

[2]Economics Department
Carleton University
Ottawa, Ontario, Canada K1S 5B6
and CIREQ, GREEN
Lynda_Khalaf@carleton.ca

[3]Canadian Economic Analysis Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
mkichian@bankofcanada.ca

Bank of Canada working papers are theoretical or empirical works-in-progress on subjects in economics and finance. The views expressed in this paper are those of the authors. No responsibility for them should be attributed to the Bank of Canada.

ISSN 1701-9397 © 2009 Bank of Canada

Acknowledgements

We would like to thank Vadim Marmer, Sophocles Movroeidis, Rhys Mendes, Ulrich Müller, and conference participants at the 2006 NBER and NSF Time Series Conference, the 2006 European Conference of the Econometrics Community (EC^2), the 2007 Computational and Financial Econometrics conference, the 2007 Canadian Economic Association meetings, as well as participants at the Bank of Canada seminar series and the 2007 Carleton University Structural Modeling and Weak Identification workshop for helpful suggestions and comments. Timothy Grieder 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).

Abstract

Weak identification is likely to be prevalent in multi-equation macroeconomic models such as in dynamic stochastic general equilibrium setups. Identification difficulties cause the breakdown of standard asymptotic procedures, making inference unreliable. While the extensive econometric literature now includes a number of identification-robust methods that are valid regardless of the identification status of models, these are mostly limited-information-based approaches, and applications have accordingly been made on single-equation models such as the New Keynesian Phillips Curve.

In this paper, we develop a set of identification-robust econometric tools that, regardless of the model's identification status, are useful for estimating and assessing the fit of a system of structural equations. In particular, we propose a vector auto-regression (VAR) based estimation and testing procedure that relies on inverting identification-robust multivariate statistics. The procedure is valid in the presence of endogeneity, structural constraints, identification difficulties, or any combination of these, and also provides summary measures of fit. Furthermore, it has the additional desirable features that it is robust to missing instruments, errors-in-variables, the specification of the data generating process, and the presence of contemporaneous correlation in the disturbances.

We apply our methodology, using U.S. data, to the standard New Keynesian model such as the one studied in Clarida, Gali, and Gertler (1999). We find that, despite the presence of identification difficulties, our proposed method is able to shed some light on the fit of the considered model and, particularly, on the nature of the NKPC. Notably our results show that (i) confidence intervals obtained using our system-based approach are generally tighter than their single-equation counterparts, and thus are more informative, (ii) most model coefficients are significant at conventional levels, and (iii) the NKPC is preponderantly forward-looking, though not purely so.

JEL classification: C52, C53, E37
Bank classification: Inflation and prices; Econometric and statistical methods

Résumé

Les modèles macroéconomiques à équations multiples, comme les modèles d'équilibre général dynamiques et stochastiques, tendent à donner lieu à des problèmes d'identification qui compromettent l'usage de techniques asymptotiques standard et la fiabilité de l'inférence statistique. Si l'abondant corpus de travaux économétriques propose aujourd'hui plusieurs méthodes robustes en matière d'identification qui gardent leur validité que le modèle soit bien ou mal identifié, ces méthodes supposent néanmoins souvent une information incomplète. Leur application s'est par conséquent trouvée limitée à des modèles à équation unique tels que la nouvelle courbe de Phillips keynésienne.

Les auteurs élaborent un ensemble d’outils économétriques qui permet d’estimer et d’évaluer la qualité de l’ajustement d’un système d’équations structurelles peu importe les conditions d’identification de ce dernier. Ils proposent notamment une procédure d’estimation et de test qui fait appel à une autorégression vectorielle et inverse le résultat des tests d’inférence robuste de type multivarié. Cette procédure est valide qu’il y ait endogénéité, contrainte structurelle ou problème d’identification, ou encore une combinaison quelconque de ces éléments. Elle offre par ailleurs des mesures sommaires de l’adéquation statistique. Elle conserve sa validité en l’absence de certains instruments, en présence d’erreurs sur les variables ou de corrélation contemporaine des perturbations et peu importe la spécification du processus générateur de données.

À l’aide de données américaines, les auteurs appliquent leur méthode à une variante du nouveau modèle keynésien type analogue à celle analysée par Clarida, Gali et Gertler (1999). En dépit des problèmes d’identification, leur approche cerne mieux l’adéquation statistique du modèle étudié ainsi que, en particulier, la nature de la nouvelle courbe de Phillips keynésienne. Les résultats montrent, d’une part, que le recours à un système d’équations multiples permet d’obtenir des intervalles de confiance plus informatifs car généralement plus étroits que ceux issus de modèles à équation unique; d’autre part, que la majorité des coefficients sont significatifs aux niveaux habituels; enfin, que sans être totalement prospective, la nouvelle courbe de Phillips keynésienne l’est très fortement.

Classification JEL : C52, C53, E37
Classification de la Banque : Inflation et prix; Méthodes économétriques et statistiques

1. Introduction

Optimization-based macroeconomic models, and, in particular, dynamic stochastic general equilibrium (DSGE) setups, are popular nowadays for analyzing a multitude of macroeconomic questions such as the effects of monetary policy. But as models of this sort become increasingly complex, featuring many types of markets, various rigidities, and different nonlinearities, the decision of whether to use a *limited* or *full* information (*LI* or *FI*) approach for estimation becomes a central question for model developers. Indeed, there appears to be a conflict in the conclusions of available published studies based on one or the other method; for instance, Galí, Gertler, and Lopez-Salido (2005), and Linde (2005) report opposite outcomes with regard to the importance of the forward-looking component of the New Keynesian Phillips Curve (NKPC) equation.

The LI/FI trade-off is an enduring econometric problem, often presented as one of weighing specification bias versus efficiency, but there are also other concerns. In particular, advances in econometrics regarding weak-instruments and weak-identification have revealed that the latter plague LI and FI methods equally, thus presenting a set of new challenges for applied researchers.

The macroeconomic literature acknowledges the LI/FI trade-off to some extent, often presenting it as one of deciding between Instrumental Variable (IV) or maximum likelihood estimation (MLE). Furthermore, published studies in the field are also familiar with the fact that weak instruments effects are critical to IV-based model performance. However, the implications of weak-identification on MLE seem to be less understood, and indeed often confused with issues related to very large estimated standard errors or poorly-approximated test statistics cut-off points. While it may be argued that likelihood-ratio (LR) criteria have more attractive finite sample properties than, for example, IV-based Wald-type ones, and in particular, size correction techniques have a much better chance of success with LR statistics (see Dufour 1997), it should be emphasized that standard MLE and full-information maximum likelihood (FIML) inference are not immune to weak-identification problems.

The complications arise largely because nonlinearities can impose discontinuous parameter restrictions that cause the breakdown of standard asymptotic procedures. Given the connection between the parameters of the underlying theoretical model and those of the estimated econometric model[1], and given the identifying constraints imposed on the model,

[1]See Galí, Gertler, and Lopez-Salido (2005) and Fernandez-Villaverde, Rubio-Ramirez, and Sargent (2005) on the importance of maintaining these constraints.

econometric versions of macroeconomic models are often highly nonlinear.[2] The more rich and complex the macroeconomic model, the more likely it is that standard regularity conditions will not fully hold. In this case, even when MLE is used for the estimation, resorting to usual t-type significance tests or Wald-type confidence intervals will lead to the same problems that plague GMM and linear or nonlinear IV;[3] see the surveys of Stock, Wright, and Yogo (2002) and Dufour (2003). As may be checked from these studies, identification difficulties will not always lead to huge regular standard errors that would alert the researcher to the problem. Instead, spuriously tight confidence intervals could occur, often concentrated on wrong parameter values, thus leading to wrong inference.

Weak-instruments and weak-identification concerns have led to the development of so-called *identification-robust* procedures, *i.e.* procedures that achieve significance or confidence-level control (at least asymptotically) whether the statistical model is weakly or strongly identified, or whether instruments are weak or strong.[4] To a certain extent, and within the context of single-equation models, such procedures are gaining credibility in macroeconomics, although some of the findings of these studies challenge the fit of popular models including the NKPC; see, for example, Mavroeidis (2004), Mavroeidis (2005) and Dufour, Khalaf, and Kichian (2006). Yet, despite the considerable volume of the associated econometric literature, identification-robust methods for multi-equation systems are still scarce (including the literature on GMM cited above which is sufficiently general to cover systems of equations) compared to methods that are available for single-equation models. Thus, it is not surprising that, in applied work, studies have addressed possible weak identification relying on single-equation approaches.

In this paper, we look at whether proponents of full information estimation are justified in their claims, and examine how well the approach stands up to the weak-identification test. In this regard, our contribution is twofold: one methodological, and one substansive.

First, we develop a set of identification-robust econometric tools that are useful for estimating and assessing the fit of a system of structural equations. In particular, we

[2] Even within the context of a single linear simultaneous equation, where identification is achieved through "exclusion" restrictions, the latter imply nonlinearity. This is easy to see when one derives the reduced-form or the structural likelihood function.

[3] We specifically mean Wald-type confidence intervals of the form [estimate $\pm$(asymptotic standard error) $\times$ (asymptotic critical point)], intervals based on the *delta*-method, and even ones based on various bootstraps.

[4] See, for example, Dufour (1997), Dufour (2003), Staiger and Stock (1997), Wang and Zivot (1998), Zivot, Startz, and Nelson (1998), Dufour and Jasiak (2001), Kleibergen (2002), Kleibergen (2005), Stock, Wright, and Yogo (2002), Moreira (2003), Dufour and Taamouti (2005), Dufour and Taamouti (2007), and Andrews, Moreira, and Stock (2006).

propose a vector auto-regression (VAR) based estimation and testing procedure that relies on identification-robust multivariate statistics. The procedure is valid (in the sense of significance-level control) in the presence of endogeneity, structural constraints, identification difficulties, or any combination of these, and also provides summary measures of fit. Furthermore, it has the additional desirable features that it is robust to missing instruments, errors-in-variables, the presence of contemporaneous correlation in the disturbances, and the specification of the data generating process (DGP).[5] Finally, these advantages hold while the constraints on the parameters and/or error terms implied by the underlying theoretical model are formally taken into account.

The methodology works through the combined use of the econometric model for the structure, and an unrestricted VAR for the instrumental underlying data generating process. More specifically, and, in the case of, for example, a DSGE framework, the five key components of our method are: (1) an instrumental model (the VAR); (2) a structural general equilibrium model (underlying theory); (3) an econometric multi-equation linearized model that links the former to the latter (the estimable DSGE structure) allowing for possible measurement error (the data), (4) multivariate statistics that summarize the information combining these three components (the pivots), and (5) multivariate measures of model fit related to the latter statistics (the J-type criteria).

Second, we apply these tools, using U.S. data, to assess the fit of the standard New Keynesian model. This fundamental structure has been extensively studied in the literature (see, for example, Clarida, Gali, and Gertler 1999), and forms the building block of many other more complex models (see, for instance, Woodford (2003), Christiano, Eichenbaum, and Evans (2005), Del Negro, Schorfheide, Smets, and Wouters (2007), to mention a few.) To allow for comparisons between our newly-proposed system-based multivariate method and univariate ones, we also consider the univariate method applied by Dufour, Khalaf, and Kichian (2006), and the univariate linear IV method from Dufour and Taamouti (2005). Each method integrates and assesses, to a different degree, the model's structural restrictions.

The empirical results may be summarized as follows. Although identification problems are present, our proposed method is able to shed some light on the fit of the considered model and, particularly, on the nature of the NKPC. In particular, we find that (i) multi-equation confidence intervals are generally tighter than their univariate counterparts, and thus more informative, (ii) most model coefficients are significant at conventional levels, with the exception two, and that (iii) the NKPC is preponderantly forward-looking, though not

[5] Unlike, for example, Nason and Smith (2003) or Linde (2005), we do not need to specify the full DGP. In other words, our method of evaluation is system-based but does necessarily have to be strictly FI.

purely so.

The paper is organized as follows. In section 2, we introduce the model we assess. Our methodology is discussed in section 3. Data and empirical results are presented in section 4. We conclude in section 5. Finally, a technical Appendix complements the methodology section.

2. Framework

Though our method is applicable to more complex structures, we consider here a variant of the standard New Keynesian model. The latter, extensively studied by Clarida, Gali, and Gertler (1999), forms the building block of numerous recent fundamental models, and for our purposes is tractable enough to allow comparisons between our proposed multivariate approach and available univariate ones (see later sections).

Specifically, we follow the setup in Linde (2005) that consists of a system of three equations: an NKPC equation, an aggregate demand equation and an interest rate rule:

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

where, for $t = 1, \ ..., \ T$, π_t is aggregate inflation, y_t is the output gap, and R_t is the nominal interest rate, and $\varepsilon_{\pi,t}$, $\varepsilon_{\pi,t}$ and $\varepsilon_{R,t}$ are random disturbances. The parameter constraints reflect an underlying macroeconomic model. For notational clarity, we will call the vector

$$\theta = \left(\begin{array}{ccccccccc} \omega_f, & \gamma, & \beta_f, & \beta_r, & \gamma_\pi, & \gamma_y, & \rho_1, & \rho_2, & \rho_3 \end{array} \right)'$$

the model's "deep" parameters.

For estimation purposes, we consider the econometric model

$$\begin{array}{c} \pi_t = \omega_f \pi_{t+1} + (1-\omega_f)\,\pi_{t-1} + \gamma y_t + \epsilon_{\pi,t} \\ y_t = \beta_f y_{t+1} + \sum_{i=1}^{4} (1-\beta_f)\,\beta_{y,i} y_{t-i} - \beta_r\,(R_t - \pi_{t+1}) + \epsilon_{y,t} \\ R_t = \gamma_\pi \left(1-\sum_{i=1}^{3}\rho_i\right)\pi_t + \gamma_y \left(1-\sum_{i=1}^{3}\rho_i\right) y_t + \sum_{i=1}^{3}\rho_i R_{t-i} + \epsilon_{R,t} \end{array} \tag{2}$$

where, due to the rational expectation hypothesis, the error terms now integrate expectation error. In this respect, even though Linde (2005) assumes a diagonal covariance matrix, we allow for possible contemporaneous error cross-correlations.

Assuming Gaussian errors, the model is readily estimable via LI or FI maximum likelihood, and parameter estimates, standard errors, as well as regular LR-type test criteria (for assessing the constrained model against, say, an unrestricted VAR), can all be easily derived. However, if the confidence interval and hypothesis tests that result from such 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.[6]

It is important to understand the fundamental reason behind such failures.[7] Nonlinear constraints complicate statistical analysis in a non-trivial way because associated transformations may be discontinuous. That is, some or all of the parameters may become identifiable only on a subset of the parameter space. In such contexts, in order to have good statistical coverage, any valid method for constructing a confidence set (CS) should allow for possibly-unbounded outcomes. Stated differently, any method that, by construction, leads to a confidence *interval* with *bounded limits*, will necessarily have poor coverage (Dufour 1997).[8] Therefore, intervals of the form {estimate ± (asymptotic standard error) × (asymptotic critical point)}, including the delta-method, are fundamentally wrong and cannot be size-corrected. Furthermore, identification difficulty does not necessarily imply that asymptotic approximations to critical points are poor, or that asymptotic standard errors are large. Indeed, the opposite may occur, with tight confidence intervals concentrated on wrong parameter values.

Identification-robust methods typically rely on appropriate pivots, *i.e.* statistics whose null distributions are invariant to the model's identification status. In particular, generalized Anderson-Rubin procedures that involve *inverting* proper pivotal tests are considered.[9] Inverting a test yields the set of parameter values that are not rejected by this test. The geometrics of such inversions typically allow for unbounded solutions—a pre-requisite for ensuring reliable coverage. While a large econometric literature has documented the superiority of such methods, multi-equation models have not been directly addressed.

Here we propose a multivariate extension of the Anderson-Rubin test, that when inverted, will yield a CS whose significance level can be controlled (at least asymptotically) in the presence of endogeneity and nonlinear parameter constraints, whether identification is weak or not. Inverting this test numerically produces the set of parameter values that are not rejected by this test, and the least-rejected parameters are the so-called Hodges-Lehmann

[6] Regularity conditions do not hold or hold only weakly when there are identification difficulties.

[7] Please refer to the econometric literature cited in the introduction for further formal discussions.

[8] It is shown that the method that proves the validity of confidence *intervals* typically excludes the parameter discontinuity regions entailed by the nonlinear functions under consideration.

[9] See, for example, Dufour (1997), Dufour (2003), Staiger and Stock (1997), Wang and Zivot (1998), Zivot, Startz, and Nelson (1998), Dufour and Jasiak (2001), Dufour and Taamouti (2005), and Dufour and Taamouti (2007).

point estimates (see Hodges and Lehmann 1963, 1983, and Dufour, Khalaf, and Kichian 2006). The test inversion may also generate an empty CS. This can be interpreted as a significant J-type test, providing an overall assessment of the structural model restrictions.

3. Methodology

In this section, we describe the methodology as it applies to model (2). For clarity of presentation our discussion is mostly descriptive but nonetheless formal; complete formulae and further references are relegated to the Appendix.

3.1 Transforming the Regression

To obtain a confidence set with level $1-\alpha$ for the deep parameter θ, we invert an identification-robust test (presented below) associated with the null hypothesis

$$H_0: \ \theta = \theta_0 \tag{3}$$

where θ_0 is given by

$$\theta_0 = \left(\ \omega_f^0, \ \ \gamma^0, \ \ \beta_f^0, \ \ \beta_r^0, \ \ \gamma_\pi^0, \ \ \gamma_y^0, \ \ \rho_1^0, \ \ \rho_2^0, \ \ \rho_3^0 \ \right)',$$

and where the parameter values with the zero superscript are known values. Formally, this implies collecting the values θ_0 that are not rejected by the test (i.e., for which the test is not significant at level α). In what follows, we first introduce the test that is inverted and then explain how the former step is performed. Let

$$Z_t = (Z_{1t}, Z_{2t})', \quad Z_{1t} = (\pi_{t-1}, R_{t-1}, R_{t-2}, R_{t-3})', \quad Z_{2t} = (y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4})'.$$

Z_t so defined consists of all predetermined variables in the system, and which we denote as the set of "internal" instruments to reflect their model dependence. We also consider a set of q additional "external" instruments, denoted $\widetilde{Z}_t$, that we use to correct for measurement errors.

Now consider the transformed regression, which, in reference to the univariate econometric literature on the Anderson-Rubin test, we call the Anderson-Rubin Multivariate Linear Regression (AR-MLR):

$$\begin{aligned} &\pi_t^* \ \text{ on } \{Z_t \text{ and } \widetilde{Z}_t\}, \qquad (4)\\ &y_t^* \text{ on } \{Z_t \text{ and } \widetilde{Z}_t\},\\ &R_t^* \ \text{ on } \{Z_t \text{ and } \widetilde{Z}_t\}, \end{aligned}$$

where

$$\begin{aligned} \pi_t^* &= \pi_t - \omega_f^0 \pi_{t+1} - (1-\omega_f^0)\pi_{t-1} - \gamma^0 y_t, \\ y_t^* &= y_t - \beta_f^0 y_{t+1} + \beta_r^0 \left(R_t - \pi_{t+1}\right), \\ R_t^* &= R_t - \left(1 - \sum_{i=1}^{3} \rho_i^0\right)\left(\gamma_\pi^0 \pi_t + \gamma_y^0 y_t\right) - \sum_{i=1}^{3} \rho_i^0 R_{t-i}. \end{aligned} \tag{5}$$

Our notation assumes that the three equations in (4) are treated as a system, allowing for error cross-correlations. Under the null hypothesis [specifically (2)-(3)], the coefficients of Z_t and $\widetilde{Z}_t$ in the first and last equations, and of Z_{1t} and $\widetilde{Z}_t$ in the middle equation of (4) should be zero. Hence, testing for such a zero null hypothesis on these coefficients provides a test of (3). The intuition is simple: the structural equation (2) that faces identification difficulties is mapped, through our approach, into the standard regression (4). The latter constitutes a regular framework where identification constraints are no longer needed because the right-hand side regressors are not "endogenous". Therefore, usual statistics for testing the exclusion of regressors can be applied in a straightforward manner.

3.2 An Identification-Robust Test

The test criterion that we use is one of the most popular statistics in SURE analysis (see Dufour and Khalaf 2003 and the references therein). Specifically, we consider the SURE-F criterion denoted $\mathcal{W}$ described in equation (10) of the Appendix. We can obtain a valid p-value for $\mathcal{W}$ using an F asymptotic distribution.[10] $\mathcal{W}$ applied to (4) is asymptotically pivotal whether (2) is weakly or strongly identified, and its asymptotic distribution is standard, depending on the sample size and on the number of predetermined variables and instruments used in the test. No further nuisance parameters intervene, and in particular, the asymptotic null distribution does not depend on the unknown variance-covariance matrix. This result obtains because the statistical reduced form AR-MLR (4) allows the test problem to be conducted within the classical multivariate linear regression statistical framework. The latter does not require any identification constraints in contrast to the original simultaneous equation system (2) that does require them. Because $\mathcal{W}$, as applied to (4), is asymptotically pivotal irrespective of the model's identification status—a property that is not shared by IV-based Wald statistics and by GMM-based J-tests—, the confidence set for θ that is obtained

[10]The $\mathcal{W}$ statistics were analyzed by Dufour and Khalaf (2003). In a system with three equations, its F-based asymptotic approximation was shown to be relatively more stable (in terms of size control) than the χ^2 counterpart. In this paper, we rely on that result.

by inverting this test will have a correct asymptotic level whether (2) is weakly or strongly identified. Our approach thus provides an attractive solution to identification difficulties.

In addition, and as with single equation Anderson-Rubin type methods, our procedure has two further "built-in" advantages. First, relatively wide confidence sets reveal weak identification. Second, if the confidence set is empty at some chosen significance level (which occurs when all economically-relevant values of the model's deep parameters are rejected at this level), then the model can soundly be rejected. This provides an identification-robust alternative to the standard GMM-based J-test. Formally, observe that the cut-off points for the $\mathcal{W}$ statistic introduced above are the same for any value θ_0 under test. As may be checked from the Appendix, the null distribution of $\mathcal{W}$ depends on the sample size, the number of equations and the number of constraints, but not on θ_0 *per se*. Taking the model in Section 2 as an example, the approximate limiting null distribution for the $\mathcal{W}$ statistic is $F(m, 3(T-k))$ with

$$m = 2(8+q) + (4+q)$$

where q as defined above is the number external instruments [if any] used and $k = 8+q$ is the number of regressors per equation in (4). So if we define

$$\overline{\mathcal{W}} = \min_{\theta_0} \mathcal{W},$$

referring the latter to an $F(m, 3(T-k))$ cut-off point (say at level α) provide an identification robust J-test, since

$$\min_{\theta_0} \mathcal{W} \geq F_\alpha(m, 3(T-k)) \Leftrightarrow \mathcal{W} \geq F_\alpha(m, 3(T-k)), \quad \forall \theta_0 \tag{6}$$

where $F_\alpha(.)$ denotes the α-level cut-off point under consideration. In other words, (6) implies that the $F(m, 3(T-k))$ distribution provides valid and identification-robust conservative bounds on the null distributions of $\overline{\mathcal{W}}$.

The latter specification check can be carried out before the test inversion step to save computation time; if the outcome is not significant [*i.e.* if $\min_{\theta_0} \mathcal{W} < F_\alpha(m, 3(T-k))$], then we can be sure that the associated confidence sets for θ will not be empty. Such specification tests can clearly be very useful tools for modelers, whether they are applied on their own or in conjunction with the test inversion problem. In view of the underlying nonlinearity, the latter minimizations must be performed numerically. We recommend a global optimization procedure such as Simulated Annealing because there is no reason to expect that $\mathcal{W}$ is a smooth function of θ_0.

3.3 Test Inversion Procedure

The test inversion procedure that we present in this section must also be conducted numerically. We suggest and apply two such procedures: First, using a grid search over the economically-meaningful set of values for θ, we sweep the choices for θ_0; as is illustrated in the Appendix, we do not need to consider the unknown variance-covariance matrix of disturbances as a nuisance parameter. For each choice considered, we compute test statistics and their associated p-values. The parameter vectors for which the p-values are greater than the level α thus constitute a confidence set with level $1 - \alpha$.

Alternatively, it is possible to construct projection-based confidence sets. These can be obtained for any linear combination of θ, of the form $a'\theta$ where a is a non-zero vector, by minimizing and maximizing (for example using simulated annealing) the function $a'\theta$ over θ such that $\mathcal{W} < \chi^2_\alpha(m)$. Components of θ are defined by setting a to the corresponding selection vector (consisting of zeros and ones).[11]

To find point estimates within our CS, we look for the values of θ_0 that lead to the largest p-value. These values are the most compatible with the data, or, alternatively, correspond to the "least rejected" model. Such an approach underlies the principles of the Hodges-Lehmann estimation method; see Hodges and Lehmann (1963); Hodges and Lehmann (1983). Whereas uniqueness (as obtained through the usual point estimation approach) is not granted, analyzing the economic information content of these least rejected models provides very useful model diagnostics.

3.4 Other Advantages

It is important to note that our method automatically corrects for errors-in-variable (generated regressor) problems under the same maintained assumptions on the reduced form. Furthermore, missing instruments will not invalidate our fundamental results. In other words, if our test does not account for all explanatory variables that define the reduced form, the significance level will not be affected. This also means that a full definition of the fundamental DGP is not required. These properties hold while the structural implications of the underlying theoretical model are maintained.

Similarly, note that the added instruments $\widetilde{Z}_t$ are not strictly necessary, since, in the

[11] For a description of a similar procedure in a univariate setting, see e.g. Dufour and Jasiak (2001).

context of the regression:

$$\begin{aligned} \pi_t^* &\quad \text{on} \quad Z_t, \\ y_t^* &\quad \text{on} \quad Z_t, \\ R_t^* &\quad \text{on} \quad Z_t, \end{aligned} \tag{7}$$

we can test for the exclusion of Z_t in the first and third equations, jointly with the exclusion of Z_{1t} from the second equation. The instruments $\widetilde{Z}_t$ are therefore used to correct for measurement errors.

The above inference method is system-based, yet is not strictly FI. In the same vein, the tests that we invert have a likelihood-based justification, yet they are not strictly FIML-based. This is formally shown in the Appendix where we also demonstrate robustness (in the sense of significance level control) to misspecification of the DGP underlying model (2). In particular, we show that the $\mathcal{W}$ statistic is compatible with a general class of reduced forms. For instance, in the context of the model that we are considering, all we need is: (i) to assume that inflation, output, and the interest rate variable can jointly be explained, up to possibly contemporaneously-correlated disturbances, by their own lags (via some linear or nonlinear VAR form), (ii) a number of predetermined variables, which may or may not come from the theoretical model (intra-model or external instruments), and (iii) possibly a set of further exogenous or predetermined variables which were not included in the test [i.e. exogenous or predetermined variables that intervene in the fundamental data generating process yet were "missed" in the sense or "not considered" by the econometrician. Most importantly, our exposition in the Appendix implies that the latter missing instruments have no incidence on the test's validity.

Since the model we consider does not imply cross-equation constraints, it is possible (and valid) to apply the univariate approach of Dufour, Khalaf, and Kichian (2006) on an equation-by-equation basis. Invariance to contemporaneous correlation of disturbances derives from the results of Dufour and Khalaf (2002), and, as may be checked from the Appendix, the underlying VAR instrumental model is also valid if one focuses on the implications of each structural equation one at a time, using a univariate statistic [as long as we work within the regression (4)]. Of course, the derived confidence sets across equations will not be simultaneous (global size control is not warranted), yet each remains asymptotically level-correct. For the same reason, one may also relax all constraints and estimate each regression equation from the system (2) as a linear simultaneous equation, using the methodology from Dufour and Taamouti (2005). In this fashion, it is possible to analyze how results are affected as more restrictions are relaxed while still maintaining endogeneity and possible errors-in-variables.

To conclude, we note that our distributional assumptions are motivated, on the one hand by convenience and on the other, by the available literature (Bayesian or classical; see the discussion in the rejoinder to Del Negro, Schorfheide, Smets, and Wouters (2007)). While non-normality is not a major issue here [since asymptotic cut-offs can be used in the context of (4)] extending our approach to account for time dependence or heteroskedasticity is conceptually straightforward. For example, one may rely on robust Wald statistics associated with regression (4). For that matter, any test statistic that fits the hypothesized error distributional assumptions will be identification-robust if conveniently applied to regression (4).[12]

4. Empirical Results

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. 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), 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 as in Linde (2005), and which is included for comparison purposes. We then take the log of both these output gap series.

Our 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.[13] Finally, as in Linde (2005), all our data is demeaned prior to estimation.

We first examine whether values in the vicinity of those reported by Linde (2005), and which were obtained with FIML, are supported by our methods. A grid search is conducted for coefficients covering about three standard errors around the obtained estimated values,

[12] Note that our methodology can conceptually also be adapted to allow for parameter time-variation and non-stationary variables. However such extensions are beyond the scope of the present paper.

[13] Wage and commodity price inflation were also in the instrument sets of Gali, Gertler, and Lopez-Salido (2001), and Galí, Gertler, and Lopez-Salido (2005).

with and without external instruments, and using the standard output gap measure.[14] We find that the joint confidence set is entirely empty, both with only intra-model instruments, and with intra- and extra-model instruments. This indicates that this econometric model is soundly rejected at the five per cent level.

We next conduct an unrestricted search over all of the admissible parameter space and calculate the J-type tests that are based on $\min_{\theta_0} \mathcal{W}$ (as described above). These are accomplished with each measure of the output gap (i.e., the standard or the real-time), and for both cases of when only intra-model or intra- and extra-model instruments are considered. Results with the standard output gap measure conform with our restricted grid search: the $\min_{\theta_0} \mathcal{W}$-based J-test is significant with either instrument sets. In contrast, when the real-time gap measure is used, the model is no longer rejected at the 5 per cent level, again for either instrument set. The test p-values are thus 0.0850 when intra-model instruments are used, and 0.0787 when both intra and extra-model instruments are considered.

Focusing on the model with the real-time gap, and where both internal and external instruments are used, we now proceed with the derivations of projections for the parameters of interest using numerical methods. After inverting our proposed multivariate test, the results are as follows:

- The NKPC equation:

  [0.6064, 0.9019] for ω_f, [-0.0368, 0.0126] for γ.

- The IS equation:

  [0.3415, 0.6456] for β_f, [-0.0668, 0.0293] for β_r.

- The Taylor rule:

  [2.7445, 3.1583] for γ_π, [3.1539, 3.3912] for γ_y,
  [0.9886, 1.1357] for ρ_1, [-0.5664,-0.3537] for ρ_2, [0.1649, 0.3593] for ρ_3.

[14] More specifically, the parameter ranges that we use are: [0.9, 1.2], [-0.8, -0.2], [0.2, 0.4], for ρ_1^0, ρ_2^0, and ρ_3^0, respectively (making sure that their sum remains inferior to one), [0.2, 0.4] for ω_f^0, [0.04, 0.06] for γ^0, [0.40, 0.46] for β_f^0, [0.08, 0.10] for β_r^0, [0.8, 1.1] for γ_π^0, and [0.5, 1.5] for γ_y^0. The corresponding incremental values for the numerical search are 0.1 for ρ_1^0, ω_f^0 and γ_y^0, 0.05 for ρ_2^0 and ρ_3^0, and β_r^0, 0.01 for β_f^0 and γ_π^0, and 0.005 for γ^0.

Within our confidence set, we can find the least-rejected parameter combination. Our Hodges-Lehmann point estimates are: $\widehat{\omega_f} = 0.7309$, $\widehat{\beta_f} = 0.4891$, $\widehat{\gamma_\pi} = 2.9252$, $\widehat{\gamma_y} = 3.2756$, while $\widehat{\gamma}$ and $\widehat{\beta_r}$ are both zero. Finally, the point estimate for sum of the interest rate autoregressive terms equals 0.2914.

The results reveal some striking features. First, there are identification difficulties associated with our considered model. This is apparent from the fact that some parameters have fairly wide ranges for their projections (particularly, parameters ω_f and β_f). Second, the projections are not symmetric. In other words, some parameter values within the projection range are more compatible with the data than other points. This is not apparent from the projections *per se* (though this information is readily available to the researcher) but can be observed when we examine the point estimates with respect to the projections. For example, the most probable point for γ_y is closer to the upper limit of its projection space than to the midpoint of that range. Third, despite the identification difficulties, the econometric model is quite informative about certain features of the economy. That is, in general terms, none of the projections are unbounded, or reach the limits of their admissible parameter space. Most of the model coefficients (except for two) are significant and coefficients have the expected signs.

We would like to focus now particularly on the conclusions with regard to the NKPC equation, as the debate on whether the curve is mostly forward- or backward-looking is far from settled. Proponents of the FIML approach such as Linde (2005) argue (for example, using simulation exercises), that full-information approaches are more likely to pin down true model parameter values. However, their conclusions are drawn in studies that are absence of any consideration for identification concerns. Similarly, proponents of limited-information approaches argue that when one estimates the closed form of a model, carefully mapping the structural form with the closed form, then limited-information methods produce valid outcomes (see, Sbordonne 2005 and Galí, Gertler, and Lopez-Salido 2005). However, the conclusions of the latter are also obtained without any reference to identification difficulties. Indeed, much of the debate seems rather to have focused on misspecification issues.

Our newly-proposed method (as well as the comparisons to existing univariate methods; see below) allows us to shed some light on this debate. In this respect, the results of our inference point decisively to a preponderantly forward-looking inflation equation. Though the range is fairly wide for the ω_f parameter, the weight on the forward-looking term is at least 60 per cent, and at most 90 per cent. This also implies that the NKPC is not purely forward-looking, as the backward-looking term is between 10 and 40 per cent. Thus, our results concur with the limited-information-based results of Roberts (2001) and of Galí, Gertler, and Lopez-Salido (2005), and are at variance with the FIML-based outcome of Linde (2005) that concludes that the inflation process is pre-ponderantly backward-looking. At the same time, however, we find that the coefficient on the output gap term in the NKPC is not significant (our projection range includes zero, and indeed our point estimate *is* zero).

For comparison purposes, we also apply two existing (yet fairly recent) univariate identification-robust methods to the NKPC. One is the method applied by Dufour, Khalaf, and Kichian (2006), and the other is the univariate linear IV method proposed in Dufour and Taamouti (2005). The data, instruments, and variables used remain the same as above. The only difference with the more general set-up is that certain model restrictions are not imposed. Accordingly, each univariate method integrates and assesses the model's structural restrictions to a different degree.

First we apply the method of Dufour, Khalaf, and Kichian (2006) to the NKPC. In this case, the restriction that the sum of the backward and forward-looking components of inflation sum to one is retained, and the instruments for the estimation now include all of the included predetermined variables in the multi-equation structure, as well as the considered external instruments. The results yield the ranges [0.345, 0.995] for the ω_f, and [-0.075, 0.055] for γ parameters, respectively.[15] Both projection regions are wider than those obtained using our multivariate approach. Notably, with this method, it is no longer possible to ascertain that the NKPC is mostly forward-looking. Interestingly, the coefficient on the gap term is not significant, as was the case with the multiequation approach results.

Next, we drop the restriction that the sum of the inflation lead and lag sum to one, and apply the method proposed in Dufour and Taamouti (2005). Here, the search space for the ω_f and γ parameters are between minus infinity and plus infinity. In this case, we find that the projection region for ω_f is [0.8649, 1.1908], while for γ, it is [-0.0905, 0.0539]. Once again, it can be concluded that the curve is forward-looking preponderantly, but, according to this method, values of ω_f above one are also admissible. As for the range of the coefficient on the output gap, it is slightly wider than with the previous two methods, and similar to the previous results, it includes zero.

In summary, we see that the multi-equation confidence intervals are generally tighter than their univariate counterparts, and thus more informative. These results are compatible with efficiency gains associated with a systems-based approach. Most notably, our results support a preponderantly forward-looking NKPC whereas the confidence set for the coefficient on the

[15] For numerical tractability, the upper end of the search region for ω_f was 0.995.

output gap term is found to cover zero whether we use multivariate or univariate methods. In addition, our results point to the existence of both forward- and backward-looking components in the IS curve, and to an insignificant coefficient on the real interest rate, which indicate that expectation-based terms seem to be the driving variables in the three-equation system. Finally, all parameters of the interest rate rule are significant, and there is evidence for smoothing behaviour in interest rates.

4. Conclusion

Taken in comparison with our earlier work (Dufour, Khalaf, and Kichian 2006, our findings indicate that LI methods, though suffering from weak-identification problems, nonetheless provide some information on the U.S. inflation process.

These results are of course specific to the models analyzed, yet they call for caution in interpreting available FIML results based on standard econometric techniques. The fact remains that, though our results are model-specific, our new methodology is, in principle, applicable beyond the specific model that was analyzed here and numerical burdens are not more demanding then the current state-of-the-art in the literature.

DSGE modelers are often confronted, among others, with the following enduring questions: (1) Should we construct large scale econometric models to capture full structural macro-economic models or should we instead focus on smaller models which address a few relevant features of interest? (2) Models are approximations i.e. "wrong" by construction, so to what extent should specification and observational equivalence issues ultimately matter given that macro-economic data are scarce? The literature is constantly struggling with such questions and many of the available econometric answers have recently taken a Bayesian perspective. Whether frequentists or Bayesians, if economists are to address such questions via productive use of econometric methods, they must endeavor to apply and develop procedures for which error probabilities can be controlled precisely. Our paper throws some light [we clearly do not claim we resolve such broad and fundamental questions] on the matter from a frequentist perspective: we propose a methodology which allows to focus on a sub-model of choice, yet it is provably robust to many characteristics of the underlying full model including full identification, missing instruments or error-in-variable problems. Pursuing identification-robust multivariate approaches is therefore a worthy research objective. These will be important to the academic community and policy makers since they might very well show that models for which doubt had been cast in the past or results that have led to unclear policy recommendations could in fact be better understood given an adequate methodology.

References

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.

Christiano, L., M. Eichenbaum, and M. Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113: 1–45.

Clarida, R., J. Gali, and M. Gertler. 1999. "The Science of Monetary Policy: A New Keynesian Perspective." *Journal of Economic Literature* 37: 1661–1707.

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: 123–162.

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(4): 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. 2002. "Simulation Based Finite and Large Sample Tests in Multivariate Regressions." *Journal of Econometrics* 111(2): 303–322.

———. 2003. "Finite Sample Tests in Seemingly Unrelated Regressions." In *Computer-Aided Econometrics*, edited by D.E.A. Giles, chapter 2, 11–35. New York: Marcel Dekker.

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.

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.

Fernandez-Villaverde, J., J. Rubio-Ramirez, and S. Sargent. 2005. *A,B,C's (And D's) For Understanding VARS.* Technical report, University of Pennsylvania.

Gali, J., M. Gertler, and J.D. Lopez-Salido. 2001. "European Inflation Dynamics." *European Economic Review* 45: 1237–1270.

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.

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

———. 1983. "Hodges-Lehmann Estimators." In *Encyclopedia of Statistical Sciences, Volume 3*, edited by N.L. Johnson, S. Kotz, and C. Read, 642–645. New York: John Wiley & Sons.

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.

Linde, J. 2005. "Estimating New Keynesian Phillips Curves: A Full Information Maximum Likelihood Approach." *Journal of Monetary Economics* 52: 1135–49.

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 ot the Phillips curve." *Journal of Money, Credit and Banking* 37: 421–449.

Moreira, M.J. 2003. "A Conditional Likelihood Ratio Test for Structural Models." *Econometrica* 71(4): 1027–1048.

Nason, J.M. and G.W. Smith. 2003. *Identifying the New Keynesian Phillips Curve.* Technical report, Univeristy of British Columbia and Queen's University.

Roberts, J. 2001. *How Well Does the New Keynesian Sticky-Price Model Fit the Data?* Technical report, Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series 2001-13, Washington.

Sbordonne, A. 2005. “Do Expected Future Marginal Costs Drive Inflation Dynamics.” *Journal of Monetary Economics* 52: 1183–97.

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

Staiger, D. and J.H. Stock. 1997. “Instrumental Variables Regression with Weak Instruments.” *Econometrica* 65(3): 557–586.

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(4): 518–529.

Theil, H. 1971. *Principles of Econometrics.* New York: John Wiley & Sons.

Wang, J. and E. Zivot. 1998. “Inference on Structural Parameters in Instrumental Variables Regression with Weak Instruments.” *Econometrica* 66(6): 1389–1404.

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

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.

Appendix

Consider the multivariate regression

$$Y = XB + U \tag{8}$$

where $Y = [Y_1, \ \ldots \ , \ Y_T]'$ is the $T \times n$ matrix of observations on n dependant variables, $X = [X_1, \ \ldots \ , \ X_T]'$ is the $T \times k$ matrix of regressors and $U = [U_1, \ \ldots \ , \ U_T]'$ is a $T \times n$ matrix of disturbances. For instance, (4) can be written in this form, with

$$Y_t = (\pi_t^*, y_t^*, R_t^*)', \quad X_t = \left(Z_t', \widetilde{Z}_t'\right)', \quad U_t = (\epsilon_{\pi,t}, \epsilon_{y,t}, \epsilon_{R,t})', \quad t = 1, ..., T, \tag{9}$$

so $n = 3$, $k = 8 + q$. Excluding Z_{1t} and $\widetilde{Z}_t$ from all equations and Z_{2t} from the first and third equation of (4) may be tested using the usual SURE-type F tests. In our context, the statistic takes the following form. Let $\widehat{B}$ denote the OLS estimator of the coefficients of (4), and let $\widehat{b} = vec(\widehat{B})$. Furthermore, we define:

$$\begin{aligned} \pi^* &= (\pi_1^*, ..., \pi_T^*)', \quad y^* = (y_1^*, ..., y_T^*)', \quad R^* = (R_1^*, ..., R_T^*)' \\ \mathcal{Y}^* &= \begin{bmatrix} \pi^* \\ y^* \\ R^* \end{bmatrix}, \mathcal{X}^* = \begin{bmatrix} X & 0 & 0 \\ 0 & X & 0 \\ 0 & 0 & X \end{bmatrix}. \end{aligned}$$

Then we consider the SURE Wald-type statistic[16]

$$\mathcal{W} = \left(\frac{3(T-k)}{m}\right) \frac{(A\widehat{b})'\left[A\left(\mathcal{X}^{*\prime}\left(\hat{\Sigma}^{-1} \otimes I_n\right)\mathcal{X}^*\right)^{-1}A'\right]^{-1}(A\widehat{b})}{(\mathcal{Y}^* - \mathcal{X}^*\widehat{b})'\left(\hat{\Sigma}^{-1} \otimes I_n\right)(\mathcal{Y}^* - \mathcal{X}^*\widehat{b})} \tag{10}$$

where A is the $m \times 3k$ selection matrix with $m = 2k + 4 + q$

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

(Theil, 1971, Chapter 6) suggests that the $F(m, 3(T-k))$ provides a good approximation to the null distribution of $\mathcal{W}$. Dufour and Khalaf (2003) confirm this claim in the context of a three-equations SURE system.

The test conducted in this framework supposes that the (unrestricted) reduced form for the system is given, up to an error term, by some function of: (i) the predetermined variables

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

in the system, (ii) the extra instruments used in the test, and (iii) possibly a set of further explanatory variables which were not used in the test. So by conducting the test of (3) in the context of (2) as a test of (**??**) in the context of (4), as described, we obtain a p-value that it is in fact robust to the specification of the fundamental DGPs under consideration, to measurement errors and excluded instruments. To see this, suppose that the reduced form takes the unrestricted VAR specification

$$\begin{array}{r}\pi_t = a_\pi \pi_{t-1} + \sum_{i=1}^{3} b_{\pi,i} R_{t-i} + \sum_{i=1}^{4} c_{\pi,i} y_{t-i} + \varpi_\pi' Q_t + \nu_{\pi,t} \\ y_t = a_y \pi_{t-1} + \sum_{i=1}^{3} b_{y,i} R_{t-i} + \sum_{i=1}^{4} c_{y,i} y_{t-i} + \varpi_y' Q_t + \nu_{y,t} \\ R_t = a_R \pi_{t-1} + \sum_{i=1}^{3} b_{R,i} R_{t-i} + \sum_{i=1}^{4} c_{R,i} y_{t-i} + \varpi_R' Q_t + \nu_{R,t}\end{array}. \tag{11}$$

where $Q_t = \left(\widetilde{Z}_t', \widetilde{Q}_t'\right)'$, and $\widetilde{Q}_t$ are a set of relevant explanatory variables; these may include further lags of the endogenous variables, and/or further predetermined or exogenous variables. Let

$$\begin{aligned}
\omega_f^* &= \omega_f - \omega_f^0, \quad \gamma^* = \gamma - \gamma^0, \quad \beta_f^* = \beta_f - \beta_f^0, \quad \beta_r^* = \beta_r - \beta_r^0, \quad \gamma_\pi^* = \gamma_\pi - \gamma_\pi^0; \\
\gamma_y^* &= \gamma_y - \gamma_y^0, \quad \rho_1^* = \rho_1 - \rho_1^0, \quad \rho_2^* = \rho_2 - \rho_2^0, \quad \rho_3^* = \rho_3 - \rho_1^0; \\
\beta_{y,i}^* &= \beta_{y,i}(1-\beta_f), \quad i = 1, ..., 4; \quad \rho_i^* = \rho_i - \rho_i^0, \quad i = 1, ..., 3; \\
\gamma_\pi^* &= \left(1 - \sum_{i=1}^{3} \rho_i\right) \gamma_\pi - \left(1 - \sum_{i=1}^{3} \rho_i^0\right) \gamma_\pi^0; \\
\gamma_y^* &= \left(1 - \sum_{i=1}^{3} \rho_i\right) \gamma_y - \left(1 - \sum_{i=1}^{3} \rho_i^0\right) \gamma_y^0.
\end{aligned}$$

Substituting (11) into (4) leads, for the inflation equation, to the following: where $Q_t = \left(\widetilde{Z}_t', \widetilde{Q}_t'\right)'$, and $\widetilde{Q}_t$ are a set of relevant explanatory variables; these may include further lags of the endogenous variables, and/or further predetermined or exogenous variables. Let

$$\begin{aligned}
\omega_f^* &= \omega_f - \omega_f^0, \quad \gamma^* = \gamma - \gamma^0, \quad \beta_f^* = \beta_f - \beta_f^0, \quad \beta_r^* = \beta_r - \beta_r^0, \quad \gamma_\pi^* = \gamma_\pi - \gamma_\pi^0; \\
\gamma_y^* &= \gamma_y - \gamma_y^0, \quad \rho_1^* = \rho_1 - \rho_1^0, \quad \rho_2^* = \rho_2 - \rho_2^0, \quad \rho_3^* = \rho_3 - \rho_1^0; \\
\beta_{y,i}^* &= \beta_{y,i}(1-\beta_f), \quad i = 1, ..., 4; \quad \rho_i^* = \rho_i - \rho_i^0, \quad i = 1, ..., 3; \\
\gamma_\pi^* &= \left(1 - \sum_{i=1}^{3} \rho_i\right) \gamma_\pi - \left(1 - \sum_{i=1}^{3} \rho_i^0\right) \gamma_\pi^0; \\
\gamma_y^* &= \left(1 - \sum_{i=1}^{3} \rho_i\right) \gamma_y - \left(1 - \sum_{i=1}^{3} \rho_i^0\right) \gamma_y^0.
\end{aligned}$$

Substituting (11) into (4) leads, for the inflation equation, to the following:

$$
\begin{aligned}
\pi_t^* = \pi_{t-1} & \left[\omega_f^* \left(a_\pi^2 + b_{\pi,1} a_R + c_{\pi,1} a_y - 1\right) + \gamma^* \left(a_y\right)\right] \\
& + R_{t-1} \left[\omega_f^* \left(a_\pi b_{\pi,1} + b_{\pi,1} b_{R,1} + c_{\pi,1} b_{y,1} + \gamma^* \left(b_{y,1}\right)\right)\right] \\
& + R_{t-2} \left[\omega_f^* \left(a_\pi b_{\pi,2} + b_{\pi,1} b_{R,2} + c_{\pi,1} b_{y,2} + b_{\pi,2}\right) + \gamma^* \left(b_{y,2}\right)\right] \\
& + R_{t-3} \left[\omega_f^* \left(a_\pi b_{\pi,3} + b_{\pi,1} b_{R,3} + c_{\pi,1} b_{y,3} + b_{\pi,3}\right) + \gamma^* \left(b_{y,3}\right)\right] \\
& + y_{t-1} \left[\omega_f^* \left(a_\pi c_{\pi,1} + b_{\pi,1} c_{R,1} + c_{\pi,1} c_{y,1}\right) + \gamma^* \left(c_{y,1}\right)\right] \\
& + y_{t-2} \left[\omega_f^* \left(a_\pi c_{\pi,2} + b_{\pi,1} c_{R,2} + c_{\pi,1} c_{y,2} + c_{\pi,2}\right) + \gamma^* \left(c_{y,2}\right)\right] \\
& + y_{t-3} \left[\omega_f^* \left(a_\pi c_{\pi,3} + b_{\pi,1} c_{R,3} + c_{\pi,1} c_{y,3} + c_{\pi,3}\right) + \gamma^* \left(c_{y,3}\right)\right] \\
& + y_{t-4} \left[\omega_f^* \left(a_\pi c_{\pi,4} + b_{\pi,1} c_{R,4} + c_{\pi,1}^* c_{y,4} + c_{\pi,4}\right) + \gamma^* \left(c_{y,4}\right)\right] \\
& + \left(\omega_f^* a_\pi \varpi_\pi' + \omega_f^* b_{\pi,1} \varpi_R' + \left(\omega_f^* c_{\pi,1} + \gamma^*\right) \varpi_y'\right) \widetilde{Z}_t + \xi_{\pi,t}
\end{aligned}
$$

where

$$
\begin{aligned}
\xi_{\pi,t} \quad = \quad & \omega_f^* a_\pi \nu_{\pi,t} + \omega_f^* b_{\pi,1} \nu_{R,t} + \left(\omega_f^* c_{\pi,1} + \gamma^*\right) \nu_{\pi,t} \\
& + \left(\omega_f^* a_\pi \varpi_\pi' + \omega_f^* b_{\pi,1} \varpi_R' + \left(\omega_f^* c_{\pi,1} + \gamma^*\right) \varpi_y'\right) \widetilde{Q}_t \\
& + \omega_f^* \nu_{\pi,t+1} + \omega_f^* \varpi_\pi' Q_{t+1} + \epsilon_{\pi,t}.
\end{aligned}
$$

The term $\xi_{\pi,t}$ includes, in addition to the errors terms, the explanatory variables that were missing from the multivariate regression (4). Turning to the output equation, we have

$$
\begin{aligned}
y_t^* = \pi_{t-1} & \left[\beta_f^* \left(a_\pi a_y + a_y c_{y,1} + a_R b_{y,1}\right) + \beta_r^* \left(a_\pi^2 + a_y c_{\pi,1} + a_R b_{\pi,1} - 1\right)\right] \\
& + R_{t-1} \left[\beta_f^* \left(b_{\pi,1} a_y + b_{y,1} c_{y,1} + b_{R,1} b_{y,1} + b_{y,2}\right) + \beta_r^* \left(b_{\pi,1} a_\pi + b_{y,1} c_{\pi,1} + b_{R,1} \left(b_{\pi,1} - 1\right) + b_{\pi,2}\right)\right] \\
& + R_{t-2} \left[\beta_f^* \left(b_{\pi,2} a_y + b_{y,2} c_{y,1} + b_{R,2} b_{y,1} + b_{y,3}\right) + \beta_r^* \left(b_{\pi,2} a_\pi + b_{y,2} c_{\pi,1} + b_{R,2} \left(b_{\pi,1} - 1\right) + b_{\pi,3}\right)\right] \\
& + R_{t-3} \left[\beta_f^* \left(b_{\pi,3} a_y + b_{y,3} c_{y,1} + b_{R,3} b_{y,1}\right) + \beta_r^* \left(b_{\pi,3} a_\pi + b_{y,3} c_{\pi,1} + b_{R,3} \left(b_{\pi,1} - 1\right)\right)\right] \\
& + y_{t-1} \left[\beta_f^* \left(c_{\pi,1} a_y + c_{y,1}^2 + c_{R,1} b_{y,1} + c_{y,2}\right) + \beta_r^* \left(c_{\pi,1} a_\pi + c_{y,1} c_{\pi,1} + c_{R,1} \left(b_{\pi,1} - 1\right) + c_{\pi,2}\right) + \beta_{y,1}^*\right] \\
& + y_{t-2} \left[\beta_f^* \left(c_{\pi,2} a_y + c_{y,2} c_{y,1} + c_{R,2} b_{y,1} + c_{y,3}\right) + \beta_r^* \left(c_{\pi,2} a_\pi + c_{y,2} c_{\pi,1} + c_{R,2} \left(b_{\pi,1} - 1\right) + c_{\pi,3} + \beta_{y,2}^*\right)\right] \\
& + y_{t-3} \left[\beta_f^* \left(c_{\pi,3} a_y + c_{y,3} c_{y,1} + c_{R,3} b_{y,1} + c_{y,4}\right) + \beta_r^* \left(c_{\pi,3} a_\pi + c_{y,3} c_{\pi,1} + c_{R,3} \left(b_{\pi,1} - 1\right) + c_{\pi,4} + \beta_{y,3}^*\right)\right] \\
& + y_{t-4} \left[\beta_f^* \left(c_{\pi,4} a_y + c_{y,4} c_{y,1} + c_{R,4} b_{y,1}\right) + \beta_r^* \left(c_{\pi,4} a_\pi + c_{y,4} c_{\pi,1} + c_{R,4} \left(b_{\pi,1} - 1\right)\right) + \beta_{y,4}^*\right] \\
& + \left[\beta_f^* \left(a_y \varpi_\pi' + c_{y,1} \varpi_y'\right) + \beta_r^* \left(a_\pi \varpi_\pi' + c_{\pi,1} \varpi_y'\right)\right] \widetilde{Z}_t + \xi_{y,t}
\end{aligned}
$$

where

$$\xi_{y,t} = \left(\beta_f^* \varpi_y' + \beta_r^* \varpi_\pi'\right) Q_{t+1} + \left[\beta_f^* \left(a_y \varpi_\pi' + c_{y,1} \varpi_y'\right) + \beta_r^* \left(a_\pi \varpi_\pi' + c_{\pi,1} \varpi_y'\right)\right] \widetilde{Q}_t$$
$$+\beta_r^* \nu_{\pi,t+1} + \left(\beta_f^* c_{y,1} + \beta_r^* c_{\pi,1}\right) \nu_{\pi,t}$$
$$+\beta_f^* \nu_{y,t+1} + \left(\beta_f^* a_y + \beta_r^* a_\pi\right) \nu_{y,t}$$
$$+ \left(\beta_f^* b_{y,1} + \beta_r^* \left(b_{\pi,1} -\right) 1\right) \nu_{R,t} + \epsilon_{y,t}.$$

Finally, the interest rate equation corresponds to:

$$\begin{aligned} R_t^* = \; & \pi_{t-1} \left[\gamma_\pi^* a_\pi + \gamma_y^* a_y\right] \\ & + R_{t-1} \left[\gamma_\pi^* b_{\pi,1} + \gamma_y^* b_{y,1} + \rho_1^*\right] \\ & + R_{t-2} \left[\gamma_\pi^* b_{\pi,2} + \gamma_y^* b_{y,2} + \rho_2^*\right] \\ & + R_{t-3} \left[\gamma_\pi^* b_{\pi,3} + \gamma_y^* b_{y,3} + \rho_3^*\right] \\ & + y_{t-1} \left[\gamma_\pi^* c_{\pi,1} + \gamma_y^* c_{y,1}\right] \\ & + y_{t-2} \left[\gamma_\pi^* c_{\pi,2} + \gamma_y^* c_{y,2}\right] \\ & + y_{t-3} \left[\gamma_\pi^* c_{\pi,3} + \gamma_y^* c_{y,3}\right] \\ & + y_{t-4} \left[\gamma_\pi^* c_{\pi,4} + \gamma_y^* c_{y,4}\right] \\ & + \left[\gamma_\pi^* \varpi_\pi' + \gamma_y^* \varpi_y'\right] \widetilde{Z}_t + \xi_{R,t} \end{aligned}$$

where

$$\xi_{R,t} = \left[\gamma_\pi^* \varpi_\pi' + \gamma_y^* \varpi_y'\right] \widetilde{Q}_t + \gamma_\pi^* \nu_{\pi,t} + \gamma_y^* \nu_{y,t} + \epsilon_{R,t}.$$

We thus see that that under the null hypothesis

$$\xi_{\pi,t} = \epsilon_{\pi,t}, \quad \xi_{y,t} = \epsilon_{y,t}, \quad \xi_{R,t} = \epsilon_{R,t}, \quad i = 1, ..., 4,$$

and the null model collapses to

$$\pi_t^* = \epsilon_{\pi,t}, \tag{12}$$

$$y_t^* = \sum_{i=1}^{4} \beta_{y,i}^* y_{t-i} + \epsilon_{y,t}, \tag{13}$$

$$R_t^* = \epsilon_{R,t}, \tag{14}$$

which justifies the tests we apply. The above derivations validate our test procedure given usual (unconstrained) reduced form assumption on the macro-economic aggregate under consideration.