



ISSN 2282-6483

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

**Under the same (Chole)sky:
DNK models, timing restrictions
and recursive identification of
monetary policy shocks**

Giovanni Angelini
Marco M. Sorge

Quaderni - Working Paper DSE N°1160



Under the same (Chole)sky: DNK models, timing restrictions and recursive identification of monetary policy shocks*

Giovanni Angelini[†]

Marco M. Sorge[‡]

April 11, 2021

Abstract

Recent structural VAR studies of the monetary transmission mechanism have voiced concerns about the use of recursive identification schemes based on short-run exclusion restrictions. We trace out the effects on impulse propagation of informational constraints embodying classical Cholesky-timing restrictions in otherwise standard Dynamic New Keynesian (DNK) models. By reinforcing internal propagation mechanisms and enlarging a model's equilibrium state space, timing restrictions may produce a non-trivial moving average component of the equilibrium representation, making finite order VARs a poor approximation of true adjustment paths to monetary impulses, albeit correctly identified. They can even serve as an independent source of model-based nonfundamentalness, thereby hampering shock identification via VAR methods. This notwithstanding, restricted DNK models are shown to feature (i) invertible equilibrium representations for the observables and (ii) fast-converging VAR coefficient matrices under empirically tenable parameterizations. This alleviates concerns about identification and lag truncation bias: low-order Cholesky-VARs do well at uncovering the transmission of monetary impulses in a truly Cholesky world.

Keywords: DSGE models, Timing restrictions, Vector autoregression, Cholesky identification

JEL Classification: C3, E3

*We wish to thank Efreem Castelnovo, Chetan Dave, Luca Fanelli, Alessia Paccagnini and seminar participants (CSEF) for helpful comments and discussions. Any remaining errors are our own.

[†]Department of Economics, University of Bologna, g.angelini@unibo.it

[‡]Corresponding author: University of Salerno, University of Göttingen and CSEF - DISES, Via Giovanni Paolo II 132 84084 Fisciano (SA), msorge@unisa.it

Non - Technical Summary

Understanding how changes in the monetary policy instruments (e.g. the monetary base and/or the short term nominal interest rate) affect real variables such as aggregate output and employment, is key to evaluating the effectiveness of monetary policy against its stabilization goals. In standard New Keynesian models of the monetary transmission mechanism, a monetary tightening, in the form of an unexpected shock to the interest rate, produces real effects in the presence of staggered price setting, for it depresses aggregate output, which in turn puts downward pressure on inflation. It is often argued that such monetary policy shocks affect real variables only with a delay. This view has informed a large number of econometric explorations of the transmission mechanism that rely on reduced form autoregressive models. It is nonetheless less clear what this view would bring about in theory-based representations of the dynamic interactions between private agents and the monetary authority. The present paper interrogates theoretical New Keynesian models in order to trace out the effects of transmission delays that originate from the lack of up-to-date information on the part of private agents about the occurrence of unforeseen monetary tightenings. Our main results are as follows: (i) information-based transmission delays can be a serious threat to the proper estimation of the true monetary transmission mechanism, for it is likely to require sample (time series) data of length that are unavailable in practice; these delays can even prevent identification of monetary shocks, for they are not retrievable from observed data; (ii) information-constrained New Keynesian models do not suffer from these issues: standard reduced form autoregressive systems that comply with the "delayed effects" view of policy shocks perform well in detecting the true monetary transmission mechanism.

1 Introduction

Recursive identification schemes based on short-run exclusion restrictions have traditionally been used to identify the macroeconomic effects of monetary policy shocks (e.g. Sims (1980), Christiano et al. (1999)). Making the recursive scheme operative in the VAR context is accomplished by generating a Cholesky decomposition of the variance-covariance matrix associated with the reduced-form residuals, with the monetary policy shock ordered last (Kilian (2013)). Given its ability to exactly identify the dynamic responses of all variables included in the VAR system to the monetary policy shock, the Cholesky assumption is empirically appealing, and yet calls for a conceivable structural interpretation of the recursive ordering.¹

The present paper develops a formal analysis of informational constraints embodying Cholesky-timing restrictions in otherwise standard Dynamic Stochastic General Equilibrium (DSGE) models; and explores the econometric relationship between such models and recursively identified VARs in controlled Monte Carlo experiments. Timing restrictions specify an information structure under which control variables are set on the basis of an increasing sequence of nested information sets, that constrain the endogenous adjustment paths to exogenous impulses (e.g. Kormilitsina (2013)). By reinforcing internal propagation mechanisms and enlarging a model’s equilibrium state space, these informational constraints may dramatically change rational expectations cross-equation restrictions (CERs) to the point of (i) generating nonfundamental representations for the observables, or (ii) inducing (invertible) non-trivial moving average (MA) components vis-à-vis their counterparts free of timing restrictions. Even when fundamental, equilibrium representations of DSGE models with timing restrictions may thereby exhibit slowly converging VAR polynomial matrices at longer horizons. In these cases, finite order VARs would stand as a poor approximation of the true VARMA structure because of truncation effects, that need not originate from small sample bias of the estimator, see e.g. Kapetanios et al. (2007), Ravenna (2007), Poskitt and Yao (2017)).

We identify conditions for existence of fundamental equilibrium representations of DSGE models with timing restrictions, that rely on the underlying information structure and the reduced-form coefficients of the model’s solution under conventional (unrestricted) timing.² We then employ a variety of model laboratories to evaluate the performance of

¹The performance of Cholesky-VARs has recently been shown to be competitive with that of alternatives such as traditional sign restrictions in a Monte Carlo context (Wolf (2020)).

²We will henceforth use the labels “unrestricted model” and “model without timing restrictions”, as well as the labels “restricted model” and “model with timing restriction”, interchangeably.

Cholesky-VARs in impulse response analysis. Our core findings show that DNK models with timing restrictions do not suffer from nonfundamentalness issues under empirically tenable parameterizations, and typically display fast decaying coefficients of the ensuing VAR(∞) representation. Even in cases where the timing of decisions deeply affect the sign, the magnitude and the persistence of dynamic adjustment paths to a monetary policy surprise, low-order Cholesky-VARs are found to perform remarkably well in retrieving structural monetary impulse responses in a truly Cholesky world.

It is well known that a broad class of DSGE models allow contemporaneous responses of the full set of model variables to structural disturbances. This lack of conformity with Cholesky-style identifying assumptions raises two crucial questions for empirical work. First, were a given unrestricted DSGE model to serve as the actual data generating process, what kind of distortion would estimated dynamic responses to macroeconomic shocks exhibit if derived from a Cholesky-identified VAR? Second, and relatedly, when taking a DSGE model to the data, would statistical techniques based on impulse response matching (e.g. the minimum distance estimation approach popularized by Rotemberg and Woodford (1997)) represent a valid guidance in the development of economic theory by offering reliable estimates of structural (deep) parameters, i.e. those invariant to policy changes?

Against these questions, Canova and Pina (2005) and Carlstrom et al. (2009) both exploit small-scale DSGE models as data generating processes in order to inspect the identifying properties of short-run exclusion restrictions in the VAR realm. Relying on calibrated versions of a limited participation framework and a sticky price-sticky wage economy, Canova and Pina (2005) report evidence of severe mis-identification in estimated dynamic responses to monetary policy disturbances as well as in variance decomposition outcomes. Carlstrom et al. (2009) theoretically explore the mapping from the true parameterization of Dynamic New Keynesian (DNK) models to the impulse predictions of Cholesky-type VARs when it comes to estimating macroeconomic reactions to a monetary policy shock. They show that, conditional on the selected parameterization, short-run exclusion restrictions may fail to identify unexpected, temporary increases in the policy rate and even point to a muted response of output and price inflation to the latter at all time horizons, even when the model-implied dynamic adjustment path starts away from zero and displays a large degree of persistence. What is more, price and output puzzles are shown to be fairly likely to arise under small perturbations of the model's parameters, that leave the theoretical impulse responses almost unchanged.³

³The apparent inconsistency between the standard timing in DSGE economies and the recursive order-

Castelnuovo (2016) contributes to the debate by proposing a formal comparison of impulse response functions (IRFs) generated by a recursively identified VAR estimated on real-world data vis-à-vis the IRFs associated with recursively identified VAR estimated on artificial series produced by an estimated medium-scale DSGE model. Upon controlling for typical issues of the VAR framework (e.g. sample size, lag order selection and model specification), the author concludes that the documented close-to-muted monetary shock responses over the Great Moderation period are an artifact of the Cholesky-style identification scheme rather than an actual feature of the data generating process.⁴

On the relationship between DSGE models and recursively identified VARs, Altig et al. (2011) document the good performance of Cholesky-VARs in identifying monetary impulse responses in a three-shock business cycle model featuring a lagged transmission mechanism. Carlstrom et al. (2009) show that the theoretical IRFs produced by the standard, unrestricted DNK small-scale model and its version featuring timing restrictions are qualitatively similar at all time periods subsequent to the one where the monetary policy shock materializes; estimated IRFs delivered by a recursively identified VAR rather display, *in population* (i.e. assigning the econometrician full knowledge of the model's reduced form structure other than the true structural shocks), considerably muted patterns, conditional on the underlying parameterization – see Figs. 1 to 5.⁵

While this evidence clearly supports the existence of a severe mis-identification issue in Cholesky-identified VARs when the true data generating process does not warrant a Cholesky structure, it might also suggest that timing restrictions in DNK models would not produce significant changes in the qualitative properties of the monetary transmission mechanism. We argue that this conclusion is unwarranted. In a rather general setting, we show that timing assumptions that reproduce the standard Cholesky identification strategy can generate substantial differences in the CERs and the persistence properties of the theoretical infinite VAR representation of the observables relative to their unrestricted counterparts. This is so since DSGE models exhibiting timing restrictions necessarily give rise to VARMA equilibrium representations for the observables. From an inferential point of view, even when impulse responses are rightly constrained by the Cholesky scheme to

ing of variables in Cholesky-VARs has spurred interest in the development of structural macroeconomic models exhibiting some degree of recursiveness, e.g. Rotemberg and Woodford (1997) and Christiano et al. (2005).

⁴See Wolf (2020) for similar remarks on what short-run exclusion restrictions are likely to identify in VAR models, when the data generating process is the equilibrium representation of a structural macroeconomic model.

⁵To be clear, all the results in Carlstrom et al. (2009) depend on neither small sample length nor truncation bias (the DNK model admits a finite second order VAR representation and the econometrician is assumed to select the same lag order and to precisely estimate all the autoregressive coefficients).

be an exact match to the theoretical ones on impact, a finite order Cholesy-VAR can still prove a very inaccurate approximation of the true adjustment paths of endogenous variables to monetary policy shocks (e.g. Ravenna (2007) and Poskitt and Yao (2017)).

We explore this issue using two different DNK structures as data generating processes (DGP). The first one is the sticky-price framework popularized by Benati and Surico (2009) and fruitfully employed in policy and academic circles to conduct business cycle analysis.⁶ The second is the DNK model with an active fiscal-passive monetary policy regime advanced by Leeper and Leith (2016) in their thorough investigation of the fiscal and monetary origins of inflation.

While virtually any DSGE model can be adapted to feature information-based timing restrictions, there are at least two good reasons for our choice of the mentioned DNK model laboratories. First, since structural (deep) parameters of DSGE models should be invariant to the timing of decisions as enforced by informational constraints, calibrating these parameters to the posterior medians reported in Benati and Surico (2009) allows us to feed our baseline Cholesky-VARs with artificial data generated by an empirically validated framework, as far as the sticky-price monetary model is concerned. Second, notwithstanding their simplicity, both DNK models are flexible enough to allow us to scrutinize the role of internal propagation mechanisms in generating persistent effects of timing restrictions, and how this feeds back into the identification problem sketched above.

As mentioned, a second contribution of this paper is to subject the connection between timing restrictions and nonfundamentalness to formal investigation. It is well-known that identifying and quantifying the relative importance of an economy's true (structural) shocks via structural VAR analysis require that the observed variables contain sufficient information to recover the unobserved state variables of interest which are assumed to produce the observables. In particular, if a given DSGE model's solution (i.e. the data generating process) does not admit an invertible (or fundamental) representation, then there exists no (linear) invertible mapping from the VAR innovations to the structural shocks of the underlying economy. Obviously, nonfundamentalness implies non-existence of a finite-order VAR approximation of the true model's dynamics. Since the empirical

⁶Lubik and Schorfheide (2004), Fanelli (2012) and Castelnuovo and Fanelli (2015) have studied the consequences of changes in the monetary policy stance for the U.S. business cycle and macroeconomic stability; Boivin and Giannoni (2006) and Benati and Surico (2009) have investigated salient empirical facts from the U.S. Great Moderation; Castelnuovo and Surico (2010) put forward a new theoretical interpretation of the VAR evidence on the positive response of prices to a monetary policy shock; Fanelli and Sorge (2017) shed light on the implications of equilibrium indeterminacy for the forecast ability of standard reduced-form prediction models.

question addressed herein is whether the structural monetary shocks of a restricted DSGE model can be retrieved from a finite-order VAR on the observable variables, we need to address the question whether, and under what conditions, timing restrictions serve as a source of nonfundamentality.

Differently from RE structures with persistently dispersed information (e.g. Kasa (2000), Kasa et al. (2014)), the specification of information-based timing restrictions does not involve an infinite regress of expectations, and the underlying model's representation will be finite dimensional. However, while being fundamental in terms of the innovations to agents' information sets, this representation can prove nonfundamental (with respect to any set of observables) in terms of the structural shocks.⁷

We construct (necessary and sufficient) conditions that pin down existence of fundamental representations of general DSGE models featuring timing restrictions. We next show that such conditions are fulfilled by standard DNK environments under empirically plausible parameterizations, thus validating the use of VAR-based approaches to evaluation of monetary impulse responses. Inspection of the evolution of the coefficients of the theoretical VAR(∞) reveals that they rapidly decay towards zero, mitigating concerns about lag truncation bias arising from the adoption of low-order VAR systems: our analysis shows that VAR(p) systems with $p \leq 3$ are enough to recover qualitative and quantitative characteristics of monetary impulse responses in restricted DNK models, for a sufficiently large sample size.⁸

Our paper clearly speaks to the rapidly increasing literature on imperfect information in models of the business cycle. On the one hand, scholars have been interested in exploring the role of imperfect information in fueling the propagation of structural shocks in otherwise standard stochastic environments – e.g. Nimark (2008), Nimark (2014), Angeletos and La'O (2013), Acharya et al. (2017). On the other, attention has been focused on the conditions under which imperfect information qualifies as an alternative mechanism that supports self-fulfilling expectations and sunspot fluctuations in models that would rather exhibit a determinate equilibrium in the presence of perfect and sym-

⁷It should be emphasized that DSGE models with timing restrictions due to informational constraints do not involve any kind of *parameter variation*, either exogenous or endogenous; time-invariant structures will then dictate optimal linearized decision rules featuring time-invariant coefficients for the enlarged set of state variables. See Kulish and Pagan (2017) for a method for constructing and estimating solutions for linearized models with (actual or perceived) structural changes; Canova et al. (2020) for an exploration of inferential issues related to time-varying structural macroeconomic models; and Cavicchioli (2020) for the analysis of the invertibility conditions for DSGE models with Markovian coefficients.

⁸Truncation effects reflect existence of both an approximation error and an estimation error, that arise even when the identification scheme is consistent with the structure of the macroeconomic model that generates the observables, see e.g. Ravenna (2007) and Poskitt and Yao (2017).

metrically shared information – e.g. Lubik et al. (2020). As emphasized in Sorge (2020), information-based timing restrictions naturally embed an informational asymmetry across economic agents, which in turn affects how beliefs are formed with respect to the stochastic unfolding of economic variables. While requiring rather different mutual consistency conditions for the beliefs of asymmetrically informed agents to coordinate into an RE equilibrium, differential information processing in this setting generically preserves the saddle-path properties of the underlying model economy: if the unrestricted model displays a determinate equilibrium, so will its restricted counterpart. Timing restrictions can however produce nonfundamental or close to nonfundamental (in the time series sense) equilibrium representations, thereby calling for a careful analysis of the mapping between restricted DSGE structures and VAR models.

The paper proceeds as follows: Section (2) motivates our analysis by means of a simple analytical example. Section (3) formally discusses how to construct first-order approximate solutions to general DSGE models with Cholesky-timing restrictions. Section (4) provides easy-to-check conditions under which restricted DSGE models admit a fundamental equilibrium representation for the observables. Section (5) lays out the DNK model laboratories, that are next used to explore the ability of Cholesky-VARs to identify and estimate theoretical IRFs generated by a monetary policy shock (Section (6)). Section (7) offers concluding remarks.

2 An illustrative example

Answering the question of whether the structural shocks of a given multivariate macroeconomic model can be recovered from a VAR analysis requires understanding if, and under what conditions, such a model can be represented as a reduced-form VAR. To build intuition on how timing restrictions influence the mapping from the DSGE structure to its equilibrium representation, let us consider the following simple bi-variate RE system

$$\mathcal{E} [y_{1,t} - \alpha y_{1,t+1} - y_{2,t} - x_{1,t} | \mathbb{I}_{1,t}] = 0, \quad \alpha > 0, \quad (1)$$

$$\mathcal{E} [y_{2,t} - \beta y_{2,t+1} - x_{1,t} - x_{2,t} | \mathbb{I}_{2,t}] = 0, \quad \beta > 0, \quad (2)$$

$$x_{1,t} = \rho_1 x_{1,t-1} + \varepsilon_{1,t}, \quad |\rho_1| < 1, \quad (3)$$

$$x_{2,t} = \rho_2 x_{2,t-1} + \varepsilon_{2,t}, \quad |\rho_2| < 1, \quad (4)$$

$$\varepsilon_{1,t} \sim NID(0, 1), \quad \varepsilon_{2,t} \sim NID(0, 1) \quad (5)$$

where $y_t = (y_{1,t}, y_{2,t})'$ are endogenous (control) variables, $x_t = (x_{1,t}, x_{2,t})'$ are exogenous (state) variables, and $\mathcal{E}[\cdot|\mathbb{I}_{i,t}]$ denotes the expectation operator accounting for potential informational constraints embedded in the conditioning sets $\mathbb{I}_{i,t}$, $i = 1, 2$ (all variables are defined on a common filtered probability space).

In the unrestricted case, $\mathbb{I}_{1,t}$ and $\mathbb{I}_{2,t}$ both coincide with the smallest closed linear subspace \mathbb{I}_t spanned by the semi-infinite history of all the model's variables up to time t , i.e. $\mathbb{I}_t = \mathbb{V}_t(y^t) \vee \mathbb{V}_t(x^t)$. Letting $E_t[\cdot] = \mathcal{E}[\cdot|\mathbb{I}_t]$, the RE system (1)-(5) can be cast in the conventional form

$$y_t = AE_t y_{t+1} + Bx_t \quad (6)$$

where

$$A = \begin{pmatrix} \alpha & \beta \\ 0 & \beta \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad (7)$$

When $\alpha, \beta \in (0, 1)$, the number of explosive roots associated with A^{-1} equals the number of jump variables, and thus the model features a locally unique (determinate) equilibrium of the form

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \frac{2-\beta\rho_1}{(1-\alpha\rho_1)(1-\beta\rho_1)} & \frac{1}{(1-\alpha\rho_2)(1-\beta\rho_2)} \\ \frac{1}{1-\beta\rho_1} & \frac{1}{1-\beta\rho_2} \end{pmatrix} \begin{pmatrix} x_{1,t} \\ x_{2,t} \end{pmatrix} \quad (8)$$

which generically admits a finite-order VAR representation.⁹

Assume now the control variable $y_{1,t}$ is optimally set prior to the realization of the exogenous variable $x_{2,t}$. This information-based timing restriction is simply captured by letting $\mathbb{I}_{1,t} = \mathbb{V}_t(y_1^t, y_2^{t-1}) \vee \mathbb{V}_t(x_1^t, x_2^{t-1}) \subset I_{2,t} = \mathbb{V}_t(y^t) \vee \mathbb{V}_t(x^t)$. Provided $\alpha, \beta \in (0, 1)$, the restricted model will exhibit the same determinacy properties as its unrestricted counterpart (Sorge (2020)). Notice however that, in contrast with the unrestricted case, the endogenous variable $y_{1,t}$ varies with the shock process $x_{2,t}$ (and functions of it) only with delay, and yet immediately adjusts in reaction to the optimal estimate of the latter on the basis of information contained in $\mathbb{I}_{1,t}$; this implies that, generically, $\mathcal{E}[y_{2,t}|\mathbb{I}_{1,t}] \neq y_{2,t}$.

A straightforward application of the method of undetermined coefficients reveals that the RE solution under timing restrictions has the following representation

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \frac{2-\beta\rho_1}{(1-\alpha\rho_1)(1-\beta\rho_1)} & 0 & \frac{\rho_2}{(1-\alpha\rho_2)(1-\beta\rho_2)} \\ \frac{1}{1-\beta\rho_1} & \frac{1}{1-\beta\rho_2} & 0 \end{pmatrix} \begin{pmatrix} x_{1,t} \\ x_{2,t} \\ x_{2,t-1} \end{pmatrix} \quad (9)$$

⁹Specifically, provided the square coefficient matrix in (8) is non-singular, the determinate solution is in VAR(1) form, see Morris (2016).

which admits the factorized MA form

$$y_t = M_1(L) \cdot M_2(L) \varepsilon_t, \quad (10)$$

$$M_1(L) = \begin{pmatrix} \frac{2-\beta\rho_1}{(1-\alpha\rho_1)(1-\beta\rho_1)} & \frac{\rho_2}{(1-\alpha\rho_2)(1-\beta\rho_2)} \cdot L \\ \frac{1}{1-\beta\rho_1} & \frac{1}{1-\beta\rho_2} \end{pmatrix}, \quad (11)$$

$$M_2(L) = \begin{pmatrix} (1-\rho_1 L)^{-1} & 0 \\ 0 & (1-\rho_2 L)^{-1} \end{pmatrix} \quad (12)$$

Existence of a causal VAR representation of the RE model obtains is warranted when the determinant of the filter $M_1(L)$ only vanishes within the unit circle. For some parametric configurations of the RE system (1)-(5) this will be the case, and yet the inversion of the MA filter will result in VAR polynomial matrices whose coefficients slowly decline toward zero as the lag counter increases. As an example, when $\alpha = 0.8$, $\beta = 0.69$, $\rho_1 = 0.45$ and $\rho_2 = 0.84$, the evolution of the coefficients of the theoretical VAR matrices with and without timing restrictions at different lags is markedly different, see Figure (1). Under these circumstances, the true DGP is a VARMA with a non-trivial MA component, and fitting a finite-order VAR to data generated from the restricted RE model, albeit correctly identified, may produce highly inaccurate estimates of impulse responses, for they involve non-linear (at horizons larger than one) functions of biased estimates of the truncated VAR coefficients.

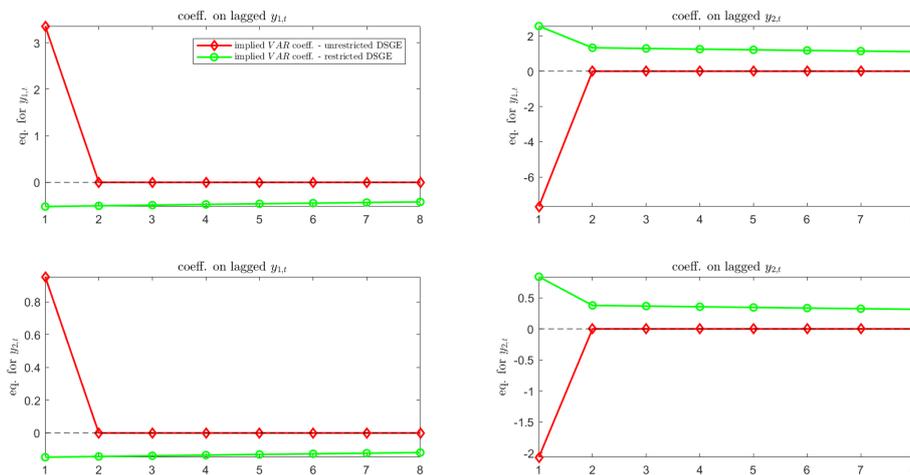


Figure 1: Evolution of coefficients of the theoretical VAR representation of the RE model (1)-(5) with and without timing restrictions.

Non-fundamentality is also an issue here. In fact, for generic parametric configura-

tions, such as e.g. $\beta \geq \alpha = 0.8$, $\rho_1 \in (0.1, 0.3)$, $\rho_2 \in (0.9, 0.99)$, there exists no (linear) invertible mapping from the empirical VAR innovations to the structural shocks of the RE model. As a result, the theoretical impulse responses to shocks $\varepsilon_{i,t}$ ($i = 1, 2$) cannot be retrieved from the history of observables y_t , no matter which identification strategy is adopted.

3 Environment: DSGE models with timing restrictions

Once reduced to its first-order optimality conditions, DSGE models under RE are typically described by a system of n_F expectational stochastic difference equations of the form

$$E_t f(y_{t+1}, y_t, x_{t+1}, x_t; \sigma) = 0 \quad (13)$$

where the random processes (y_t) and (x_t) are defined on the same probability space, and E_t is the standard (conditional) expectation operator associated with the underlying probability measure. The n_y -dimensional vector y collects the model's endogenous jump variables, whereas the n_x -dimensional vector x contains endogenous predetermined as well as exogenous states. Finally, the scalar $\sigma \geq 0$ is used to scale the size of aggregate uncertainty surrounding the economy, see Schmitt-Grohé and Uribe (2004).

3.1 Unrestricted model

To ease notation, let the prime superscript denotes one-step ahead variables. In the standard unrestricted case, policy functions for all the endogenous variables depend on all the state variables x . Time-invariant, analytic solutions to (13) are in the form

$$y = g(x, \sigma), \quad x' = h(x, \sigma) + \sigma \epsilon' \quad (14)$$

where the elements of the n_x -dimensional vector ϵ are i.i.d. zero-mean, unit variance innovations (e.g. structural shocks).

As shown in Schmitt-Grohé and Uribe (2004), up to first order certainty equivalence holds generically, and therefore σ does not enter the linearly perturbed model's dynamics for endogenous variables, i.e. one has

$$y = g_x x, \quad x' = h_x x + \sigma \epsilon' \quad (15)$$

where g_x and h_x are conformable matrices of first-order derivatives of the maps $g(x, \sigma)$ and $h(x, \sigma)$ with respect to x , evaluated at the non-stochastic steady state (\bar{y}, \bar{x}) that solves (13) when $\sigma = 0$.

3.2 Restricted model

Let us now turn to studying the implications of information-based timing restrictions. To this end, we consider the most basic case where any given time period is split into two informational subperiods, according to the timing of the model's variables. Formally, the control and state vectors are accordingly partitioned as

$$y = [y_u; y_r], \quad x = [x_u; x_r] \quad (16)$$

where the n_{x_u} -dimensional vector x_u consists of endogenous predetermined as well as exogenous variables which materialize in the beginning of the first subperiod, x_r contains n_{x_r} exogenous variables with realizations in the second subperiod, y_u is the n_{y_u} -dimensional vector of fully endogenous jump variables, i.e. endogenous variables which are conditioned on all the state variables x . Finally, the n_{y_r} -dimensional vector y_r collects partially endogenous variables, which are decided upon in the first subperiod, when realizations of only a subset of state variables are known. To apply Kormilitsina (2013)'s solution approach, the RE system (13) is partitioned as follows

$$f = [f^0; f^1; f^{x_r}] \quad (17)$$

so that the sub-system f^0 includes n_{y_r} equations determining endogenous variables y_r , the sub-system f^1 includes n_{y_u} equations that determine endogenous variables y_u and n_{x_u} equations determining the dynamics of the states x_u , and the sub-system f^{x_r} describes the evolution of exogenous shocks x_r , represented as a first-order stationary autoregressive process

$$x'_r = Px_r + \sigma \epsilon'_{x_r}, \quad \epsilon_{x_r} \sim i.i.d.N(0, V_{\epsilon_{x_r}}) \quad (18)$$

where P is a stable square matrix of autoregressive coefficients, and ϵ'_{x_r} collects the n_{x_r} shocks associated with the states x_r .

The RE system with timing restrictions can be equivalently rewritten as

$$\mathcal{E}_t[f(y', x', y, x)] = 0 \quad (19)$$

and its recursive solution represented in general form as

$$y_u = \hat{g}(x_u, x_r, x_{r,-1}, \sigma), \quad y_r = \hat{j}(x_u, x_r, x_{r,-1}, \sigma), \quad x'_u = \hat{h}(x_u, x_r, x_{r,-1}, \sigma) + \sigma \epsilon'_{x_r} \quad (20)$$

Endogenous (jump) variables in y_r only react to the conditional forecast of states in x_r (a function of previous period variables $x_{r,-1}$), as the latter do not belong in the first subperiod information set. By the same token, endogenous (jump) variables in y_u are a function of y_r – a state variable in the second informational subperiod – and thus of lagged states $x_{r,-1}$. Notice that the timing restrictions only involve exogenous variables x_r which are uncorrelated with other exogenous variables in x ; also, all the x_r variables are not observed in the first subperiod, hence the filtering problem does not require using the variance-covariance matrix of the ϵ_{x_r} shocks in order to compute an optimal (in the mean-square sense) estimate of unobserved states.

As established in Kormilitsina (2013), linearly perturbed DSGE models with timing restrictions comply with the certainty equivalence principle, so that the first-order approximation to (20) is

$$\begin{aligned} y_u &= \hat{g}_{x_u} x_u + \hat{g}_{x_r} x_r + \hat{g}_{x_{r,-1}} x_{r,-1}, \\ y_r &= \hat{j}_{x_u} x_u + \hat{j}_{x_{r,-1}} x_{r,-1}, \\ x'_u &= \hat{h}_{x_u} x_u + \hat{h}_{x_r} x_r + \hat{h}_{x_{r,-1}} x_{r,-1} + \sigma \epsilon'_{x_u} \end{aligned} \quad (21)$$

or in more compact form

$$y = \hat{g}_x \begin{pmatrix} x_u \\ x_r \\ x_{r,-1} \end{pmatrix}, \quad x' = \hat{h}_x \begin{pmatrix} x_u \\ x_r \\ x_{r,-1} \end{pmatrix} + \sigma \epsilon' \quad (22)$$

where

$$\hat{g}_x = \begin{pmatrix} \hat{g}_{x_u} & \hat{g}_{x_r} & \hat{g}_{x_{r,-1}} \\ \hat{j}_{x_u} & 0_{n_{y_r} \times n_{x_r}} & \hat{j}_{x_{r,-1}} \end{pmatrix}, \quad \hat{h}_x = \begin{pmatrix} \hat{h}_{x_u} & \hat{h}_{x_r} & \hat{h}_{x_{r,-1}} \\ 0_{n_{x_r} \times n_{x_r}} & P & 0_{n_{x_r} \times n_{x_r}} \end{pmatrix} \quad (23)$$

Sorge (2020) formally shows that, provided a given full rank condition is met, the restricted model (19) admits a unique, dynamically stable first-order approximate solution whenever its unrestricted counterpart (13) does. In this case, the solution to the restricted model can be readily constructed via linear transformations of (15), however computed (e.g.

Klein (2000); Christiano (2002); King and Watson (2002); Sims (2002)).¹⁰

Upon partitioning the equilibrium coefficient matrices (g_x, h_x) in (15) as follows

$$g_x = \begin{pmatrix} g_{x_u} & g_{x_r} \\ j_{x_u} & j_{x_r} \end{pmatrix}, \quad h_x = \begin{pmatrix} h_{x_u} & h_{x_r} \\ 0 & P \end{pmatrix} \quad (24)$$

we can easily map the full information coefficient matrices into those appearing in (22), i.e.

$$\hat{g}_x = \begin{pmatrix} g_{x_u} & g_{x_r} + (\nabla(f^1)^{-1} f_{y_r}^1 j_{x_r})_{n_{y_u}} & -(\nabla(f^1)^{-1} f_{y_r}^1 j_{x_r} P)_{n_{y_u}} \\ j_{x_u} & 0_{n_{y_r} \times n_{x_r}} & j_{x_r} P \end{pmatrix}, \quad (25)$$

$$\hat{h}_x = \begin{pmatrix} h_{x_u} & h_{x_r} + [\nabla(f^1)^{-1} f_{y_r}^1 j_{x_r}]_{n_{x_u}} & [-\nabla(f^1)^{-1} f_{y_r}^1 j_{x_r} P]_{n_{x_u}} \\ 0 & P & 0 \end{pmatrix} \quad (26)$$

where $\nabla(f^1)$ denotes the Jacobian of the sub-system f^1 with respect to the vector $[x'_u, y_u]$, $f_{y_r}^1$ is the matrix of partial derivatives of f^1 with respect to the partially endogenous variables collected in the vector y_r , and $[M]_m$ is used to denote the selection of the first (or last) m rows of some matrix M .

We remark that virtually any DSGE model can be adapted to feature timing restrictions, for they can be directly applied to equilibrium conditions (whether nonlinear or linearized) in a way that does not violate any of the RE postulates; as a matter of fact, the restricted equilibrium representation (22) is equivalent to the first-order approximation of the same underlying model where nested information sets are specified at the outset. In the Appendix we report full details on the solution algorithm developed in Kormilitsina (2013), and apply it to the illustrative example discussed in Section (2).

4 Timing restrictions and VAR representations of DSGE models

The foregoing arguments clarify that the information-based timing restrictions engender an enlarged state space as well as an increased degree of backward dependence in policy

¹⁰Hespeler and Sorge (2019) devise a simple model economy where timing restrictions force a mismatch between the number of fully endogenous variables and the number of equations that are required to pin down them. Whenever this is the case, generic singularity of the Jacobian $\nabla(f^1)$ obtains for it reflects a fundamental failure of regularity requirements which are implicitly invoked for the algorithmic implementation of computational routines in standard (unrestricted) RE systems, see e.g. Klein (2000).

functions. Since all the non-zero eigenvalues of the companion matrix

$$h_x^\dagger = \begin{pmatrix} h_{x_u} & h_{x_r} + [\nabla(f^1)^{-1} f_{y_r}^1 j_{x_r}]_{n_{x_u}} & [-\nabla(f^1)^{-1} f_{y_r}^1 j_{x_r} P]_{n_{x_u}} \\ 0 & P & 0 \\ 0 & I & 0 \end{pmatrix} \quad (27)$$

are those of h_{x_u} and P , the first-order approximate solution under timing restrictions is dynamically stable (Sorge, 2020). The ensuing equilibrium MA representation for (all) the endogenous variables included in the y vector is

$$y_t = C(L)x_t = C(L)(I - A(L))^{-1} \sigma \epsilon_t = B(L)\epsilon_t \quad (28)$$

where L stands for the conventional lag operator and

$$C(L) = (\hat{g}_{x_u}, \hat{g}_{x_r} + \hat{g}_{x_r, -1}L), \quad A(L) = (\hat{h}_{x_u}, \hat{h}_{x_r} + \hat{h}_{x_r, -1}L) \quad (29)$$

Recall that the IRFs of the elements of y_t to a unit shock in one of the elements of ϵ_t occurring at some time $t = s$ are given by the sequences of corresponding MA coefficients in (28) from time s onward. These coefficients functionally depend on those entering the $C(L)$ filter, which in turn depends on the elected structure of timing restrictions as embodied in the partition-based Jacobian $\nabla(f^1)$. As a consequence, theoretical impulse response functions (IRFs) associated with a DSGE model under timing restrictions generically differ from those implied by its unrestricted counterpart over all time horizons.

The equilibrium MA representation (28) may in principle fail to invert into a causal autoregressive one, thereby preventing VAR methods from recovering the truly structural economic shocks. Non-fundamentalness is known to be a generic issue in square systems, i.e. when $n_y = n_x$, e.g. (Alessi et al., 2011).

The reduced form of the general model (13) under timing restrictions will be non-fundamental (and thus non-invertible in the past) if and only if the determinant of the filter $B(L)$ vanishes within the unit circle. By virtue of the stability property of $A(L)$, an easy-to-check condition for fundamentalness of the first-order approximate solution can be stated as follows¹¹

Condition 1 (Fundamentalness). *Let $n_y = n_x$. Then, for any given informational partition (16)-(17), the equilibrium representation under timing restrictions is fundamental*

¹¹See Rozanov and Rozanov (1967) for a comprehensive discussion of fundamentalness and invertibility properties for stationary random processes.

if and only if

$$\text{Det} \begin{pmatrix} g_{x_u} & g_{x_r} + [\nabla(f^1)^{-1} f_{y_r}^1 j_{x_r}]_{n_{y_u}} - [\nabla(f^1)^{-1} f_{y_r}^1 j_{x_r} P]_{n_{y_u}} \cdot z \\ j_{x_u} & j_{x_r} P \cdot z \end{pmatrix} = 0 \implies |z| \geq 1 \quad (30)$$

Even when fundamentalness is warranted by the model's parameterization, the equilibrium dynamics (28) need not admit a finite-order (causal) VAR representation, opening room to lag truncation bias in IRFs estimation exercises, e.g. Chari et al. (2008) and Socorsi (2016). As is known, approximating VAR(∞) representations with finite-order VAR systems may severely distort estimation the structural IRFs, whatever the identification strategy adopted by the researcher (Ravenna (2007), Poskitt and Yao (2017)). Notice that the state space representation

$$y_t = C(L)A(L)x_{t-1} + C(L)\sigma\epsilon_t, \quad (31)$$

$$x_t = A(L)x_{t-1} + \sigma\epsilon_t \quad (32)$$

is tied directly to the CERs of the RE model under timing restrictions and thus involves the state vector x_t of the smallest dimension possible for replicating the dynamic properties of the reduced form equilibrium process for the observables. The following condition will therefore fully characterize existence of a finite order VAR representation for the observables in the square case (Franchi and Vidotto, 2013)

Condition 2 (Finite order VAR representation). *Let $n_y = n_x$ and define the matrices*

$$I^\circ = \begin{pmatrix} I_{n_x \times n_x} \\ 0_{n_{x_r} \times n_x} \end{pmatrix}, \quad (33)$$

$$g_x^\dagger = \begin{pmatrix} g_{x_u} & g_{x_r} + [\nabla(f^1)^{-1} f_{y_r}^1 j_{x_r}]_{n_{y_u}} \\ j_{x_u} & 0_{n_{y_r} \times n_{x_r}} \end{pmatrix}, \quad (34)$$

$$g_x^\circ = \begin{pmatrix} g_{x_u} & g_{x_r} + [\nabla(f^1)^{-1} f_{y_r}^1 j_{x_r}]_{n_{y_u}} - [\nabla(f^1)^{-1} f_{y_r}^1 j_{x_r} P]_{n_{y_u}} \\ j_{x_u} & 0_{n_{y_r} \times n_{x_r}} \quad j_{x_r} P \end{pmatrix} \cdot h_x^\dagger \quad (35)$$

Then, provided g_x^\dagger is non-singular, for any given informational partition (16)-(17), a finite order VAR equilibrium representation under timing restrictions exists if and only if the $(n_x + n_{x_r})$ -dimensional square matrix

$$F := h_x^\dagger - I^\circ \cdot (g_x^\dagger)^{-1} \cdot g_x^\circ \quad (36)$$

is nilpotent – i.e. all its eigenvalues are equal to zero.

We emphasize that both conditions (1) and (2) are expressed in terms of the reduced form coefficients of the RE solution to the unrestricted model (13), and can thus be checked whatever the solution algorithm employed, once the informational partition (17) has been devised.¹²

5 Model laboratories

As mentioned, we employ two distinct model laboratories to conduct our numerical experiments: the hybrid sticky-price model put forward in Benati and Surico (2009) in their VAR analysis of the so-called Great Moderation; and a dynamic production model with an passive monetary-active fiscal policy regime, popularized in Leeper and Leith (2016) that expands on the standard forward-looking DNK model by allowing fiscal policy to influence the real economy.

5.1 Benati and Surico (2009)’s model

The model is given as follows

$$g_t = \gamma E_t g_{t+1} + (1 - \gamma)g_{t-1} - \delta(i_t - E_t \pi_{t+1}) + \omega_t^g \quad (37)$$

$$\pi_t = \frac{\beta}{1 + \beta\alpha} E_t \pi_{t+1} + \frac{\alpha}{1 + \beta\alpha} \pi_{t-1} + \kappa g_t + \omega_t^\pi \quad (38)$$

$$i_t = \rho i_{t-1} + (1 - \rho)(\varphi_\pi \pi_t + \varphi_g g_t) + \omega_t^i \quad (39)$$

where

$$\omega_t^j = \rho_j \omega_{t-1}^j + \epsilon_t^j, \quad |\rho_j| < 1, \quad \epsilon_t^j \sim \text{WN}(0, \sigma_j^2), \quad j = g, \pi, i \quad (40)$$

and expectations are conditional on the information set \mathcal{I}_t , i.e. $E_t \cdot := E(\cdot \mid \mathcal{I}_t)$. The variables g_t , π_t , and i_t stand for the output gap, inflation, and the nominal interest rate, respectively; γ is the weight of the forward-looking component in the dynamic IS curve; α is price setters’ extent of indexation to past inflation; δ is households’ intertemporal elasticity of substitution; κ is the slope of the Phillips curve; ρ , φ_π , and φ_g are the interest

¹²We provide Matlab code that performs this task, that expands on the perturbation approach put forward in Kormilitsina (2013). Notice that, given that the state space representation (31) is in minimal form, a necessary and sufficient condition for the process for y_t to admit an infinite order causal VAR representation is that F be a stable matrix, i.e. its eigenvalues are all strictly below one in modulus, see Fernández-Villaverde et al. (2007) and Franchi and Paruolo (2015).

rate smoothing coefficient, the long-run coefficient on inflation, and that on the output gap in the monetary policy rule, respectively; finally, ω_t^g , ω_t^π and ω_t^i in eq. (40) are the mutually independent, AR(1) disturbances and ϵ_t^g , ϵ_t^π and ϵ_t^i are the structural shocks.

The DNK model (37)-(38)-(39) is submitted to Cholesky-timing restrictions according to which (i) the monetary policy shocks are orthogonal to the non-policy variables (g_t, π_t) , and (ii) these non-policy variables are thus predetermined with respect to the policy instrument, i.e. the nominal interest rate. A relevant characteristic of this model environment is that it always predicts conventional macroeconomic reactions to a monetary policy shock for this entails recessionary and deflationary effects that are only delayed when timing restrictions are at work. Monetary impulse responses are thus qualitatively similar across the two information structures (restricted versus unrestricted), and eventually follow the same exact pattern when the effect of the policy shock dies out.

Cholesky-timing assumptions in Benati and Surico (2009)'s model are enforced by the following assignment of variables:

$$\begin{aligned} y_u &= i, & y_r &= [g, \pi]' \\ x_u &= [g_{-1}, \pi_{-1}, \omega^g, \omega^\pi]', & x_r &= \omega^i \end{aligned} \tag{41}$$

Let us remark that the recursive structure of the equilibrium representation of the restricted DNK model will have non-policy variables (output gap and inflation) not react to the monetary shock ω^i on impact; and yet both non-policy variables are free to adjust in response to the demand shock ω^g and the cost-push shock ω^π . This is different from the fully recursive ordering of variables in the Cholesky-VAR, where the non-policy variable ordered first is restricted not to react to the shock ordered second. Model-wise, perfectly aligning the DNK structure and the Cholesky-VAR would require using a multiperiod informational partition that reflects the presence of different layers for the partially endogenous variables that are allowed to react to only a subset of the shock other than the monetary policy one, see Kormilitsina (2013). This is however not an issue for our empirical exercises, since the Cholesky scheme is known to be sufficient for the identification of the dynamic response of all variables in the VAR system to the monetary shock, provided the interest rate equation is ordered last, e.g. Christiano et al. (1999).

5.2 Leeper and Leith (2016)'s model

Based on the standard DNK structure, Leeper and Leith (2016)'s model considers a representative agent cashless economy featuring monopolistic competition in the goods

market and nominal rigidities (staggered price setting). One-period nominal bonds B_t sell at price $1/i_t$, where i_t is also the monetary policy instrument; bonds maturity is measured by the rate of decay $\rho \geq 0$ ($\rho = 0$ means one period maturity). Government purchases are zero and the primary government surplus s_t is assumed to evolve exogenously (reflecting the presence of lump-sum taxes). Monetary and fiscal policies are subject to structural disturbances, in addition to shocks to the dynamic IS equation and the NKPC, as in Wolf (2020). Once linearly approximated around the zero inflation non-stochastic steady state, the model dynamics are described by the following equations

$$g_t = E_t[g_{t+1}] - \sigma(i_t - E_t[\pi_{t+1}]) + \omega_t^g \quad (42)$$

$$\pi_t = \beta E_t[\pi_{t+1}] + \kappa g_t + \omega_t^\pi \quad (43)$$

$$i_t = \varphi_\pi \pi_t + \varphi_g g_t + \omega_t^i \quad (44)$$

$$P_t^m = -i_t + \beta \rho E_t[P_{t+1}^m] \quad (45)$$

$$b_t^m = \beta^{-1} b_{t-1}^m + (\rho - 1) P_t^m + (1 - \beta^{-1}) s_t - \beta^{-1} \pi_t \quad (46)$$

$$\omega_t^j = \rho_j \omega_{t-1}^j + \epsilon_t^j, \quad |\rho_j| < 1, \quad j = g, \pi, i \quad (47)$$

$$s_t = \rho_s s_{t-1} + \epsilon_t^s, \quad |\rho_s| < 1 \quad (48)$$

where (45) is no-arbitrage condition linking bond prices to the one-period nominal interest rate and (46) is the flow Government budget identity. The shocks $\epsilon_t^g, \epsilon_t^\pi, \epsilon_t^i$ and ϵ_t^s are taken to be mutually independent white noise sequences.

Following Leeper and Leith (2016) and Wolf (2020), we focus attention on the passive monetary - active fiscal policy regime. The system then delivers RE equilibria for the endogenous variables g_t (output gap), π_t (inflation rate), i_t (nominal interest rate), P_t^m (bond price) and b_t^m (real face value of outstanding debt); determinacy of the RE equilibrium is warranted when e.g. $\varphi_\pi \in [0, 1 - \frac{1-\beta}{\kappa} \varphi_g]$.

As shown in Leeper and Leith (2016), monetary policy shocks that raise the nominal interest rate and depresses output in the short-run can bring about a rise in inflation, whose magnitude and persistence vary non-monotonically with the strength of reaction of the policy rule to fluctuations in the inflation rate. In the passive monetary-active fiscal regime, higher nominal interest rates trigger strong inflationary expectations that overtake the deflationary push flowing from lower output through the Phillips curve (wealth effect). This positive inflation impact response to monetary surprises is pronounced for relatively short debt maturities (ρ equal or close to zero) for they enhance, all else equal, the market value of debt.

Under the conventional Cholesky-timing assumption, we observe that (i) the monetary

policy shocks are orthogonal to the non-policy variables (g_t, π_t) , and (ii) these non-policy variables are thus predetermined with respect to the nominal interest rate. Valuation of debt is also taken not to reflect unexpected changes in the nominal interest rates. This feature, together with the presence of long debt and of high persistence of structural shocks other than the cost-push one, generates a significant (although delayed) contraction in both the inflation rate and the output gap, that eventual rise when the real interest rate and bond prices start declining at longer horizons.

Methodology-wise, the restricted version of the model is fully characterized by the following assignment of variables

$$y_u = [i], \quad y_r = [g, \pi, P^m, b^m]' \quad (49)$$

$$x_u = [s, \omega^g, \omega^\pi, b_{-1}^m]', \quad x_r = [\omega^i] \quad (50)$$

6 Numerical experiments

We now conduct a number of numerical experiments aimed at comparing DNK model-implied impulse responses to those produced by a VAR in (possibly a subset of) the model's variables, in which the monetary policy shock is identified by a Cholesky decomposition of the variance-covariance matrix of the reduced-form residuals. Specifically, we consider a VAR specification involving three observables generated by each of the two DNK models described above – the output gap, the inflation rate and the nominal interest rate – and adopt the Cholesky identification approach by restricting the monetary policy shocks not to have a contemporaneous impact on non-policy variables; since the policy instrument (here, the interest rate) is allowed to react on impact to other structural disturbances, the equation for the nominal interest rate is placed last in the VAR ordering. Notice that, different from Carlstrom et al. (2009), our Cholesky-VAR econometrician is not assumed to know the exact VAR representation of the DNK model generating the data, and is thus concerned with the estimation of the VAR matrix coefficients; parameter identification may thus contaminate the inference on the monetary impulse responses as identified via short-run exclusion restrictions. Lag length selection is data-driven and appeals to the Bayesian information criterion (BIC), see e.g. Lütkepohl (2005).

In order to generate artificial time series for the observables (our DGPs), we first calibrate Benati and Surico (2009)'s model with their posterior estimates over the Great Period¹³ and Leeper and Leith (2016)'s model with the parameters reported in Wolf

¹³See last column of Table 1, Carlstrom et al. (2009), p. 1640.

(2020). Then, *for both the restricted and the unrestricted versions of the model*, artificial data samples are generated by simulating the model’s determinate solution when shock realizations are independently drawn from the assumed densities at any given period. For each simulation a Cholesky-VAR is specified, its lag length selected via the BIC and autoregressive coefficients estimated by a standard Maximum Likelihood technique; estimated IRFs to a normalized monetary policy shock are finally computed and stored.

All our figures below report average Cholesky-VARs monetary impulse responses and plot them against the DGP-consistent IRFs, for both the restricted and the unrestricted version of each model. Operationally, we use $K = 100$ repetitions, $H = 15$ as the IRF horizon, and $T = 1000$ as length of the artificial data sample, with a burn-in of 200 observations.¹⁴

Figures (4) and (7) deliver our comparison results for the benchmark calibration. In the same spirit of Carlstrom et al. (2009), we also re-run the above described numerical exercises by varying either the persistence parameters of exogenous shock processes (Figures (5) and (8)), or some of the key structural parameters that govern the degree of endogenous persistence (hence, the strength of the internal propagation mechanism) of the underlying model (Figures (6) and (9)). Tables (1) collects all the parameter choices for the DNK models under scrutiny; for all of these empirically plausible parameterizations, condition 1 is fulfilled, whereas condition 2 is not: thus, lag truncation bias will necessarily arise when employing a finite-order VAR specification to estimate the equilibrium infinite-order VAR representation of the DNK model with timing restrictions.¹⁵

We argue, however, that truncation effects on the approximating VAR performance are moderate, even when timing restrictions endogenously propagate over time. First, notice that truncation does not induce an identification error due to functional dependencies of the identifying matrix relating structural shocks and reduced-form innovation on the VAR coefficients, since impact exclusion restrictions are employed as an identification

¹⁴Of course, short data samples and measurement errors can undermine the precision of estimates IRFs as in any other inferential problem. Given our focus on the ability of Cholesky-VARS in correctly identifying the true monetary impulses responses when timing restrictions re-shape a model’s internal propagation mechanism, we endow our econometrician with a sufficiently large sample to perform her task. Further investigation of these issues (available on request) suggests that (i) a small sample size produces a standard downward bias and yet does not affect the ability of Cholesky-VARS to closely reproduce the essential shape of the true IRFs generated by restricted models, and (ii) measurement errors generates attenuation effects on point estimates of impulse-response coefficients, lowering all else equal the power of significance tests against the null of a zero response.

¹⁵Notice that, provided $\rho \neq 0$, our expanded version of Leeper and Leith (2016)’s model involves five observables and only four shocks; the ensuing equilibrium MA representation is therefore non-square and *tall*. Fundamentalness in tall systems is known to be a generic property in absence of measurement errors, that we explicitly rule out for our purpose of focusing exclusively on model-implied nonfundamentalness.

Benati and Surico (2009)											
	γ	δ	β	α	κ	ρ	ϕ_π	ϕ_g	ρ_g	ρ_π	ρ_i
<i>Figure 4</i>	0.74	0.12	0.99	0.06	0.044	0.83	1.75	1.15	0.8	0.42	0.4
<i>Figure 5</i>	0.74	0.12	0.99	0.06	0.044	0.83	1.75	1.15	0.9	0.9	0
<i>Figure 6</i>	1	0.12	0.99	0	0.044	0	1.75	1.15	0.8	0.42	0.4

Leeper and Leith (2016)										
	σ	β	κ	φ_π	φ_g	ρ	ρ_g	ρ_π	ρ_i	ρ_s
<i>Figure 7</i>	1	0.95	0.3	0.5	0.5	0	0.3	0.4	0.6	0.6
<i>Figure 8</i>	1	0.95	0.3	0.5	0.5	0.95	0.3	0.4	0.6	0.6
<i>Figure 9</i>	1	0.95	0.3	0.5	0.5	0.95	0.9	0	0.9	0.9

Table 1: *Alternative parameterizations for Benati and Surico (2009) and Leeper and Leith (2016)'s DNK model.*

device (e.g. Ravenna (2007)). Second, the estimated VAR matrices in the finite-order empirical specification will be less prone to pure truncation bias for the true VAR representation features fast decaying coefficients at relatively short horizons. To see this, consider the benchmark calibration of Benati and Surico (2009)'s model, reported in the first row of Table 1, that identifies inertia in the endogenous behavior on the part of households and firms as a non-negligible source of time series persistence. In the presence of Cholesky-timing restrictions, the model admits an infinite-order VAR representation whose polynomial matrices asymptotically converge to zero as a function of the largest (in modulus) eigenvalue of the F matrix defined in (36), see e.g. Ravenna (2007), which is equal to $2.3012e - 06$. The following Figures show the evolution of the coefficients of the theoretical VAR matrices with and without timing restrictions at different lags (Figure (2)), and also the largest eigenvalue (in modulus) of the F matrix associated with the restricted DGP as a function of the deep parameters δ and κ , all else equal (Figure (3)). It is apparent that, although reflecting existence of an approximation error, the truncated VAR system serves as a reasonably good description of the true equilibrium reduced form of the restricted DNK model.¹⁶

¹⁶Notice that Leeper and Leith (2016)'s model suffers from stochastic singularity if all the five endogenous variables are taken to be observable. This is not an issue for our applications, since we do not engage in system-based structural estimation exercises of any kind.

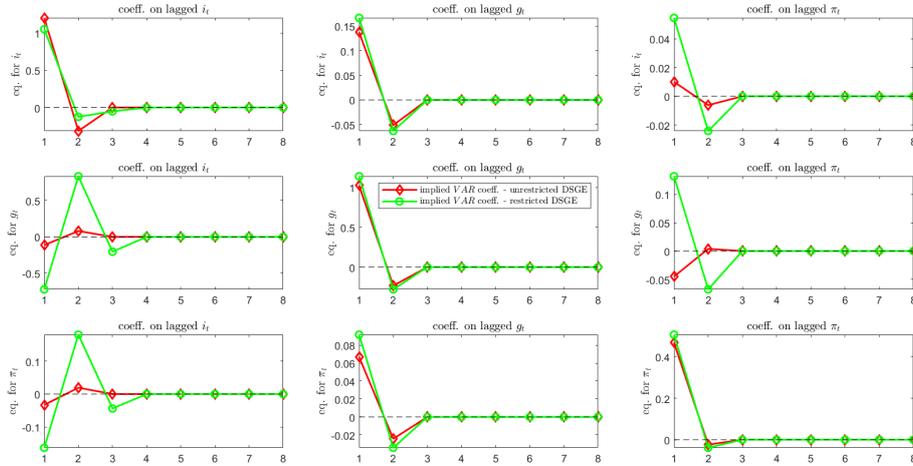


Figure 2: Evolution of coefficients of the theoretical VAR representation of Benati and Surico (2009)'s model with timing restrictions (green) and without timing restrictions (red).

6.1 Benati and Surico (2009) vs. Cholesky-VARs

Figures (4) to (6) contrast the DNK-implied monetary impulse responses with those generated by Cholesky-VARs, when Benati and Surico (2009)'s model is used the DGP. In all the figures, the solid red lines are the true IRFs in the unrestricted model; the solid green lines are the true IRFs in the restricted model; the dashed blue lines are the IRFs identified using Cholesky-VARs when the DGP is the unrestricted model; the solid yellow lines are the IRFs identified using Cholesky-VARs when the DGP is the restricted model. The BIC suggests adopting VAR systems with at most three lags to approximate the true DGP dynamics.

Inspection of Figure (4) reveals the following. First, and in line with findings from Carlstrom et al. (2009) and Castelnuovo (2016), when the unrestricted DNK model generates the data, the Cholesky-identified IRFs systematically underestimate the true responses of all the three endogenous variables to a monetary shock, while also displaying a puzzling upward pattern of variation for the output gap at relatively short horizons. Second, on the assumption that the restricted DNK model serves as the actual DGP, the Cholesky-identified IRFs are correctly signed and similar in shape to their theoretical counterparts, with moderate differences in the estimated magnitudes.

Figure (5) confirms this conclusion by showing that the Cholesky identification strategy produces puzzling positive reactions of inflation and output gap at short horizons when the unrestricted DNK model serves as DGP, and demand- and supply-side struc-

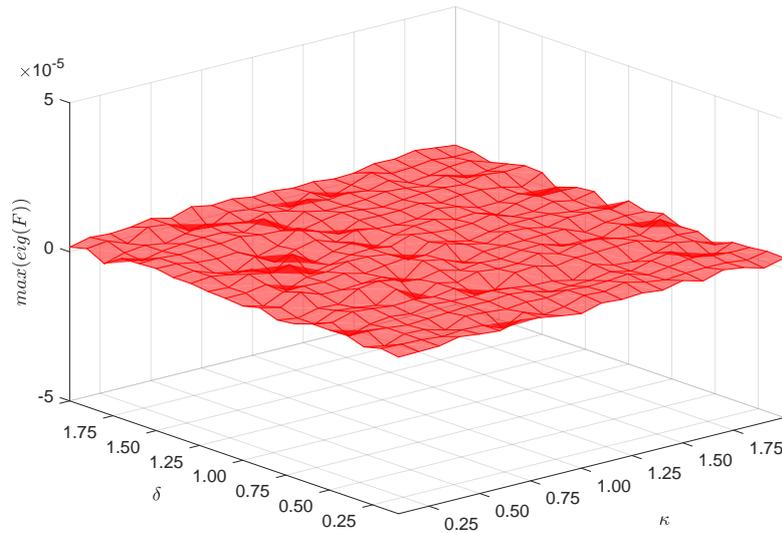


Figure 3: Largest eigenvalue of F matrix in Benati and Surico (2009)'s model under timing restrictions, as a function of households' intertemporal elasticity of substitution (σ) and the slope of the Phillips curve (κ).

tural disturbances feature a large degree of persistence ($\rho_g = \rho_\pi = 0.9$) vis-à-vis a serially uncorrelated monetary policy shock ($\rho_i = 0$). A recursive identification scheme in a non-recursive world thus severely distorts estimates of the adjustment paths of model variables to a monetary policy surprise; it however delivers a fairly reasonable empirical representation of the true IRFs when the model complies with the Cholesky recursive assumptions, even though the autoregressive coefficients of the empirical VAR are estimated with error due to lag truncation bias.

Figure (6) finally illustrates that these astonishing discrepancies in the relative performance of Cholesky-VARs across the restricted and the unrestricted model environments are not due to the presence of a strong internal propagation mechanism that transmits the effects of structural disturbances within the model and over time: even when the DNK model is purely forward-looking ($\gamma = 1$ and $\alpha = 0$) and yet has no recursive structure, a serious mis-representation of true monetary impulse responses will arise when the VAR system is submitted to short-run exclusion restriction. This is not the case for a restricted model-consistent DGP, although the magnitude of short to medium-horizon effects of monetary policy shocks are not fully retrieved (a finding in line with Wolf (2020)).

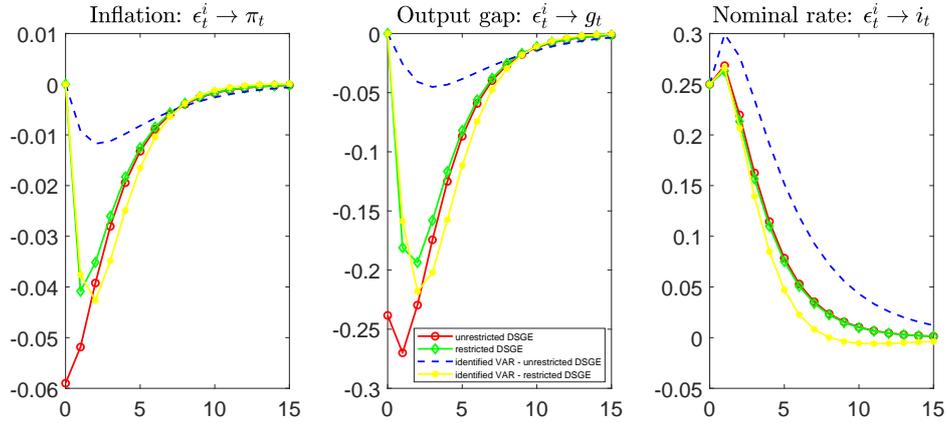


Figure 4: Impulse response functions (IRFs) to a monetary shock for Benati and Surico (2009)'s model with $\rho_g = 0.8$, $\rho_\pi = 0.42$ and $\rho_i = 0.4$.

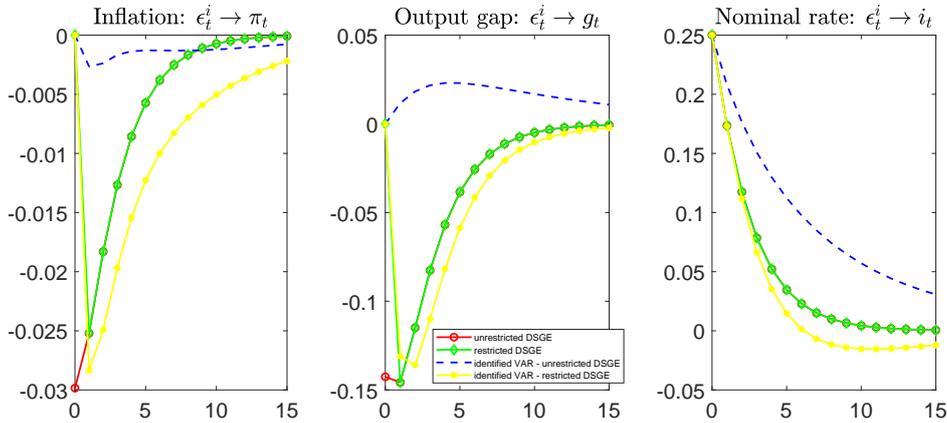


Figure 5: Impulse response functions (IRFs) to a monetary shock for the Benati and Surico (2009)'s model with $\rho_g = 0.9$, $\rho_\pi = 0.9$ and $\rho_i = 0$.

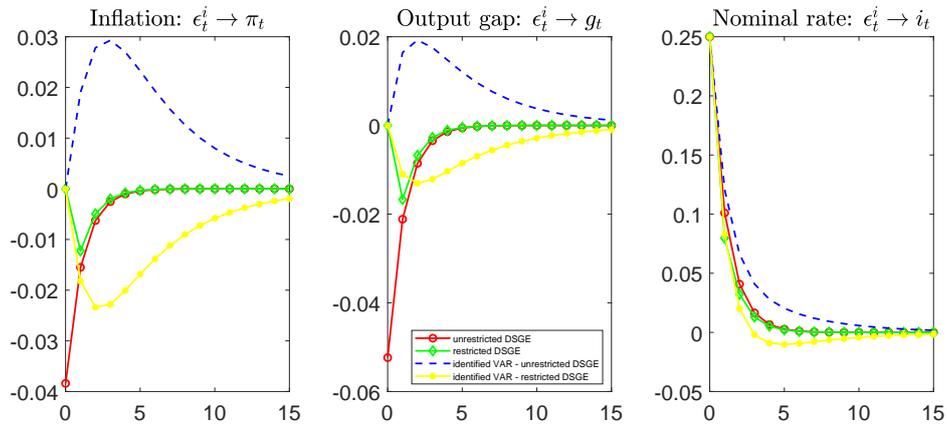


Figure 6: Impulse response functions (IRFs) to a monetary shock for the Benati and Surico (2009)'s model with $\rho_g = 0.8$, $\rho_\pi = 0.4$, $\rho_i = 0.4$, $\gamma = 1$, $\alpha = 0$, and $\rho = 0$.

6.2 Leeper and Leith (2016) vs. Cholesky-VARs

Figures (7) to (9) report the DNK-implied monetary impulse responses against Cholesky-VARs identified ones, when Leeper and Leith (2016)'s model is used as the underlying DGP. As before, in all the figures the solid red lines are the true IRFs in the unrestricted model; the solid green lines are the true IRFs in the restricted model; the dashed blue lines are the IRFs identified using Cholesky-VARs when the DGP is the unrestricted model; the solid yellow lines are the IRFs identified using Cholesky-VARs when the DGP is the restricted model. Again, adoption of the lag length selection information criterion results in the estimation of VAR(p) systems with $p \leq 2$.

Figures (7) and (8) reinforce our claim about the good approximating performance of Cholesky-VARs when applied to a DGP exhibiting a Cholesky-style recursive structure. Most remarkably, Figure (9) shows that information-based timing restrictions dramatically alter the true monetary impulse responses relative to the model's unrestricted counterpart: while both inflation and output gap jump on impact above their long-run values and then remain positive over shorter horizons (up to four periods), they stay in the negative territory when restricted not to react contemporaneously to an unexpected change in the nominal interest rate; and these dynamic adjustment patterns then flip sign at medium to long horizons in the convergence process to steady state.

This strikingly different behavior of the unrestricted and restricted versions of the same underlying DNK structure dramatically impact the ability of Cholesky-VARs to uncover the true monetary impulse responses: estimated IRFs are almost flat at zero for the unrestricted model, whereas are closely replicated if generated by the restricted model.¹⁷

¹⁷This observation also qualifies the results in Wolf (2020) as it points to the possibility of severe mis-identification of monetary IRFs in cases where a standard Cholesky identification scheme is imposed on data generated by a non-recursive model of monetary-fiscal policy interaction.

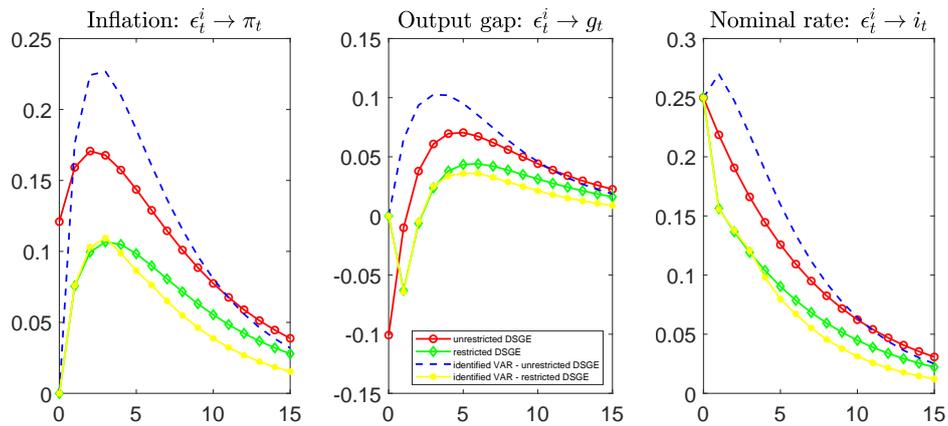


Figure 7: Impulse response functions (IRFs) to a monetary shock for the Leeper and Leith (2016)'s model with $\rho = 0$.

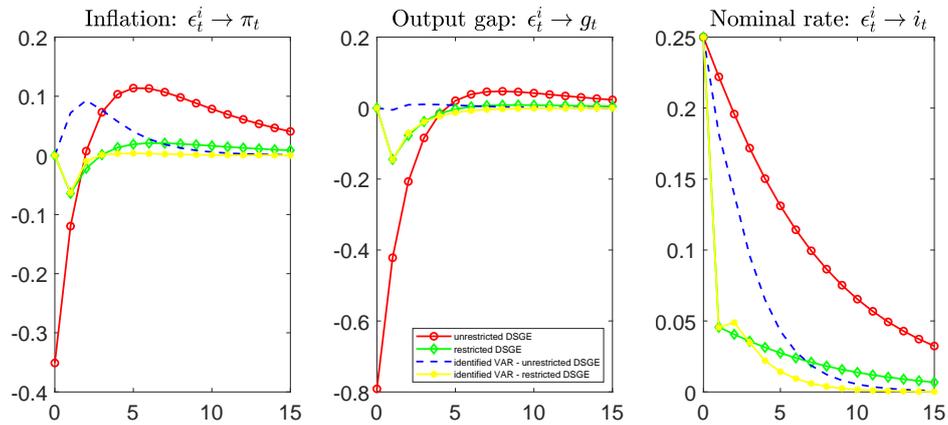


Figure 8: Impulse response functions (IRFs) to a monetary shock for Leeper and Leith (2016)'s model with $\rho = 0.95$.

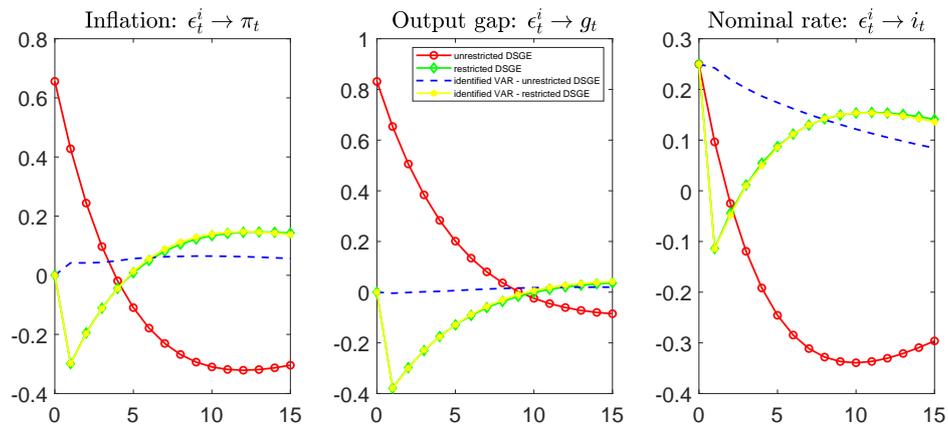


Figure 9: Impulse response functions (IRFs) to a monetary shock for Leeper and Leith (2016)'s model with $\rho = 0.95$ and $\rho_\pi = 0$.

7 Concluding remarks

This paper assesses Cholesky-VARs against model-based data generating processes in controlled DSGE environments. When the building blocks of structural macroeconomic models prove inherently at odds with Cholesky-timing restrictions, the ensuing discrepancy between theoretical and Cholesky VAR-implied dynamic responses can be remarkable, and absolutely impossible to phase out via the impulse response matching procedure.

Moving from the analytical characterization of the effects of Cholesky-timing restrictions on the equilibrium representation of otherwise standard DNK models, this paper establishes that these restrictions are a potential source of non-trivial VARMA equilibrium representations as well as of model-based nonfundamentalness, making inference from VAR systems unreliable.

Simulation results nonetheless suggest that concerns about identification and lag truncation bias are not as serious as one may suspect. The results of our numerical experiments invariably provide support to the view that Cholesky-VARs perform well in identifying the actual monetary impulse responses in restricted DNK structures, no matter whether timing restrictions display long-lasting effects and/or predict substantially different response patterns (in terms of sign, shape and magnitude) relative to the unrestricted benchmark case.

References

- Acharya, S., Benhabib, J., and Huo, Z. (2017). The anatomy of sentiment-driven fluctuations. Technical report, National Bureau of Economic Research.
- Alessi, L., Barigozzi, M., and Capasso, M. (2011). Non-fundamentalness in structural econometric models: A review. *International Statistical Review*, 79(1):16–47.
- Altig, D., Christiano, L. J., Eichenbaum, M., and Linde, J. (2011). Firm-specific capital, nominal rigidities and the business cycle. *Review of Economic dynamics*, 14(2):225–247.
- Angeletos, G.-M. and La’O, J. (2013). Sentiments. *Econometrica*, 81(2):739–779.
- Benati, L. and Surico, P. (2009). Var analysis and the great moderation. *American Economic Review*, 99(4):1636–52.
- Boivin, J. and Giannoni, M. P. (2006). Has monetary policy become more effective? *The Review of Economics and Statistics*, 88(3):445–462.

- Canova, F., Ferroni, F., and Matthes, C. (2020). Detecting and analyzing the effects of time-varying parameters in dsge models. *International Economic Review*, 61(1):105–125.
- Canova, F. and Pina, J. P. (2005). What var tell us about dsge models? In *New Trends in Macroeconomics*, pages 89–123. Springer.
- Carlstrom, C. T., Fuerst, T. S., and Paustian, M. (2009). Monetary policy shocks, choleski identification, and dnk models. *Journal of Monetary Economics*, 56(7):1014–1021.
- Castelnuovo, E. (2016). Modest macroeconomic effects of monetary policy shocks during the great moderation: An alternative interpretation. *Journal of Macroeconomics*, 47:300–314.
- Castelnuovo, E. and Fanelli, L. (2015). Monetary policy indeterminacy and identification failures in the us: Results from a robust test. *Journal of Applied Econometrics*, 30(6):924–947.
- Castelnuovo, E. and Surico, P. (2010). Monetary policy, inflation expectations and the price puzzle. *The Economic Journal*, 120(549):1262–1283.
- Cavicchioli, M. (2020). Invertibility and var representations of time-varying dynamic stochastic general equilibrium models. *Computational Economics*, 55(1):61–86.
- Chari, V. V., Kehoe, P. J., and McGrattan, E. R. (2008). Are structural vars with long-run restrictions useful in developing business cycle theory? *Journal of Monetary Economics*, 55(8):1337–1352.
- Christiano, L. J. (2002). Solving dynamic equilibrium models by a method of undetermined coefficients. *Computational Economics*, 20(1):21–55.
- Christiano, L. J., Eichenbaum, M., and Evans, C. L. (1999). Monetary policy shocks: What have we learned and to what end? *Handbook of macroeconomics*, 1:65–148.
- Christiano, L. J., Eichenbaum, M., and Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of political Economy*, 113(1):1–45.
- Fanelli, L. (2012). Determinacy, indeterminacy and dynamic misspecification in linear rational expectations models. *Journal of Econometrics*, 170(1):153–163.

- Fanelli, L. and Sorge, M. M. (2017). Indeterminate forecast accuracy under indeterminacy. *Journal of Macroeconomics*, 53:57–70.
- Fernández-Villaverde, J., Rubio-Ramírez, J. F., Sargent, T. J., and Watson, M. W. (2007). Abs (and ds) of understanding vars. *American Economic Review*, 97(3):1021–1026.
- Franchi, M. and Paruolo, P. (2015). Minimality of state space solutions of dsge models and existence conditions for their var representation. *Computational Economics*, 46(4):613–626.
- Franchi, M. and Vidotto, A. (2013). A check for finite order var representations of dsge models. *Economics Letters*, 120(1):100–103.
- Hespeler, F. and Sorge, M. M. (2019). Solving rational expectations models with informational subperiods: a comment. *Computational Economics*, 53(4):1649–1654.
- Kapetanios, G., Pagan, A., and Scott, A. (2007). Making a.. match: Combining theory and evidence in policy-oriented macroeconomic modeling. *Journal of Econometrics*, 136(2):565–594.
- Kasa, K. (2000). Forecasting the forecasts of others in the frequency domain. *Review of Economic Dynamics*, 3(4):726–756.
- Kasa, K., Walker, T. B., and Whiteman, C. H. (2014). Heterogeneous beliefs and tests of present value models. *Review of Economic Studies*, 81(3):1137–1163.
- Kilian, L. (2013). Structural vector autoregressions. In *Handbook of research methods and applications in empirical macroeconomics*. Edward Elgar Publishing.
- King, R. G. and Watson, M. W. (2002). System reduction and solution algorithms for singular linear difference systems under rational expectations. *Computational economics*, 20(1-2):57–86.
- Klein, P. (2000). Using the generalized schur form to solve a multivariate linear rational expectations model. *Journal of economic dynamics and control*, 24(10):1405–1423.
- Kormilitsina, A. (2013). Solving rational expectations models with informational subperiods: A perturbation approach. *Computational Economics*, 41(4):525–555.
- Kulish, M. and Pagan, A. (2017). Estimation and solution of models with expectations and structural changes. *Journal of Applied Econometrics*, 32(2):255–274.

- Leeper, E. M. and Leith, C. (2016). Understanding inflation as a joint monetary–fiscal phenomenon. In *Handbook of Macroeconomics*, volume 2, pages 2305–2415. Elsevier.
- Lubik, T., Matthes, C., and Mertens, E. (2020). Indeterminacy and imperfect information.
- Lubik, T. A. and Schorfheide, F. (2004). Testing for indeterminacy: An application to us monetary policy. *American Economic Review*, 94(1):190–217.
- Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.
- Morris, S. D. (2016). Varma representation of dsge models. *Economics Letters*, 138:30–33.
- Nimark, K. (2008). Dynamic pricing and imperfect common knowledge. *Journal of monetary Economics*, 55(2):365–382.
- Nimark, K. P. (2014). Man-bites-dog business cycles. *American Economic Review*, 104(8):2320–67.
- Poskitt, D. S. and Yao, W. (2017). Vector autoregressions and macroeconomic modeling: An error taxonomy. *Journal of business & economic statistics*, 35(3):407–419.
- Ravenna, F. (2007). Vector autoregressions and reduced form representations of dsge models. *Journal of monetary economics*, 54(7):2048–2064.
- Rotemberg, J. J. and Woodford, M. (1997). An optimization-based econometric framework for the evaluation of monetary policy. *NBER macroeconomics annual*, 12:297–346.
- Rozanov, J. A. and Rozanov, Y. A. (1967). *Stationary random processes*. Holden-Day.
- Schmitt-Grohé, S. and Uribe, M. (2004). Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of economic dynamics and control*, 28(4):755–775.
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society*, pages 1–48.
- Sims, C. A. (2002). Solving linear rational expectations models. *Computational economics*, 20(1-2):1.
- Soccorsi, S. (2016). Measuring nonfundamentality for structural vars. *Journal of Economic Dynamics and Control*, 71:86–101.

Sorge, M. M. (2020). Computing sunspot solutions to rational expectations models with timing restrictions. *The BE Journal of Macroeconomics*, 1(ahead-of-print).

Wolf, C. K. (2020). Svar (mis)identification and the real effects of monetary policy shocks. *American Economic Journal: Macroeconomics*, 12(4):1–32.

Appendix

A.1 Solution algorithm for general DSGE models

The first-order approximate solutions (and their minimal state space representation) to the general DSGE model under informational constraints can be obtained via the following algorithm — see Kormilitsina (2013) and Sorge (2020) for further details.

Step 1. Compute the steady state (\bar{y}, \bar{x}) of the unrestricted RE model $E_t f(y_{t+1}, y_t, x_{t+1}, x_t; \sigma) = 0$;

Step 2. Arrange variables in y and x in vectors $[y_u, y_r]$ and $[x_u, x_r]$. Sort the equilibrium conditions into vectors f^0 , f^1 and f^{x_r} , and arrange them into the partition $f = [f^0; f^1; f^{x_r}]$ accordingly;

Step 3. Obtain matrices g_x and h_x for the unrestricted RE model, and partition them as follows

$$g_x = \begin{pmatrix} g_{x_u} & g_{x_r} \\ \dot{j}_{x_u} & \dot{j}_{x_r} \end{pmatrix}, \quad h_x = \begin{pmatrix} h_{x_u} & h_{x_r} \\ 0 & P \end{pmatrix} \quad (51)$$

where g_{x_u} is $(n_{y_u} \times n_{x_u})$ -dimensional, g_{x_r} is $(n_{y_u} \times n_{x_r})$ -dimensional, \dot{j}_{x_u} is $(n_{y_r} \times n_{x_u})$ -dimensional, \dot{j}_{x_r} is $(n_{y_r} \times n_{x_r})$ -dimensional, h_{x_u} is $(n_{x_u} \times n_{x_u})$ -dimensional and h_{x_r} is $(n_{x_u} \times n_{x_r})$ -dimensional;

Step 4. Set

$$\hat{g}_{x_u} = g_{x_u}, \quad \hat{j}_{x_u} = \dot{j}_{x_u}, \quad \hat{h}_{x_u} = h_{x_u}, \quad (52)$$

$$\hat{g}_\sigma = 0, \quad \hat{j}_\sigma = 0, \quad \hat{h}_\sigma = 0; \quad (53)$$

Step 5. Compute the partial derivatives $f_{y'}^1, f_{x'_u}^1, f_{y_u}^1$, evaluate them at the steady state (\bar{x}, \bar{y}) and check invertibility of the square matrix

$$\nabla(f^1) = [f_{y'}^1 g_{x_u} + f_{x'_u}^1, f_{y_u}^1].$$

Then compute

$$\begin{pmatrix} \hat{h}_{x_r, -1} \\ \hat{g}_{x_r, -1} \end{pmatrix} = -\nabla(f^1)^{-1} f_{y_r}^1 j_{x_r} P \quad (54)$$

$$\begin{pmatrix} \hat{h}_{x_r} \\ \hat{g}_{x_r} \end{pmatrix} = \begin{pmatrix} h_{x_r} \\ g_{x_r} \end{pmatrix} + \nabla(f^1)^{-1} f_{y_r}^1 j_{x_r} \quad (55)$$

$$\hat{j}_{x_r, -1} = j_{x_r} P \quad (56)$$

Step 6. Derive the minimal state space representation under timing restrictions as follows

$$\begin{pmatrix} x'_u \\ x'_r \\ x_r \end{pmatrix} = \begin{pmatrix} \hat{h}_{x_u} & \hat{h}_{x_r} & \hat{h}_{x_r, -1} \\ 0_{n_{x_r} \times n_{x_u}} & P & 0_{n_{x_r} \times n_{x_r}} \\ 0_{n_{x_r} \times n_{x_u}} & I_{n_{x_r} \times n_{x_r}} & 0_{n_{x_r} \times n_{x_r}} \end{pmatrix} \begin{pmatrix} x_u \\ x_r \\ x_{r, -1} \end{pmatrix} + \sigma \begin{pmatrix} \epsilon_{x_u} \\ \epsilon_{x_r} \end{pmatrix} \quad (57)$$

$$\begin{pmatrix} y_u \\ y_r \end{pmatrix} = \begin{pmatrix} \hat{g}_{x_u} & \hat{g}_{x_r} & \hat{g}_{x_r, -1} \\ \hat{j}_{x_u} & 0_{n_{y_r} \times n_{x_r}} & \hat{j}_{x_r, -1} \end{pmatrix} \begin{pmatrix} x_u \\ x_r \\ x_{r, -1} \end{pmatrix} \quad (58)$$

A.2 Solving the bi-variate example from section (2)

Consider again the model

$$\mathcal{E} [y_{1,t} - \alpha y_{1,t+1} - y_{2,t} - x_{1,t} | \mathbb{I}_{1,t}] = 0, \quad \alpha > 0, \quad (59)$$

$$\mathcal{E} [y_{2,t} - \beta y_{2,t+1} - x_{1,t} - x_{2,t} | \mathbb{I}_{2,t}] = 0, \quad \beta > 0, \quad (60)$$

$$x_{1,t} = \rho_1 x_{1,t-1} + \varepsilon_{1,t}, \quad |\rho_1| < 1, \quad (61)$$

$$x_{2,t} = \rho_2 x_{2,t-1} + \varepsilon_{2,t}, \quad |\rho_2| < 1, \quad (62)$$

$$\varepsilon_{1,t} \sim NID(0, 1), \quad \varepsilon_{2,t} \sim NID(0, 1) \quad (63)$$

under the assumption $\mathbb{I}_{1,t} = \mathbb{V}_t(y_1^t, y_2^{t-1}) \vee \mathbb{V}_t(x_1^t, x_2^{t-1}) \subset \mathbb{I}_{2,t} = \mathbb{V}_t(y^t) \vee \mathbb{V}_t(x^t)$. This timing restriction involves the following assignment of the model's variables

$$y_u = [y_2], \quad y_r = [y_1] \quad (64)$$

$$x_u = [x_1], \quad [x_r] \quad (65)$$

and thereby the following informational partition

$$f^0 = (y_{1,t} - \alpha y_{1,t+1} - y_{2,t} - x_{1,t}) \quad (66)$$

$$f^1 = \begin{pmatrix} y_{2,t} - \beta y_{2,t+1} - x_{1,t} - x_{2,t} \\ x_{1,t+1} - \rho_1 x_{1,t} - \varepsilon_{1,t+1} \end{pmatrix} \quad (67)$$

$$f^{x_r} = (x_{2,t+1} - \rho_2 x_{1,t} - \varepsilon_{2,t+1}) \quad (68)$$

Recall that the determinate solution (8) to the unrestricted model can be partitioned as follows:

$$j_{x_u} = \frac{2 - \beta\rho_1}{(1 - \alpha\rho_1)(1 - \beta\rho_1)}, \quad j_{x_r} = \frac{1}{(1 - \alpha\rho_2)(1 - \beta\rho_2)}, \quad (69)$$

$$g_{x_u} = \frac{1}{1 - \beta\rho_1}, \quad g_{x_r} = \frac{1}{1 - \beta\rho_2}, \quad (70)$$

$$h_{x_u} = \rho_1, \quad h_{x_r} = 0, \quad P = \rho_2 \quad (71)$$

It is easily seen that $f_{y_r}^1 = [0, 0]'$ and thus $\hat{h}_{x_r, -1} = \hat{g}_{x_r, -1} = 0$, $\hat{h}_{x_r} = h_{x_r}$, $\hat{g}_{x_r} = g_{x_r}$, and finally $\hat{j}_{x_r, -1} = j_{x_r}\rho_2$. This leads to the equilibrium representation

$$y_{1,t} = \frac{2 - \beta\rho_1}{(1 - \alpha\rho_1)(1 - \beta\rho_1)} x_{1,t} + \frac{\rho_2}{(1 - \alpha\rho_2)(1 - \beta\rho_2)} x_{2,t-1}, \quad (72)$$

$$y_{2,t} = \frac{1}{1 - \beta\rho_1} x_{1,t} + \frac{1}{1 - \beta\rho_2} x_{2,t} \quad (73)$$

which is in the same form as (9).

A.3 VARs and Cholesky identification

Let an n_y -dimensional vector of observable variables be represented as a canonical VAR of order k

$$y_t = A_1 y_{t-1} + \dots + A_k y_{t-k} + u_t, \quad E[u_t] = 0, \quad E[u_t u_t'] = V \quad (74)$$

where k is a non-negative integer (capturing the number of lags) and the innovations u_t are assumed to be uncorrelated with all variables dated $t - 1$ and earlier.

To uncover the dynamic response functions of y_t to fundamental (structural) economic shocks ϵ_t , researchers usually assume existence of a linear relationship between the latter and the VAR innovations, i.e.

$$B_0 u_t = \epsilon_t, \quad (75)$$

with B_0 being a square, full-rank matrix. The VAR (74) thus admits the structural representation

$$B_0 y_t = B_1 y_{t-1} + \cdots + B_k y_{t-k} + \epsilon_t, \quad B_i := B_0 \cdot A_i, \quad i = 1, \dots, q \quad (76)$$

which makes clear that the impulse response of any component of the y_t vector to a transitory, unit shock in some component of ϵ_t is a function of entries of matrices B_i , $i = 0, \dots, k$. Absent restrictions on B_0 , the identification of structural shocks ϵ_t requires additional restrictions, for data will only provide information about the response of y_t to innovations u_t . Since $E[\epsilon_t \epsilon_t'] = I_{n_y \times n_y} = B_0 V B_0'$, and upon recognizing that V can be consistently estimated from (74) and thus treated as known, the B_0 matrix will be completely identified by imposing $n_y(n_y + 1)/2$ identifying constraints.

According to the Cholesky scheme, identification is achieved by orthogonalizing the innovations u_t via a Cholesky decomposition of its variance covariance matrix of the form $PP' = V$ (where P is a conformable, lower triangular matrix with real and positive diagonal entries) and then simply imposing $B_0^{-1} = P$. This reveals that the Cholesky scheme amounts to restricting the impact reactions of some variables to structural shocks, i.e. by setting selected elements of B_0^{-1} to zero for a given ordering in y_t .



Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Strada Maggiore 45
40125 Bologna - Italy
Tel. +39 051 2092604
Fax +39 051 2092664
<http://www.dse.unibo.it>