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

Published Version:
Angelini G., Sorge M.M. (2021). Under the same (Chole)sky: DNK models, timing restrictions and recursive identification of monetary policy shocks. JOURNAL OF ECONOMIC DYNAMICS & CONTROL, 133, 104265-104265 [10.1016/j.jedc.2021.104265].

Availability:
This version is available at: https://hdl.handle.net/11585/843192 since: 2021-12-27

Published:

DOI: http://doi.org/10.1016/j.jedc.2021.104265

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

Giovanni Angelini† Marco M. Sorge‡

April 20, 2022

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. This paper characterizes the effects on impulse propagation of informational constraints embodying classical Cholesky-type timing restrictions in otherwise standard DSGE models. We formally show that timing restrictions can produce non-trivial moving average components of rational expectations solutions, or even serve as an independent source of model-based nonfundamentalness, thereby hampering impulse response analysis via VAR procedures. We then derive population conditions for existence of VAR representations of DSGE economies exhibiting timing restrictions, and numerically explore their bearing on shock identification in a range of monetary models of the business cycle. Our analysis reveals that dynamic New Keynesian models admit invertible equilibrium representations as well as 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 retrieving the true aggregate effects of monetary policy shocks in a Cholesky world.

Keywords: DSGE models, Timing restrictions, Vector autoregression, Cholesky identification
JEL Classification: C3, E3

---

*We wish to thank Efrem Castelnuovo, Chetan Dave, Luca Fanelli, Alessia Paccagnini, Matthias O. Paustian, Stefano Soccorsi, Justyna Wróblewska and seminar/workshop participants (CSEF Lunch Talk, 7^{th} RCEA Time Series Conference, 4^{th} Workshop on Macroeconomic Research) for helpful discussions on earlier versions of the manuscript. We also owe a debt of gratitude to two anonymous referees for extremely valuable comments and suggestions. Any remaining errors are our sole responsibility.

†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

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 policy rate placed between slow and fast moving variables, e.g. Kilian (2013). Albeit empirically appealing, the Cholesky approach calls for a conceivable structural interpretation of the enforced recursive ordering.[1]

The lack of conformity between the conventional timing in structural macroeconomic frameworks and the identifying assumptions in Cholesky-VARs has also spurred interest in the development of dynamic stochastic general equilibrium (DSGE) models exhibiting some degree of recursiveness, e.g. Rotemberg and Woodford (1997), Christiano et al. (2005), Boivin and Giannoni (2006), Altig et al. (2011). While resorting to peculiar (and often overly restrictive) assumptions on the timing of decisions, these studies do not engage in a thorough investigation of the DSGE-VAR mapping in the presence of Cholesky-type restrictions.

The present paper fully characterizes conditions for existence of causal VAR representations of general, multivariate DSGE models featuring informational constraints that embody classical Cholesky-timing restrictions; and explores the relevance of these conditions for empirical VAR-based exercises aimed at identifying the monetary transmission mechanism or rather at validating/estimating DSGE models via the use of impulse-response analysis. Formally, we show that information-based timing restrictions enlarge a model's equilibrium state space and modify rational expectations cross-equation restrictions (CERs), opening room to the emergence of (i) nonfundamental VARMA representations for any set of observables, or (ii) invertible non-trivial MA equilibrium components vis-à-vis their counterparts free of timing restrictions. As a result, even when impulse responses are rightly constrained by the Cholesky scheme to be an exact match to the theoretical ones on impact, a finite order Cholesy-VAR may still prove an inaccurate approximation of the true VARMA structure that shapes the endogenous adjustment paths to monetary policy shocks.

It is well-known that identifying a given DSGE model's structural shocks via VAR analysis require that the observed variables contain sufficient information to recover the

[1] The 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).

unobserved state variables which are assumed to produce the observables. In particular, if the model's solution does not admit an invertible (or fundamental) representation for the observables, then there exists no (linear) invertible mapping from the VAR innovations to the structural shocks, see e.g. Alessi et al. (2011).

Our first contribution is to uncover a novel invertibility issue in theoretical macroeconomic models that conform with a Cholesky-type structure. Differently from dynamic 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 generically 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. We isolate conditions for existence of fundamental equilibrium representations of restricted DSGE models, that solely rely on the imposed information structure and the reduced-form coefficients of the model's solution under conventional (unrestricted) timing. In the same spirit of Fernández-Villaverde et al. (2007), the procedures to check for fundamentalness and to characterize the finite-order VAR equilibrium representation of restricted DSGE models are cast in a straightforward algorithmic form, that can be used very broadly in structural macroeconomic modeling, well beyond the applications discussed next.[2]

Even when warranting invertibility, timing restrictions can enforce VARMA equilibrium representations that exhibit slowly converging VAR polynomial matrices at fairly longer horizons. In principle, this might prevent low-order VARs, of the type required by actual data availability, from being informative about the dynamic effects of monetary shocks, even though the identification scheme correctly reflects the structure of the model that generates the observables, e.g. Ravenna (2007) and Poskitt and Yao (2017). To investigate this issue, we set up a series of controlled Monte Carlo experiments on the assumption that the true data generating process (DGP) belongs to the class of restricted DSGE economies. Specifically, we rely on several specifications of two fully-fledged models of monetary policy transmission in the New Keynesian tradition. The first one is the sticky-price framework with consumption habits and inflation inertia advanced by Boivin and Giannoni (2006) for their quantitative exploration of the effectiveness of monetary policy in the U.S. post-WWII macroeconomic history. The second is the three-equation monetary business cycle model with lagged expec-

[2] We 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.

tations employed by Guerron-Quintana et al. (2017) to exemplify the properties of impulse-response matching estimators for DSGE models. We leverage the flexibility of Kormilitsina (2013)'s perturbation approach to solving restricted DSGE models in order to bring the above mentioned frameworks in closer conformity with the Cholesky-VAR recursive structure. We then submit the properly restricted models to our formal test for VAR representability, and explore what the practical consequences of failing this test are for the impulse response functions of interest.

We show that both dynamic New Keynesian (DNK) models admit a causal (possibly infinite order) VAR representation under empirically tenable 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(∞) further reveals that they rapidly decay towards zero, mitigating concerns about lag truncation bias arising from the adoption of low-order VAR systems. We then scrutinize the ability of Cholesky-VARs to uncover the true aggregate effects of monetary shocks through the lens of our model laboratories; and also assess the relative performance of the Cholesky scheme vis-à-vis the agnostic sign restrictions procedure, popularized by e.g. Canova and De Nicolo (2002) and Uhlig (2005). Our simulation results convincingly suggest that Cholesky-VARs, *per se* and as opposed to the competing identification scheme, are likely to provide tight and reliable inference about the monetary transmission mechanism in restricted DNK structures, no matter how long-lasting the dynamic effects of timing restrictions are. Sign restrictions, by contrast, may fail to correctly unveil the true size of the monetary impulse responses for they prove sensitive to the relative volatility scale of structural disturbances; a finding in line with e.g. Paustian (2007) and Castelnuovo (2016). Finally, an application to the estimation of restricted DSGE models via IRF matching techniques is also proposed.

A last remark about the scope of our analysis is in order. While we focus on the DSGE-VAR mapping in the presence of structural Cholesky-type restrictions that arise from ever-binding informational constraints, a large strand of literature has rather explored alternative approaches to shock identification that do not rely on the ordering of the variables in the VAR and thereby comply with the standard (non-recursive) timing protocol of DSGE models, such as the use of external instruments (e.g. Gertler and Karadi (2015), Stock and Watson (2018), Angelini and Fanelli (2019)), the heteroskedasticity in the data (e.g. Bacchiocchi and Fanelli (2015), Kilian and Lütkepohl (2017)), the independence and non-Gaussianity of the shocks (e.g. Lanne et al. (2017)), or direct coefficient restrictions on monetary policy rules (e.g. Arias et al. (2019)). We of course acknowledge the relevance of these research avenues, whether or not associated with the empirical validation of shocks' transmission mechanisms of DSGE frameworks. The specific questions we address and the answers we provide are not to be read as an endorsement of classic exclusion restrictions against competing alternatives for identification purposes; rather, they are meant to warn applied researchers against the potential emergence of econometric issues inherent in the reduction of restricted DSGE models into VAR form.

The paper proceeds as follows: Section (2) presents a simple example that illustrates the potentially harmful consequences of information-based timing restrictions on the VAR-based analysis of DSGE models. Section (3) formally discusses how to construct first-order approximate solutions to general DSGE economies featuring timing restrictions. Section (4) provides easy-to-check conditions under which restricted DSGE models admit a fundamental (of possibly finite-order) equilibrium representation for the observables. Section (5) lays out the DNK model laboratories, that are next used (Section (6)) to assess the ability of classical short-run exclusion restrictions to unveil the model-implied monetary transmission mechanism, and to document the relative performance of Cholesky-VARs versus a sign restrictions-type of approach. Section (7) offers concluding remarks.

1.1 Literature review

We are of course not the first to formally introduce timing restrictions in otherwise standard DSGE environments. Previous work in the field – e.g. Rotemberg and Woodford (1997), Christiano et al. (2005), Boivin and Giannoni (2006), Altig et al. (2011) – formalizes the recursiveness assumption by positing that slow moving non-policy variables (such as consumption, wages and prices) are fixed at least one period in advance, and expectations on the future evolution of such variables are conditioned on lagged information sets that fail to include current realizations of the whole set of the model's variables (endogenous and exogenous).[3]

While guaranteeing that the model's responses to a monetary innovation are zero on impact, this is actually *not* an implication of the Cholesky-timing assumptions for the latter in principle allow, in structural settings, non-monetary shocks to hit these variables on impact as well as to influence the formation of expectations (and of the endogenous forecast revisions, e.g. Sims (2002)) about the future unfolding of the

[3] Keating (1996) derives a set of general recursiveness conditions on a structural VAR model under which the Cholesky decomposition succeeds in identifying structural shock responses.

model's variables.

A relatively more conservative approach, capable of perfectly aligning a given DSGE model's structure with the Cholesky-VAR, has been put forward in Kormilitsina (2013) and generalized in Sorge (2020). It acknowledges the fact that it is not the date at which expectations are formed that matters, but rather the date and the structure of the information set upon which expectations are framed. Accordingly, this approach formalizes timing restrictions in the form of fictitious *informational sub-periods* characterized by increasing sequences of heterogeneous (across rational decision-makers) information sets and of the ensuing framing of expectations and timing of decisions. Most importantly, being rooted in the theory of perturbation of non-linear systems, Kormilitsina (2013)'s approach works *directly* with the non-linear equilibrium conditions of a given DSGE model by specifying, for each variable and/or equation, informational sub-periods (sequences of observables and conditioning information sets) and associated expectation operators; approximated (up to second-order) optimal decision rules can then be derived on the basis of the imposed information structure.

Our analysis adds to a rapidly growing strand of literature that has explored the DSGE-VAR mapping in controlled model laboratories. 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 DNK models to the impulse predictions of Cholesky-type VARs when it comes to estimating macroeconomic reactions to a monetary policy shock. They document the extent to which the Cholesky scheme is likely to produce spurious price and output puzzles under small perturbations of the model's parameters, that leave the theoretical impulse responses almost unchanged.[4]

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

[4]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).

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.[5]

Relative to the mentioned studies, we offer additional insights by devising easy-to-check (necessary and sufficient) conditions for DSGE models with timing restrictions to admit a VAR representation, and implementing them in the context of DNK models. In the presence of nonfundamentalness, a direct comparison of the DSGE's and VAR's impulse response may be misleading; even non-existence of finite-order VAR representations may complicate statistical inference (Alessi et al., 2011). Our simulation results indicate that, irrespective of whether timing restrictions deeply affect the sign, the magnitude and the persistence of dynamic adjustment paths to a monetary policy surprise, low-order Cholesky-VARs perform remarkably well in unraveling structural content in a truly Cholesky world.[6]

Our paper also speaks to the rapidly increasing literature on imperfect information in dynamic RE models. 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 small perturbations to full information generate stable RE equilibria that prove least-squares learnable (Rondina and Walker (2016)) as well as conditions under which imperfect information qualifies as a mechanism that supports self-fulfilling sunspot beliefs in economies that would rather exhibit a determinate equilibrium in the presence of perfect and symmetrically 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 specific 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

[5] See 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.

[6] This is in line with the experiments run in Altig et al. (2011), who document the good performance of Cholesky-VARs in identifying monetary impulse responses in a three-shock business cycle model featuring a lagged transmission mechanism.

equilibrium, so will its restricted analogue. 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.[7]

2 An illustrative example

Answering the question of whether the structural shocks of a given multivariarate 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

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

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

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

$$x_{2,t} = \rho_1 x_{2,t-1} + \varepsilon_{2,t}, \quad |\rho_2| < 1, \tag{4}$$

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

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, $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t})'$ are structural innovations, and $\mathcal{E}\left[\cdot|\mathbb{I}_{i,t}\right]$ 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. $\mathbb{I}_t = \mathbb{V}_t(y^t) \vee \mathbb{V}_t(x^t)$ where $y^t \equiv \{y_t, y_{t-1}, y_{t-2}, \ldots\}$ and $x^t \equiv \{x_t, x_{t-1}, x_{t-2}, \ldots\}$.

[7] It should be stressed that DSGE models with timing restrictions 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; and Canova et al. (2020) for an exploration of inferential issues related to time-varying structural macroeconomic models.

Letting $E_t[\cdot] = \mathcal{E}[\cdot|\mathbb{I}_t]$, the RE system (1)-(5) can be cast in the conventional form

$$y_t = AE_t\left[y_{t+1}\right] + Bx_t, \tag{6}$$

$$x_t = Cx_{t-1} + \varepsilon_t \tag{7}$$

where

$$A = \begin{pmatrix} \alpha & \beta \\ 0 & \beta \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{pmatrix} \tag{8}$$

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

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

which generically admits a finite-order VAR representation.[8]

Suppose 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 $\mathbb{I}_{1,t} = \mathbb{V}_t(y_1^t, y_2^{t-1}) \vee \mathbb{V}_t(x_1^t, x_2^{t-1}) \subset \mathbb{I}_{2,t} = \mathbb{V}_t(y^t) \vee \mathbb{V}_t(x^t)$. Provided $\alpha, \beta \in (0,1)$, the restricted model will exhibit the same determinacy properties as its unrestricted counterpart – see 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, $\mathcal{E}\left[y_{2,t}|\mathbb{I}_{1,t}\right] \neq y_{2,t}$.

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

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \frac{2-\beta\rho_1}{(1-\alpha\rho_1)(1-\beta\rho_1)} & 0 & \frac{\rho_2}{(1-\alpha\rho_2)(1-\beta\rho_2)} \\ \frac{1}{1-\beta\rho_1} & \frac{1}{1-\beta\rho_2} & 0 \end{pmatrix} \begin{pmatrix} x_{1,t} \\ x_{2,t} \\ x_{2,t-1} \end{pmatrix} \tag{10}$$

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

which admits the factorized MA form

$$
\begin{aligned}
y_t &= M_1(L) \cdot M_2(L)\varepsilon_t, &(11)\\
M_1(L) &= \begin{pmatrix} \frac{2-\beta\rho_1}{(1-\alpha\rho_1)(1-\beta\rho_1)} & \frac{\rho_2}{(1-\alpha\rho_2)(1-\beta\rho_2)} \cdot L \\ \frac{1}{1-\beta\rho_1} & \frac{1}{1-\beta\rho_2} \end{pmatrix}, &(12)\\
M_2(L) &= \begin{pmatrix} (1-\rho_1 L)^{-1} & 0 \\ 0 & (1-\rho_2 L)^{-1} \end{pmatrix} &(13)
\end{aligned}
$$

Existence of a causal VAR representation of the RE model obtains 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.

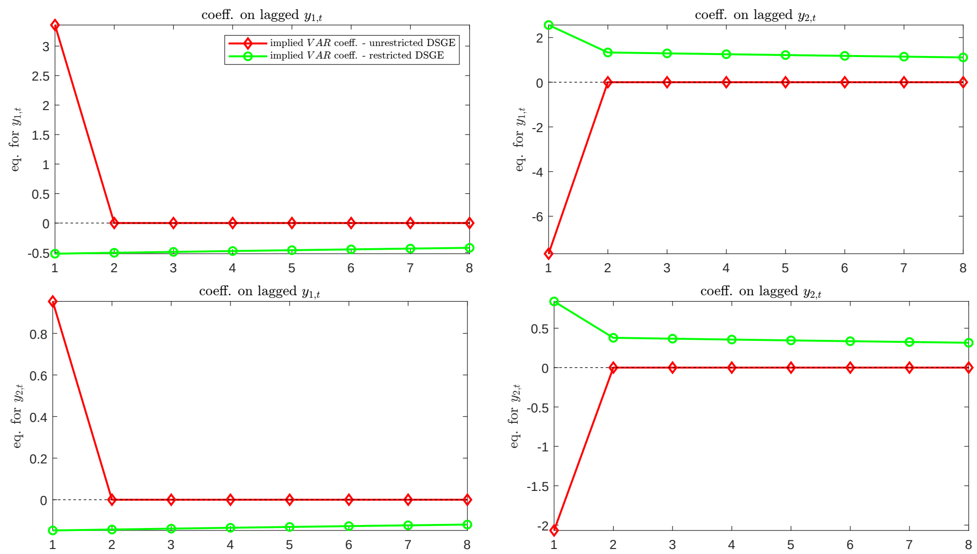

Figure 1: *Evolution of coefficients of the theoretical VAR representation of the RE model* (1)-(5) *with and without timing restrictions.*

Nonfundamentalness is also an issue here. In fact, for generic parametric configurations, 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.

To shed further light on the relevance of this novel source of nonfundamentalness bias, we next compute the error measure put forward in Soccorsi (2016) – called d_∞ – for two parameterizations of the simple RE model. This measure quantifies the distance between the nonfundamental representation of the data providing the structural shocks and its unique Wold representation, and therefore represents the limit lower bound of the error one incurs into when approximating the true (nonfundamental) VARMA structure with a finite-order VAR. The nonfundamentalness bias is accordingly expressed in terms of percentage error relative to the zero threshold that the fundamental MA representation achieves by construction. As reported in the Table (1), when the model's parameterization supports a fundamental equilibrium representation under timing restrictions, the bias is null. For a different parameterization that enforces nonfundamentalness, such a bias is by contrast quite substantial (i.e. 278% percent error); in this case, impulse-response analysis based on structural VAR models of whatever lag length, albeit correctly identified, would be invalid.[9]

Nonfundamentalness				
d_∞	α	β	ρ_1	ρ_2
0	0.80	0.69	0.45	0.84
2.78	0.80	0.80	0.20	0.95

Table 1: *Soccorsi (2016)'s measure of nonfundamentalness bias. Here* $d_\infty = \frac{||\Sigma_{u-v}||}{||\Sigma_u||}$*, where* $||\cdot||$ *denotes the Euclidean norm for square matrices,* Σ_u *is the covariance matrix of the vector of the true residuals* u_t *of the nonfundamental equilibrium representation, and* Σ_{u-v} *denotes the covariance matrix of the vector difference between these residuals and the Wold innovations* v_t *for the observables.*

[9]Forni et al. (2019) put forward a measure of informational deficiency of small-scale VARs for the identification of specific economic shock, or a subset of shocks, of interest. While close in spirit to the distance metric proposed in Soccorsi (2016), unlike the latter it provides a shock-specific indication for the informational deficiency of a given set of observables. See also Canova and Hamidi Sahneh (2018) and Beaudry et al. (2019) for alternative testing procedures for the relevance of nonfundamentalness in structural VARs.

3 Environment: DSGE models with timing restrictions

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

$$E_t f\left(y_{t+1}, y_t, x_{t+1}, x_t; \sigma\right) = 0 \tag{14}$$

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 n_x^1 endogenous predetermined variables as well as n_x^2 exogenous states ($n_x^1 + n_x^2 = n_x$). 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 (14) are in the form

$$y = g(x, \sigma), \quad x' = h(x, \sigma) + \sigma\epsilon' \tag{15}$$

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

$$y = g_x x, \quad x' = h_x x + \sigma\epsilon' \tag{16}$$

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 (14) when $\sigma = 0$.

3.2 Restricted model

In the presence of information-based timing restrictions, the general form of the multivariate RE model is

$$\mathcal{E}_t \left[f \left(y_{t+1}, y_t, x_{t+1}, x_t; \sigma \right) \right] = 0 \tag{17}$$

where $\mathcal{E}_t$ denotes the collection of (conditional) expectation operators accounting for informational constraints embedded in the conditioning information sets, i.e.

$$\mathcal{E}_t \left[f (y_{t+1}, y_t, x_{t+1}, x_t; \sigma) \right] = \begin{pmatrix} \mathcal{E} \left[f_1^{(y,x)}(y_{t+1}, y_t, x_{t+1}, x_t; \sigma) \,\middle|\, \mathbb{I}_{1,t} \right] \\ \mathcal{E} \left[f_2^{(y,x)}(y_{t+1}, y_t, x_{t+1}, x_t; \sigma) \,\middle|\, \mathbb{I}_{2,t} \right] \\ \vdots \\ \mathcal{E} \left[f_{n_y+n_x^1}^{(y,x)}(y_{t+1}, y_t, x_{t+1}, x_t; \sigma) \,\middle|\, \mathbb{I}_{n_y+n_x^1,t} \right] \\ f_1^{(x)}(x_{t+1}^2, x_t^2; \sigma) \\ f_2^{(x)}(x_{t+1}^2, x_t^2; \sigma) \\ \vdots \\ f_{n_x^2}^{(x)} \left(x_{t+1}^2, x_t^2; \sigma \right) \end{pmatrix}$$

where $f_k^{(y,x)}$ $(k \leq n_y + n_x^1)$ is the model's equation used to pin down the k-th endogenous variable (y, x^1), conditional on the equilibrium values for the other endogenous variables and the relevant states, for which model-consistent expectations (optimal projections) at date t are determined on the basis of the restricted (and in principle different across these equations) information set $\mathbb{I}_{k,t}$, $k \leq n_y$; and $f_j^{(x)}$ $(j \leq n_x^2)$, is the possibly nonlinear equation that governs the dynamics of j-th exogenous state variable x_j.

We next consider the most basic case where any given time period is split into two informational sub-periods, according to the timing of the model's variables. Formally, the control and state vectors are accordingly partitioned as

$$y = [y_u;\ y_r], \quad x = [x_u;\ x_r] \tag{18}$$

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 (14) is partitioned as follows

$$f = \left[f^0;\ f^1;\ f^{x_r}\right] \tag{19}$$

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

$$x_r' = Px_r + \sigma\epsilon_{x_r}', \quad \epsilon_{x_r} \sim i.i.d.N(0, V_{\epsilon_{x_r}}) \tag{20}$$

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 recursive solution to the restricted RE model is in the general form

$$y_u = \hat{g}(x_u, x_r, x_{r,-1}, \sigma), \quad y_r = \hat{j}(x_u, x_{r,-1}, \sigma), \quad x_u' = \hat{h}(x_u, x_r, x_{r,-1}, \sigma) + \sigma\epsilon_{x_r}' \tag{21}$$

Notice that 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 (21) is

$$\begin{aligned} y_u &= \hat{g}_{x_u} x_u + \hat{g}_{x_r} x_r + \hat{g}_{x_{r,1}} x_{r,-1}, \\ y_r &= \hat{j}_{x_u} x_u + \hat{j}_{x_{r,-1}} x_{r,-1}, \\ x'_u &= \hat{h}_{x_u} x_u + \hat{h}_{x_r} x_r + \hat{h}_{x_{r,-1}} x_{r,-1} + \sigma \epsilon'_{x_u} \end{aligned} \tag{22}$$

or in more compact form

$$y = \hat{g}_x \begin{pmatrix} x_u \\ x_r \\ x_{r,-1} \end{pmatrix}, \quad x' = \hat{h}_x \begin{pmatrix} x_u \\ x_r \\ x_{r,-1} \end{pmatrix} + \sigma \epsilon' \tag{23}$$

where

$$\hat{g}_x = \begin{pmatrix} \hat{g}_{x_u} & \hat{g}_{x_r} & \hat{g}_{x_{r,-1}} \\ \hat{j}_{x_u} & 0_{n_{y_r} \times n_{x_r}} & \hat{j}_{x_{r,-1}} \end{pmatrix}, \quad \hat{h}_x = \begin{pmatrix} \hat{h}_{x_u} & \hat{h}_{x_r} & \hat{h}_{x_{r,-1}} \\ 0_{n_{x_r} \times n_{x_r}} & P & 0_{n_{x_r} \times n_{x_r}} \end{pmatrix} \tag{24}$$

Provided the rank condition characterized in Sorge (2020) is fulfilled, the solution to the restricted model can be readily constructed via uniquely defined linear transformations of (16), however computed (e.g. Klein (2000); Christiano (2002); King and Watson (2002); Sims (2002)).
Upon partitioning the equilibrium coefficient matrices (g_x, h_x) in (16) as follows

$$g_x = \begin{pmatrix} g_{x_u} & g_{x_r} \\ j_{x_u} & j_{x_r} \end{pmatrix}, \quad h_x = \begin{pmatrix} h_{x_u} & h_{x_r} \\ 0 & P \end{pmatrix} \tag{25}$$

we can easily map the full information coefficient matrices into those appearing in (23), i.e.

$$\hat{g}_x = \begin{pmatrix} g_{x_u} & g_{x_r} + \left(\nabla(f^1)^{-1} f^1_{y_r} j_{x_r}\right)_{n_{y_u}} & -\left(\nabla(f^1)^{-1} f^1_{y_r} j_{x_r} P\right)_{n_{y_u}} \\ j_{x_u} & 0_{n_{y_r} \times n_{x_r}} & j_{x_r} P \end{pmatrix}, \tag{26}$$

$$\hat{h}_x = \begin{pmatrix} h_{x_u} & h_{x_r} + \left[\nabla(f^1)^{-1} f^1_{y_r} j_{x_r}\right]_{n_{x_u}} & \left[-\nabla(f^1)^{-1} f^1_{y_r} j_{x_r} P\right]_{n_{x_u}} \\ 0 & P & 0 \end{pmatrix} \tag{27}$$

where $\nabla(f^1)$ denotes the Jacobian of the sub-system f^1 with respect to the vector $[x'_u, y_u]$, $f^1_{y_r}$ 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.[10]

4 Timing restrictions and VAR representation 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 functions. Since all the non-zero eigenvalues of the companion matrix

$$h_x^\dagger = \begin{pmatrix} h_{x_u} & h_{x_r} + \left[\nabla(f^1)^{-1} f^1_{y_r} j_{x_r}\right]_{n_{x_u}} & \left[-\nabla(f^1)^{-1} f^1_{y_r} j_{x_r} P\right]_{n_{x_u}} \\ 0 & P & 0 \\ 0 & I & 0 \end{pmatrix} \tag{28}$$

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

$$y_t = C(L)x_t = C(L)\left(I - A(L)\right)^{-1} \sigma\epsilon_t = B(L)\epsilon_t \tag{29}$$

where L stands for the conventional lag operator and

$$C(L) = \left(\hat{g}_{x_u},\, \hat{g}_{x_r} + \hat{g}_{x_{r,-1}} L\right), \quad A(L) = \left(\hat{h}_{x_u},\, \hat{h}_{x_r} + \hat{h}_{x_{r,-1}} L\right) \tag{30}$$

Recall that the IRFs of the elements of y_t to a unit shock in one of the elements of ϵ_t occuring at some time $t = s$ are given by the sequences of corresponding MA coefficients in (29) 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 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 (29) 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

[10]To facilitate reading, we relegate to the Appendix the computational details of the solution algorithm developed in Kormilitsina (2013), and apply it to the illustrative example discussed in Section (2).

systems, i.e. when $n_y = n_x$, e.g. (Alessi et al., 2011).

The reduced form of the general model (14) 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 follows[11]

Condition 1 (**Fundamentalness**). *Let $n_y = n_x$. Then, for any given informational partition (18)-(19), the equilibrium representation under timing restrictions is fundamental if and only if*

$$Det\begin{pmatrix} g_{x_u} & g_{x_r} + \left[\nabla(f^1)^{-1} f^1_{y_r} j_{x_r}\right]_{n_{y_u}} - \left[\nabla(f^1)^{-1} f^1_{y_r} j_{x_r} P\right]_{n_{y_u}} \cdot z \\ j_{x_u} & j_{x_r} P \cdot z \end{pmatrix} = 0 \implies |z| \geq 1 \tag{31}$$

Even when fundamentalness is warranted by the model's parameterization, the equilibrium dynamics (29) 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 Soccorsi (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

$$y_t = C(L)A(L)x_{t-1} + C(L)\sigma\epsilon_t, \tag{32}$$

$$x_t = A(L)x_{t-1} + \sigma\epsilon_t \tag{33}$$

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*

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

matrices

$$I^{\circ} = \begin{pmatrix} I_{n_x \times n_x} \\ 0_{n_{x_r} \times n_x} \end{pmatrix}, \tag{34}$$

$$g_x^{\dagger} = \begin{pmatrix} g_{x_u} & g_{x_r} + \left[\nabla(f^1)^{-1} f^1_{y_r} j_{x_r}\right]_{n_{y_u}} \\ j_{x_u} & 0_{n_{y_r} \times n_{x_r}} \end{pmatrix}, \tag{35}$$

$$g_x^{\circ} = \begin{pmatrix} g_{x_u} & g_{x_r} + \left[\nabla(f^1)^{-1} f^1_{y_r} j_{x_r}\right]_{n_{y_u}} & -\left[\nabla(f^1)^{-1} f^1_{y_r} j_{x_r} P\right]_{n_{y_u}} \\ j_{x_u} & 0_{n_{y_r} \times n_{x_r}} & j_{x_r} P \end{pmatrix} \cdot h_x^{\dagger} \tag{36}$$

Then, provided $g_x^{\dagger}$ is non-singular, for any given informational partition (18)-(19), a finite order VAR equilibrium representation under timing restrictions exists if and only if the $(n_x + n_{x_r})$-dimensional square matrix

$$F := h_x^{\dagger} - I^{\circ} \cdot \left(g_x^{\dagger}\right)^{-1} \cdot g_x^{\circ} \tag{37}$$

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 (14), and can thus be checked whatever the solution algorithm employed, once the informational partition (19) has been devised.[12]

5 Model laboratories

To run our Monte Carlo experiments, we borrow the small-scale DNK models of monetary policy transmission from Boivin and Giannoni (2006) and Guerron-Quintana et al. (2017). As mentioned, both these frameworks feature overly restrictive assumptions about the timing of decisions and expectations formation, relative to those enforced by the Cholesky scheme. As a consequence, and for our purposes, we shall first show how to rework the timing of the two models via Kormilitsina (2013)'s methodology.

[12] Notice that, given that the state space representation (32) 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).

5.1 Boivin and Giannoni (2006)'s model

Building upon Rotemberg and Woodford (1997), Boivin and Giannoni (2006) develop a stylized structural macroeconomic in which private agents (households and producers) must optimize their representative objective function on the basis of information regarding the state of the economy *two periods* (for consumption decisions) or *one period* earlier (for pricing decisions). This prevents consumption (and, in equilibrium, output) and the inflation rate from reacting to current exogenous demand and supply disturbances, while also forcing forward-looking expectations on these variables not to adjust in response to changes in any of the exogenous variables, as well as in the policy instrument (the short-term nominal interest rate). What is more, the intertemporal investment-savings equation features lagged expectations (formed at date $t-2$) of the one-step ahead (log-linearized) output and of the marginal utility of additional income, which by construction do not reflect knowledge of current inflation, for the latter is determined at time t on the basis of information available at time $t-1$.

Formally, Boivin and Giannoni (2006)'s model comprises the following log-linearized equilibrium conditions[13].

- **Dynamic IS equation** (here y_t stands for aggregate output, $\hat{\lambda}_t$ is the log-linearized Lagrangian multiplier on the representative household's intertemporal budget constraint, i.e. the household's marginal utility of nominal income at time t, π_t is the inflation rate and i_t is the short-term nominal interest rate)

$$y_t = \eta y_{t-1} + \beta\eta E_{t-2} y_{t+1} - \psi E_{t-2}\hat{\lambda}_t + g_t, \tag{38}$$

$$\hat{\lambda}_t = E_t\left[\hat{\lambda}_{t+1} + i_t - \pi_{t+1}\right] \tag{39}$$

- **Aggregate supply equation** (here $\hat{s}_t$ is the log-linearized cross-firm average real marginal cost)

$$\pi_t = \gamma\pi_{t-1} + \kappa E_{t-1}\hat{s}_t + \beta E_{t-1}\left[\pi_{t+1} - \gamma\pi_t\right], \tag{40}$$

$$\hat{s}_t = \omega y_t - \hat{\lambda}_t - q_t, \quad \omega > 0 \tag{41}$$

- **Monetary policy rule**

$$i_t = \rho_1 i_{t-1} + \rho_2 i_{t-2} + \phi_\pi \pi_t + \phi_y y_t + \varepsilon_t, \tag{42}$$

[13] All variables are expressed as percentage deviations from the model's non-stochastic steady state, in equilibrium output is equal to consumption plus government expenditure.

where the exogenous demand shock g_t and supply shock q_t are functions of first- and second-order derivatives of the utility function evaluated at steady-state as well as past expectations of the preference shock that enters the functional specification of the consumption utility and the output production disutility.

Consistent with the Cholesky-timing assumptions about monetary policy surprises, we rather submit the NK structure to information-based timing restrictions under which (i) the monetary policy shocks are orthogonal to the non-policy variables (y_t, π_t) and (ii) these non-policy variables are thus predetermined with respect to the nominal interest rate, while sensitive to non-monetary shocks that occur in the same time period. Letting $\mathcal{E}_t = E(\cdot|\mathbb{I}_t)$ and

$$\mathbb{I}_t = \{y_{t-j}, \pi_{t-j}, i_{t-1-j}, g_{t-j}, q_{t-j}, \varepsilon_{t-1-j}; j \geq 0\}$$

we obtain

- **Dynamic IS equation**

$$y_t = \eta y_{t-1} + \mathcal{E}_t \left[\beta \eta y_{t+1} - \psi \sum_{j=t}^{\infty} (i_j - \pi_{j+1}) \right] + \hat{g}_t, \tag{43}$$

- **Aggregate supply equation**

$$\pi_t = \frac{\gamma}{1+\beta\gamma} \pi_{t-1} + \frac{\beta}{1+\beta\gamma} \mathcal{E}_t [\pi_{t+1}] + \frac{\kappa}{1+\beta\gamma} \left[\omega y_t - \hat{q}_t - \mathcal{E}_t \left[\hat{\lambda}_t \right] \right], \tag{44}$$

- **Monetary policy rule**

$$i_t = \rho_1 i_{t-1} + \rho_2 i_{t-2} + \phi_\pi \pi_t + \phi_y y_t + \varepsilon_t, \tag{45}$$

Notice that, since $\mathbb{I}_t$ contains current realizations of non-monetary shocks (as well as of all t-dated variables except for the policy instrument i_t and the monetary shock ε_t), it is generically the case that $E_{t-1}(\cdot) \neq \mathcal{E}_t(\cdot)$, $g_t \neq \hat{g}_t$ and $q_t \neq \hat{q}_t$; this is relevant for our Monte Carlo exercises based on artificial data generated by the model, since – for a given model – different informational assumptions typically produce different CERs in equilibrium representations; and non-monetary shocks can account for a relatively large share of the variance in non-policy variables.

We adopt two simplifying assumptions relative to Boivin and Giannoni (2006), with no loss of generality: first, we assume away government expenditure (since it only scales up the composite exogenous disturbance by a multiplicative constant, see Rotemberg and Woodford (1997)); second, we stick to a simple inertial Taylor-type feedback rule for monetary policy, by which the nominal interest rate is set on the basis of current deviations of inflation rate and aggregate output from policy targets, as well as its own lags. By the same token, we specify the stochastic properties of demand and supply shocks (along with the monetary policy one), in order to assess whether timing restrictions might induce nonfundamental equilibrium representations for the three observables (output, inflation, nominal interest rate) generated by the model. Formally, we posit[14]

$$g_t = \xi_t^1, \quad \xi_t^1 = \rho_{\xi^1}\xi_{t-1}^1 + \epsilon_t^1,\ \epsilon_t^1 \sim N.i.d.(0,\sigma_1),\ \rho_{\xi^1} \in (0,1) \tag{46}$$

$$q_t = -\xi_t^2, \quad \xi_t^2 = \rho_{\xi^2}\xi_{t-1}^2 + \epsilon_t^2,\ \epsilon_t^2 \sim N.i.d.(0,\sigma_2),\ \rho_{\xi^2} \in (0,1) \tag{47}$$

$$\varepsilon_t = \rho_\varepsilon \varepsilon_{t-1} + \tau_t, \quad \tau_t \sim N.i.d.(0,\sigma_\tau),\ \rho_\varepsilon \in (0,1) \tag{48}$$

5.2 Guerron-Quintana et al. (2017)'s model

Focusing on the prototypical three-equation NK framework, Guerron-Quintana et al. (2017) assume that households and firms both form their expectations about the future evolution of endogenous variables (output gap, inflation rate and nominal interest rate) on the basis of past information relative to the current date. Formally, the model's equilibrium is described by the following equations[15]

- **Dynamic IS equation** (here g_t is the output gap, π_t is the inflation rate, and z_t is a demand shock process)

$$g_t = E_{t-1}[g_{t+1}] - \sigma(E_{t-1}[i_t - \pi_{t+1}] - z_t) \tag{49}$$

- **NK Phillips curve**

$$\pi_t = \beta E_{t-1}[\pi_{t+1}] + \kappa g_t \tag{50}$$

[14] To make the analysis of nonfundamentalness in restricted DSGE models meaningful, we assume that the preference shock to the disutility of producing output, call it ξ_t^2, is different from the preference shock affecting the utility from consuming produced goods, call it ξ_t^1, both shifting up the marginal utility of consumption/disutility of output production.

[15] All variables are expressed as log-deviations from the model's non-stochastic steady state, in equilibrium consumption is equal to output.

- **Monetary policy rule**

$$i_t = \varphi_i i_{t-1} + (1 - \varphi_i)\left[\varphi_\pi \pi_t + \varphi_g g_t\right] + \xi_t \tag{51}$$

- **Exogenous processes for demand and monetary shock**

$$z_t = \rho_z z_{t-1} + \sigma^z \epsilon_t^z, \quad \epsilon_t^z \sim N.i.d.(0,1) \tag{52}$$

$$\xi_t = \sigma^i \epsilon_t^i, \quad \epsilon_t^i \sim N.i.d.(0,1) \tag{53}$$

Letting $\mathcal{E}_t = E[\cdot|\mathbb{I}_t]$ and $\mathbb{I}_t = \{g_{t-j}, \pi_{t-j}, i_{t-1-j}, z_{t-j}, q_{t-j}, \xi_{t-1-j}; j \geq 0\}$, the NK model is brought in closer conformity with the Cholesky-timing assumptions as follows:

- **Dynamic IS equation**

$$g_t = \mathcal{E}_t[g_{t+1}] - \sigma(\mathcal{E}_t[i_t - \pi_{t+1}] - z_t) \tag{54}$$

- **NK Phillips curve**

$$\pi_t = \beta \mathcal{E}_t[\pi_{t+1}] + \kappa g_t + q_t \tag{55}$$

- **Monetary policy rule**

$$i_t = \varphi_i i_{t-1} + (1 - \varphi_i)\left[\varphi_\pi \pi_t + \varphi_g g_t\right] + \xi_t \tag{56}$$

- **Exogenous processes for demand, supply and monetary shocks**

$$z_t = \rho_z z_{t-1} + \sigma^z \epsilon_t^z, \quad \epsilon_t^z \sim N.i.d.(0,1) \tag{57}$$

$$q_t = \rho_q q_{t-1} + \sigma^q \epsilon_t^q, \quad \epsilon_t^q \sim N.i.d.(0,1) \tag{58}$$

$$\xi_t = \sigma^i \epsilon_t^i, \quad \epsilon_t^i \sim N.i.d.(0,1) \tag{59}$$

where a first-order autoregressive supply shock process q_t (e.g. a cost-push shock) has been appended to the Phillips curve (55) to make the analysis of nonfundamentalness meaningful.[16]

[16] In fact, Guerron-Quintana et al. (2017)'s model involves thre endogenous variables and only two shocks; if the former (as we assume in our numerical experiments) are observable, then the ensuing equilibrium MA representation is 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.

Again, given that $\mathbb{I}_t$ contains the current realization of the demand shock z_t (as well as of all t-dated variables except for the policy instrument i_t and the monetary shock ε_t), it is generically the case that $E_{t-1}(\cdot) \neq \mathcal{E}_t(\cdot)$.

6 Numerical experiments

6.1 Cholesky-VARs and the monetary transmission mechanism

Our first experiments are aimed at comparing DNK model-implied impulse responses to those produced by a VAR in (possibly a subset of) the model's endogenous 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 – real output or the output gap, inflation and the policy 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 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 fix the models' parameters at the estimated values from Boivin and Giannoni (2006) and the specification reported in Guerron-Quintana et al. (2017), respectively. Then, 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 estimated average Cholesky-VARs monetary impulse responses and 90% Monte Carlo confidence bands, and plot them against the DGP-consistent IRFs. Operationally, we use $K = 1000$ repetitions, $H = 15$ as the IRF horizon, and $T = 1000$ as length of the artificial data sample, with a burn-in of 200 observations.[17] Figures (2) and (5) portray our estimation results for the baseline parameterization. We also re-run the above described numerical exercises by varying structural parameters that govern the degree of endogenous inertia and/or shock persistence – hence, the strength of the internal propagation mechanism (Figures (3), (4) and (6)), or other key characteristics of the underlying model economy – e.g. the average price duration in a sticky-price setting (Figure (7)).

Tables (2) collects all the parameter choices for the DNK economies under scrutiny. For all of these empirically plausible specifications, 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.

Boivin and Giannoni (2006)													
	β	ω	ψ	η	κ	γ	ϕ_π	ϕ_y	ρ_1	ρ_2	ρ_{ξ^1}	ρ_{ξ^2}	ρ_ε
Fig. 2	0.99	0.47	0.667	0.5	0.008	1	1.5	0	0.602	-0.055	0.9	0.9	0.9
Fig. 3	0.99	0.47	0.667	0	0.008	0	1.5	0	0.602	-0.055	0.9	0.9	0.9
Fig. 4	0.99	0.47	0.667	0.5	0.008	1	1.5	0	0.602	-0.055	0.9	0.9	0.1

Guerron-Quintana et al. (2017)											
	β	σ	φ_π	φ_g	ρ_i	ρ_z	ρ_q	κ	σ_z	σ_q	σ_i
Fig. 5	0.99	1	1.5	0.125	0.75	0.9	0.9	0.0245	0.3	0.1	0.2
Fig. 6	0.99	1	1.5	0.125	0.2	0.1	0.1	0.0245	0.3	0.3	0.2
Fig. 7	0.99	1	1.5	0.125	0.75	0.9	0.9	0.085	1	1	1

Table 2: *Alternative parameterizations for Boivin and Giannoni (2006) and Guerron-Quintana et al. (2017)'s DNK models.*

We argue, however, that truncation effects on the approximating VAR performance

[17] Of 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.

are moderate, for two main reasons. First, 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, for zero restrictions are imposed on impact coefficients rather than long-run effects (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 theoretical VAR(∞) representation of both DNK models features fast decaying coefficients at relatively short horizons. These in fact asymptotically converge to zero as a function of the largest (in modulus) eigenvalue of the F matrix defined in (37), see e.g. Ravenna (2007): for all the employed models and the chosen parameterizations, this eigenvalue is never larger than 0.00002. A truncated VAR model can therefore lend itself as a reasonably good approximation of the true equilibrium reduced form of the restricted DNK model. Figures (2) to (4) contrast the model-implied monetary impulse responses (solid blue line) with those estimated via Cholesky-VARs (solid red lines), when Boivin and Giannoni (2006)'s model is used the DGP. The BIC suggests adopting VAR systems with at most three lags to approximate the true DGP dynamics. Inspection of Figure (2) and (3) reveal clear patterns. First, notwithstanding the presence of an approximation error, the Cholesky-identified IRFs are correctly signed and similar in shape to their theoretical counterparts, with fairly precisely estimated effects at short to longer horizons. Second, this near-perfect identification is attained irrespective of whether the DSGE structure features a relatively weaker internal propagation mechanism that spread the effects of structural disturbances within the model and over time; this is apparent from Figure (3), which is based on the reduction of Boivin and Giannoni (2006)'s model to the one of Rotemberg and Woodford (1997), where no inertia in the endogenous behavior of households (via the formation of consumption habits) and firms (via price indexation to past inflation) is allowed for. Thus, when identification is valid, the VAR methodology does very well in retrieving the transmission of monetary policy shocks. Figure (4) finally illustrates that a markedly larger degree of serial correlation in the non-monetary structural shock processes ($\rho_{\xi^1} = \rho_{\xi^2} = 0.9$) relative to the monetary one ($\rho_{\varepsilon} = 0.1$) does not induce any distortion in the estimates of the adjustment paths of the model's endogenous variables to an unexpected shift in the policy rate.

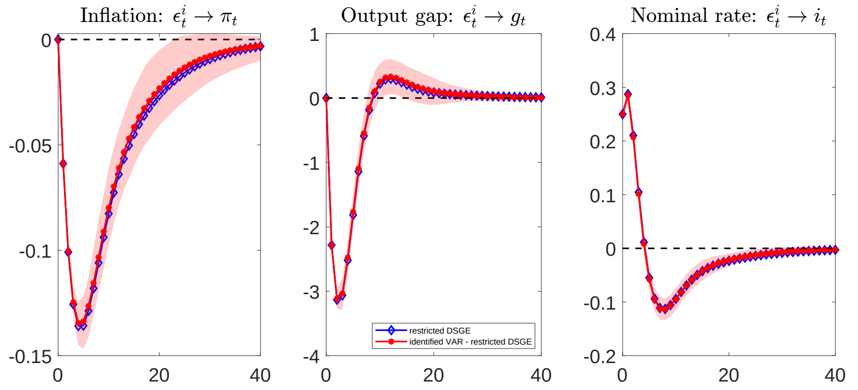

Figure 2: *Impulse response functions (IRFs) to a monetary shock for the Boivin and Giannoni (2006)'s model. See Table 2 for details.*

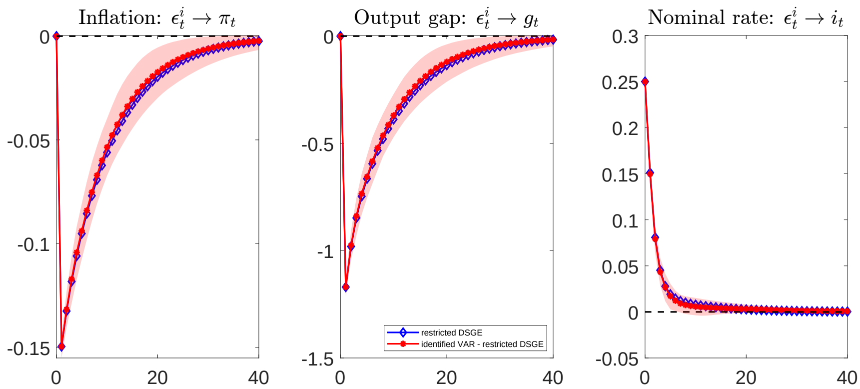

Figure 3: *Impulse response functions (IRFs) to a monetary shock for the Boivin and Giannoni (2006)'s model. See Table 2 for details.*

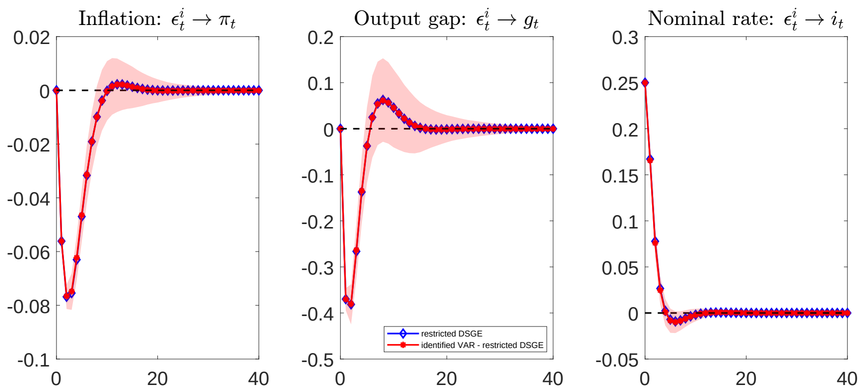

Figure 4: *Impulse response functions (IRFs) to a monetary shock for the Boivin and Giannoni (2006)'s model. See Table 2 for details.*

Compared to Boivin and Giannoni (2006)'s, the benchmark DNK model in Guerron-Quintana et al. (2017) exhibits a considerably lower degree of backward dependence. In fact, both the dynamic IS equation (54) and the NK Phillips curve (55) entail no inertial components, and the only mechanism producing endogenous persistence is represented by the partial adjustment of the current policy rate to its own lags (interest rate smoothing). This specification is not sufficiently rich to strongly propagate the effects of the serially uncorrelated monetary shock across the private sector, with impulse responses fading out at relatively short horizons; all else equal, the lower the degree of monetary policy inertia (here captured by the coefficient ρ_i) and the autocorrelation of non-monetary shocks (ρ_j, $j = z, q$), the faster the output gap and the inflation rate will revert back to their steady state levels – Figure (5) versus Figure (6). For these first two model's specifications, the signs and patterns of the model-based dynamic effects of the monetary shock are closely captured by the Cholesky-VAR, and yet there is no strong evidence that the point estimates are statistically different from zero across all the horizons.

The slope of the New Keynesian Phillips curve (55) too influences impulse response dynamics in general, and the relative response of non-policy variables to an unanticipated deviation from the systematic component of the monetary rule, in particular. This key parameter measures the sensitivity of inflation to changes in the output gap, and is a decreasing function of the degree of price-stickiness. While reduced-form estimation approaches appear to suggest that the pass-through of marginal costs into inflation is fairly limited, thereby pointing to flat Phillips curves, a wide set of estimates – ranging from 0 to 4 – for the slope coefficient have been reported in the DSGE model-based econometric literature (e.g. Schorfheide (2008)). As Figure (7) shows, a steeper Phillips curve engenders a markedly deeper drop in inflation relative to the baseline case ((5)); and the effects of the monetary shock are all more precisely estimated at short to medium horizons even in the presence of higher volatility for the structural disturbances ($\sigma_j, j = q, z, i.$)

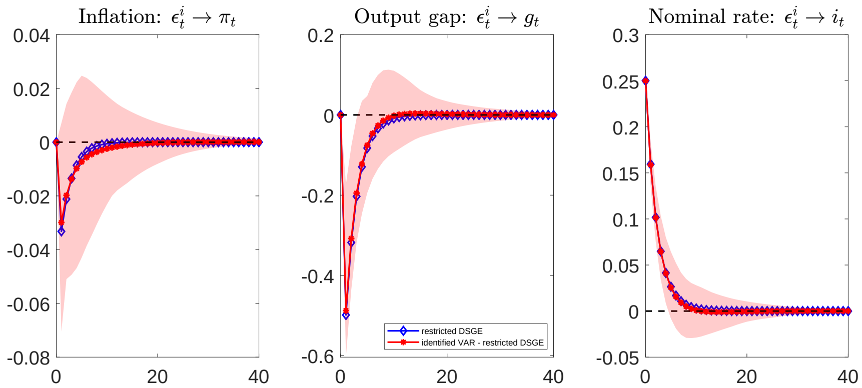

Figure 5: *Impulse response functions (IRFs) to a monetary shock for the Guerron-Quintana et al. (2017)'s model. See Table 2 for details.*

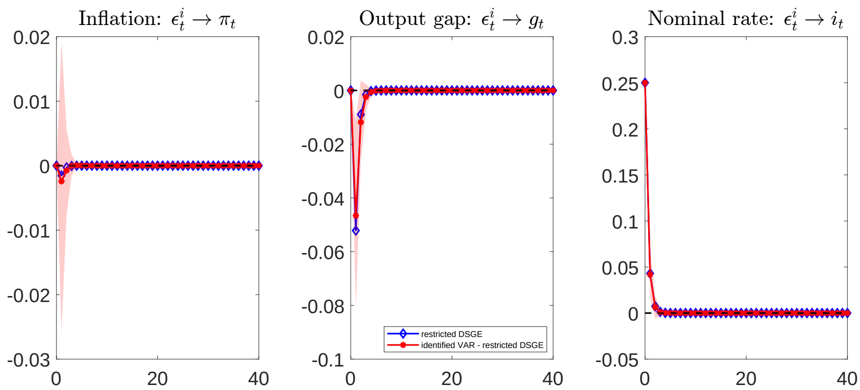

Figure 6: *Impulse response functions (IRFs) to a monetary shock for the Guerron-Quintana et al. (2017)'s model. See Table 2 for details.*

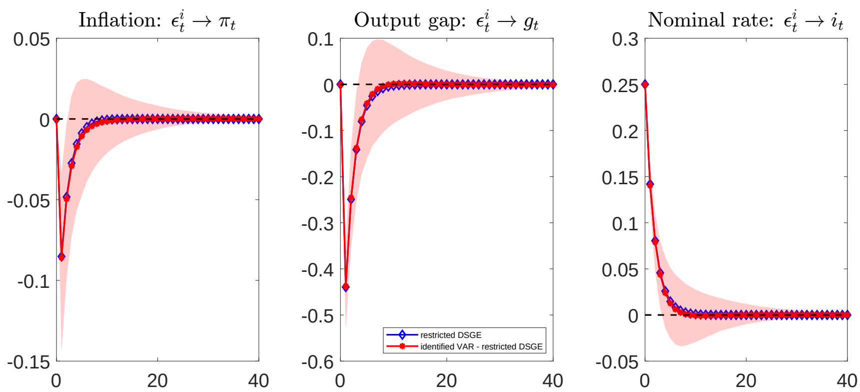

Figure 7: *Impulse response functions (IRFs) to a monetary shock for the Guerron-Quintana et al. (2017)'s model. See Table 2 for details.*

6.2 Cholesky-VARs vs. sign restrictions

Conducted in a Monte Carlo context, Wolf (2020)'s analysis shows that commonly adopted minimal assumptions underlying sign-based approaches to the identification of aggregate effects of monetary shocks may mis-identify monetary impulses as linear combinations of non-monetary structural shocks, that do lie in the identified set and yet produce a counterfactual upsurge in measures of real activity. In a similar context, Castelnuovo (2016) offers simulation evidence suggesting that VAR models identified via sign restrictions work well in uncovering the monetary transmission mechanism provided the true monetary shock accounts for a sufficiently large share of the variance in the variables whose dynamic responses the econometrician is interested in.[18]

When the true DGP complies with the same exact zero impact restrictions as enforced by the Cholesky scheme, it is reasonable to conjecture that Cholesky-VARs would not under-perform vis-à-vis sign restriction-VARs. In order to offer support to this conjecture, we exploit again Boivin and Giannoni (2006)'s model as DGP in the exact same Monte Carlo setup as described in Section (6), except for the fact that the candidate impact matrix for shock identification is not chosen to be the Cholesky factor of the covariance matrix of the reduced-form VAR residuals, but rather from one of its rotations performed via post-multiplication by random orthonormal matrices, provided they fulfill a given set of a priori sign restrictions – see e.g. Fry and Pagan (2011) for a thorough discussion of the procedure.[19]

A (positive) monetary shock is identified by imposing weak sign restrictions, in the spirit of Uhlig (2005), that constrain the response of the policy rate to be positive and the response of inflation to be negative, keeping the sign of the output response unrestricted. We run two distinct experiments, one in which Boivin and Giannoni (2006)'s baseline parameterization is retained and the other where, keeping all the other parameters fixed, the monetary shock is assumed to exhibit significantly lower volatility (standard deviation equal to 0.1 against 1 in the baseline model) as opposed to non-monetary disturbances.

Figure (8) shows that, when the three structural shocks are equally volatile (same unconditional variance), this competing identification scheme points to monetary impulse responses that nearly perfectly reproduce the theoretical ones in both qualitative

[18]See Paustian (2007) for a formal result.

[19]The reduced-form VAR coefficients and covariance matrix are here estimated via OLS. Since sign restriction-VARS are only set identified, we compute the monetary impulse responses as Monte Carlo averages of the pointwise median responses enforced, for each repetition, by one of the admissible rotations.

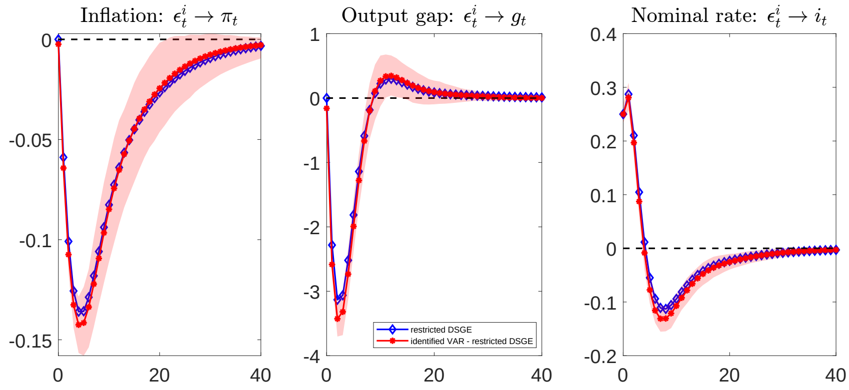

Figure 8: *Impulse response functions (IRFs) to a monetary shock for the Boivin and Giannoni (2006)'s model; same parameterization as in Figure (2), all structural shocks N.i.d. with zero mean and unit variance.*

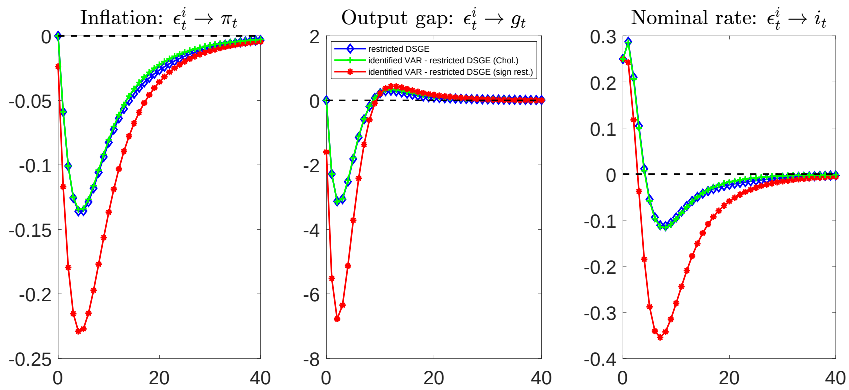

Figure 9: *Impulse response functions (IRFs) to a monetary shock for the Boivin and Giannoni (2006)'s model; same parameterization as in Figure (2), non-monetary shocks N.i.d. with zero mean and unit variance, monetary shock N.i.d. with zero mean and variance equal to* 0.1.

(sign, shape) and quantitative (magnitude, persistence) terms. When the monetary shock variance, by contrast, is relatively low, the sign restriction-VAR dramatically over-estimates the negative short- to medium-run effects of unexpected increases in the policy rate on both inflation and real output, while otherwise uncovering qualitative features of the monetary transmission mechanism in restricted DNK environments. As far as our specific DGP is concerned, the Cholesky-VAR model appears to be immune to this volatility scaling issue, doing remarkably better than its competitor.

6.3 Impulse response matching estimation

Our analysis has clarified the key differences between our information-based approach to modeling timing restrictions in DSGE settings, as a way to bring the latter in close conformity to the mechanics of recursive identification, and earlier proposals (e.g. Boivin and Giannoni (2006). Here we illustrate how the so-amended models can be safely used to derive econometric inference about the structural parameters of DSGE models via IRF matching techniques, whose goal is that of minimizing a suitably weighted average of the distance between model-implied and Cholesky-VARs impulse responses. Specifically, we replicate the impulse-response matching estimation exercise run in Guerron-Quintana et al. (2017) exploiting our restricted version of their NK model (54)-(57).

As mentioned, nonfundamentalness of the theoretical MA polynomials or non-existence of a finite-order VAR representation may engender several complications in the direct comparison of theoretical and empirical impulse responses. While timing restrictions do no cause invertibility issues in Guerron-Quintana et al. (2017), they do induce an infinite order VAR representation of the model's equilibrium. In order to treat the empirical data and the model economy symmetrically, Guerron-Quintana et al. (2017) truncate the true VAR(∞) representation to the same finite order as the one used in the empirical VAR model, and generate the population values for the impulse responses to be matched from the so-approximated VAR.[20]

To allow for comparability, we closely follow Guerron-Quintana et al. (2017)'s experimental design: first, theoretical IRFs are derived from a VAR(p) approximation of the infinite order equilibrium representation of the restricted model, with the same (arbitrary) selection of the lag order p as in the empirical VAR; second, 500 artificial data sets of length $T \in \{100, 232\}$ for inflation and the interest rate are generated from the model being evaluated at the true parameter values (fixed at the baseline parameterization, see Table (2)); third, a truncated VAR(p) model for inflation and the interest rate is fit to each of the data set and the four structural impulse response functions are estimated at horizons $0, \ldots, H$ (consistent with the Cholesky timing assumptions, we impose that the inflation rate does not react contemporaneously to the shock ordered last in the VAR, which it therefore identified as the monetary policy one); finally, impulse responses from the empirical VAR are estimated and point estimates for the

[20] A different (indirect inference) approach could be adopted instead, that replaces the true model-based IRFs with average IRFs that are obtained by repeatedly simulating data from the underlying model (for a given pick of the parameter values) and estimating a structural VAR on simulated data, to be then contrasted with the monetary IRFs generated by the Cholesky-VAR which is rather estimated on the actual sample analog.

parameter α (the probability of a firm not to be allowed to adjust their price, which enters the slope κ of the Phillips curve) are derived by minimizing a weighted average of the distance between the model-based impulse responses and those generated by the VAR, where the weighting matrix is chosen to be the inverse of the bootstrap covariance matrix estimator of the structural impulse responses. As in Guerron-Quintana et al. (2017), all the other structural parameters are assumed to be known in the estimation exercise.[21]

Reported in Table (3), our simulation evidence is largely in line with Guerron-Quintana et al. (2017)'s, with slightly lower Figures for the mean bias. As expected, a larger sample size enhances the precision of the point estimates, all else equal; on the other hand, for a fixed impulse horizon, a higher lag order for the empirical VAR system (hence for the approximating finite-order VAR representation of the structural model) does not generally improve on the estimation bias and the associated standard error.

T	p	H	$\hat{\alpha}$	bias: $\hat{\alpha} - \alpha_0$	$s.e.(\hat{\alpha})$
100	2	2	0.7514	0.0014	0.0309
100	4	2	0.7538	0.0038	0.0311
100	2	4	0.7546	0.0046	0.0355
100	4	4	0.7548	0.0048	0.0327
232	2	2	0.7511	0.0011	0.0286
232	4	2	0.7533	0.0033	0.0295
232	2	4	0.7510	0.0010	0.0353
232	4	4	0.7525	0.0025	0.0325

Table 3: *T denotes the sample size, p the VAR lag order, and H the maximum horizon of the impulse response functions.*

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

[21] We would like to refer the reader to Guerron-Quintana et al. (2017) for further details on the bootstrap version of the minimum distance estimators based on impulse-response matching.

remarkable, and likely impossible to phase out via the impulse response matching procedure.

What happens then, when one feeds a given DSGE model with the set of timing restrictions which reproduce those of Cholesky-VARs? We show that such restrictions may stand as a source of non-trivial VARMA equilibrium representations as well as of model-based nonfundamentalness, making inference from VAR systems unreliable. Our analytical results in this respect have an explicit operational content, as they can be readily applied for the formal construction of recursive (informationally constrained) model structures that admit causal VAR representations.

Adopting a model-based perspective, we then provide simulation evidence on the presence of fundamentalness and lag truncation bias in restricted DNK models of the monetary transmission mechanism. The results of our numerical experiments invariably provide support to the view that Cholesky-VARs succeed in identifying the actual monetary impulse responses in 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.

It shall be stressed once more that the line of research pursued herein relies on aligning DSGE settings with the timing protocol enforced by Cholesky-VARs. The natural question that arises is then whether the so-amended DSGE models are actually consistent with the mechanics of recursive identification. On the flip side, the rapidly increasing literature in structural macroeconomic modeling has also favored the development of alternative sVAR identification approaches – such as sVAR-IV methods and direct coefficient restrictions on policy rules – that tend to restore compatibility with the standard timing of prototypical DSGE frameworks, and can therefore be fruitfully exploited for the identification of the aggregate effects of monetary policy shocks, see e.g. Wolf (2020).

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.

Angelini, G. and Fanelli, L. (2019). Exogenous uncertainty and the identification of structural vector autoregressions with external instruments. *Journal of Applied Econometrics*, 34(6):951–971.

Arias, J. E., Caldara, D., and Rubio-Ramirez, J. F. (2019). The systematic component of monetary policy in svars: An agnostic identification procedure. *Journal of Monetary Economics*, 101:1–13.

Bacchiocchi, E. and Fanelli, L. (2015). Identification in structural vector autoregressive models with structural changes, with an application to us monetary policy. *Oxford bulletin of economics and statistics*, 77(6):761–779.

Beaudry, P., Feve, P., Guay, A., and Portier, F. (2019). When is nonfundamentalness in svars a real problem? *Review of Economic Dynamics*, 34:221–243.

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. and De Nicolo, G. (2002). Monetary disturbances matter for business fluctuations in the g-7. *Journal of Monetary Economics*, 49(6):1131–1159.

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 Hamidi Sahneh, M. (2018). Are small-scale svars useful for business cycle analysis? revisiting nonfundamentalness. *Journal of the European Economic Association*, 16(4):1069–1093.

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.

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.

Fernández-Villaverde, J., Rubio-Ramírez, J. F., Sargent, T. J., and Watson, M. W. (2007). Abcs (and ds) of understanding vars. *American Economic Review*, 97(3):1021–1026.

Forni, M., Gambetti, L., and Sala, L. (2019). Structural vars and noninvertible macroeconomic models. *Journal of Applied Econometrics*, 34(2):221–246.

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.

Fry, R. and Pagan, A. (2011). Sign restrictions in structural vector autoregressions: A critical review. *Journal of Economic Literature*, 49(4):938–60.

Gertler, M. and Karadi, P. (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics*, 7(1):44–76.

Guerron-Quintana, P., Inoue, A., and Kilian, L. (2017). Impulse response matching estimators for dsge models. *Journal of Econometrics*, 196(1):144–155.

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.

Keating, J. W. (1996). Structural information in recursive var orderings. *Journal of economic dynamics and control*, 20(9-10):1557–1580.

Kilian, L. (2013). Structural vector autoregressions. In *Handbook of research methods and applications in empirical macroeconomics*. Edward Elgar Publishing.

Kilian, L. and Lütkepohl, H. (2017). *Structural vector autoregressive analysis*. Cambridge University Press.

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.

Lanne, M., Meitz, M., and Saikkonen, P. (2017). Identification and estimation of non-gaussian structural vector autoregressions. *Journal of Econometrics*, 196(2):288–304.

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. (2003). Computing sunspot equilibria in linear rational expectations models. *Journal of Economic dynamics and control*, 28(2):273–285.

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.

Paustian, M. (2007). Assessing sign restrictions. *The BE Journal of Macroeconomics*, 7(1).

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.

Rondina, G. and Walker, T. B. (2016). Learning and informational stability of dynamic ree with incomplete information. *Review of Economic Dynamics*, 21:147–159.

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.

Schorfheide, F. (2008). Dsge model-based estimation of the new keynesian phillips curve. *FRB Richmond Economic Quarterly*, 94(4):397–433.

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 nonfundamentalness 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).

Stock, J. H. and Watson, M. W. (2018). Identification and estimation of dynamic causal effects in macroeconomics using external instruments. *The Economic Journal*, 128(610):917–948.

Uhlig, H. (2005). What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics*, 52(2):381–419.

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

Appendices (intended for online publication only)

A.1 Solution algorithm for DSGE models with timing restrictions

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

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

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

$$g_x = \begin{pmatrix} g_{x_u} & g_{x_r} \\ j_{x_u} & j_{x_r} \end{pmatrix}, \quad h_x = \begin{pmatrix} h_{x_u} & h_{x_r} \\ 0 & P \end{pmatrix} \tag{60}$$

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, j_{x_u} is $(n_{y_r} \times n_{x_u})$-dimensional, 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

$$\hat{g}_{x_u} = g_{x_u}, \quad \hat{j}_{x_u} = j_{x_u}, \quad \hat{h}_{x_u} = h_{x_u}, \tag{61}$$

$$\hat{g}_\sigma = 0, \quad \hat{j}_\sigma = 0, \quad \hat{h}_\sigma = 0; \tag{62}$$

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

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

Then compute

$$\begin{pmatrix} \hat{h}_{x_{r,-1}} \\ \hat{g}_{x_{r,-1}} \end{pmatrix} = -\nabla(f^1)^{-1} f^1_{y_r} j_{x_r} P \tag{63}$$

$$\begin{pmatrix} \hat{h}_{x_r} \\ \hat{g}_{x_r} \end{pmatrix} = \begin{pmatrix} h_{x_r} \\ g_{x_r} \end{pmatrix} + \nabla(f^1)^{-1} f^1_{y_r} j_{x_r} \tag{64}$$

$$\hat{j}_{x_{r,-1}} = j_{x_r} P \tag{65}$$

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

$$\begin{pmatrix} x'_u \\ x'_r \\ x_r \end{pmatrix} = \begin{pmatrix} \hat{h}_{x_u} & \hat{h}_{x_r} & \hat{h}_{x_{r,-1}} \\ 0_{n_{x_r} \times n_{x_u}} & P & 0_{n_{x_r} \times n_{x_r}} \\ 0_{n_{x_r} \times n_{x_u}} & I_{n_{x_r} \times n_{x_r}} & 0_{n_{x_r} \times n_{x_r}} \end{pmatrix} \begin{pmatrix} x_u \\ x_r \\ x_{r,-1} \end{pmatrix} + \sigma \begin{pmatrix} \epsilon_{x_u} \\ \epsilon_{x_r} \end{pmatrix} \tag{66}$$

$$\begin{pmatrix} y_u \\ y_r \end{pmatrix} = \begin{pmatrix} \hat{g}_{x_u} & \hat{g}_{x_r} & \hat{g}_{x_{r,-1}} \\ \hat{j}_{x_u} & 0_{n_{y_r} \times n_{x_r}} & \hat{j}_{x_{r,-1}} \end{pmatrix} \begin{pmatrix} x_u \\ x_r \\ x_{r,-1} \end{pmatrix} \tag{67}$$

A.2 Solving the bi-variate example

Consider again the model

$$\mathcal{E}\left[y_{1,t} - \alpha y_{1,t+1} - y_{2,t} - x_{1,t} | \mathbb{I}_{1,t}\right] = 0, \quad \alpha > 0, \tag{68}$$

$$\mathcal{E}\left[y_{2,t} - \beta y_{2,t+1} - x_{1,t} - x_{2,t} | \mathbb{I}_{2,t}\right] = 0, \quad \beta > 0, \tag{69}$$

$$x_{1,t} = \rho_1 x_{1,t-1} + \varepsilon_{1,t}, \quad |\rho_1| < 1, \tag{70}$$

$$x_{2,t} = \rho_1 x_{2,t-1} + \varepsilon_{2,t}, \quad |\rho_2| < 1, \tag{71}$$

$$\varepsilon_{1,t} \sim NID(0,1), \quad \varepsilon_{2,t} \sim NID(0,1) \tag{72}$$

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

$$y_u = [y_2], \quad y_r = [y_1] \tag{73}$$

$$x_u = [x_1], \quad [x_r] \tag{74}$$

and thereby the following informational partition

$$
\begin{aligned}
f^0 &= (y_{1,t} - \alpha y_{1,t+1} - y_{2,t} - x_{1,t}) && (75)\\
f^1 &= \begin{pmatrix} y_{2,t} - \beta y_{2,t+1} - x_{1,t} - x_{2,t} \\ x_{1,t+1} - \rho_1 x_{1,t} - \varepsilon_{1,t+1} \end{pmatrix} && (76)\\
f^{x_r} &= (x_{2,t+1} - \rho_2 x_{1,t} - \varepsilon_{2,t+1}) && (77)
\end{aligned}
$$

Recall that the determinate solution (9) to the unrestricted model can be partitioned as follows:

$$
\begin{aligned}
j_{x_u} = \tfrac{2-\beta\rho_1}{(1-\alpha\rho_1)(1-\beta\rho_1)}, \qquad j_{x_r} = \tfrac{1}{(1-\alpha\rho_2)(1-\beta\rho_2)}, && (78)\\
g_{x_u} = \tfrac{1}{1-\beta\rho_1}, \qquad g_{x_r} = \tfrac{1}{1-\beta\rho_2}, && (79)\\
h_{x_u} = \rho_1, \qquad h_{x_r} = 0, \qquad P = \rho_2 && (80)
\end{aligned}
$$

It is easily seen that $f^1_{y_r} = [0,0]'$ 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

$$
\begin{aligned}
y_{1,t} &= \frac{2-\beta\rho_1}{(1-\alpha\rho_1)(1-\beta\rho_1)} x_{1,t} + \frac{\rho_2}{(1-\alpha\rho_2)(1-\beta\rho_2)} x_{2,t-1}, && (81)\\
y_{2,t} &= \frac{1}{1-\beta\rho_1} x_{1,t} + \frac{1}{1-\beta\rho_2} x_{2,t} && (82)
\end{aligned}
$$

which is in the same form as (10).

A.3 Indeterminacy and timing restrictions: The example revisited

Our analysis is conducted on the assumption that the DSGE model exhibits equilibrium determinacy. Since timing restrictions are generically neutral with respect to the determinacy/indeterminacy properties of the RE equilibrium (the restricted model simply inherits those of its unrestricted counterpart), any identification and/or estimation issues that might arise when data are generated by a restricted DGP are unrelated to the appearance of arbitrary reduced form components (such as non-structural parameters and/or sunspot shocks) of the model's equilibrium representation, and can therefore be ascribed entirely to the different timing of decisions and expectations formation.

Allowing for equilibrium indeterminacy dramatically changes this picture. Perhaps surprisingly, timing restrictions can just *cease to matter* for an arbitrarily large set of indeterminate solutions, and thereby play no role in shaping patterns of impulse propagation over time.

To see this, let us revisit the stylized bi-variate RE system (1)-(5). Recall that in the conventional, unrestricted setting, the RE model has the form

$$y_t = AE_t[y_{t+1}] + Bx_t, \tag{83}$$

$$x_t = Cx_{t-1} + \varepsilon_t \tag{84}$$

where

$$A = \begin{pmatrix} \alpha & \beta \\ 0 & \beta \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{pmatrix} \tag{85}$$

Let $\alpha, \beta > 1$, then A^{-1} is a stable matrix and the RE model features an indeterminate equilibrium of the form

$$y_t = A^{-1}y_{t-1} - A^{-1}Bx_{t-1} + \eta_t \tag{86}$$

where $\eta_t = [\eta_{1,t}, \eta_{2,t}]'$ is the vector of (endogenous) forecast errors $\eta_{1,t} = y_{1,t} - E_{t-1}[y_{1,t}]$ and $\eta_{2,t} = y_{2,t} - E_{t-1}[y_{2,t}]$. By construction, no stability condition is required to restrict the forecast errors, which thus can be arbitrarily chosen within the family of covariance-stationary martingale difference sequences. Following Lubik and Schorfheide (2003), let η_t be specified as a linear function of the structural shocks ε_t and a conformable vector of reduced form sunspot shocks ζ_t^* satisfying $E_{t-1}[\zeta_t^*] = 0$

$$\eta_t = \tilde{M}\varepsilon_t + \zeta_t^* \tag{87}$$

where $\tilde{M}$ is a square, arbitrary coefficient matrix. Multiple RE equilibria in the indeterminacy region of the model's parameter space are thus indexed by the arbitrary selection of $\tilde{M}$ (*parametric indeterminacy*) and of the stochastic features of the sunspot shocks ζ_t^*, provided these are orthogonal to the information set at date $t-1$ (*stochastic indeterminacy*).

As a result, the full set of RE solutions under indeterminacy can be written in VARMA-type form

$$\left(I - A^{-1}L\right) y_t = \left(\tilde{M} - A^{-1}B\left(I - CL\right)^{-1} L\right) \epsilon_t + \zeta_t^* \tag{88}$$

where L is the standard lag operator. Notice that, since sunspot variables enlarge the (already infinite and uncountable) set of equilibrium-consistent endogenous forecast errors, they do not produce yet only amplify the indeterminacy issue. For the sake of clarity, let us focus attention on the parametric indeterminacy case (assuming e.g. $\zeta_t^* = 0$ almost surely for all t); the previous system can then be equivalently written in the following state-space form

$$y_t = G s_t, \tag{89}$$

$$s_t = F s_{t-1} + H \varepsilon_t, \tag{90}$$

where $s_t = [y_{t-1}, x_t]'$ is the expanded state vector and

$$G = \begin{pmatrix} A^{-1} & \tilde{M}\,(I - CL) - A^{-1}BL \end{pmatrix}, \; F = \begin{pmatrix} A^{-1} & -A^{-1}BL \\ 0 & C \end{pmatrix}, \; H = \begin{pmatrix} \tilde{M}L \\ I \end{pmatrix}$$

Now, in order to pin down a specific solution in the (parametrically) indeterminate set, most of the literature adopts the so-called *orthogonality assumption* $\tilde{M} = 0$, according to which the component of the forecast errors due to the structural shocks is orthogonal to the contribution of the sunspot shocks (Lubik and Schorfheide (2003)). In the present setting, both structural shocks ε_t would not affect the endogenous variables contemporaneously if $\tilde{M} = 0$. On the other hand, for any lower triangular $\tilde{M}$ matrix (e.g. $\tilde{M} = I_{2\times 2}$), timing restrictions *do not* affect the dynamic properties of the RE solution under indeterminacy. To see this, let

$$\tilde{M} = \begin{pmatrix} \tilde{m}_1 & 0 \\ \tilde{m}_3 & \tilde{m}_4 \end{pmatrix},$$

and notice that from the matrix G in the state-space representation (23)-(24) we obtain

$$\tilde{M}\,(I - CL) - A^{-1}BL = \begin{pmatrix} \tilde{m}_1(1-\rho_1 L) - \alpha^{-1}L & 0 \\ \tilde{m}_3(1-\rho_1 L) - \beta^{-1}L & \tilde{m}_4(1-\rho_2 L) - \beta^{-1}L \end{pmatrix},$$

Lower triangularity of this matrix implies that, for any arbitrary choice of the coefficients $(\tilde{m}_1, \tilde{m}_2, \tilde{m}_3)$, the endogenous variable $y_{1,t}$ does not move with the state $x_{2,t}$ contemporaneously. Intuitively, for this whole family of indeterminate solutions, backward-looking dynamics prevail, thereby removing the incidence of the unobserved state filtering problem that the decision-maker controlling $y_{1,t}$ faces in the presence of timing restrictions. As a result, these restrictions add no further constraint on the equilibrium dynamics for such a variable and, *a fortiori*, for the informationally unconstrained one ($y_{2,t}$): these indeterminate solutions can thus exhibit a fundamental MA representation no matter whether informational constraints would cause their determinate counterpart to display nonfundamentalness.[22]

A.4 VARs and Cholesky identification

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

$$y_t = A_1 y_{t-1} + \ldots + A_k y_{t-k} + u_t, \qquad E[u_t] = 0,\ E\left[u_t u_t'\right] = V \tag{91}$$

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.

$$B_0 u_t = \epsilon_t, \tag{92}$$

with B_0 being a square, full-rank matrix. The VAR (91) thus admits the structural representation

$$B_0 y_t = B_1 y_{t-1} + \cdots + B_k y_{t-k} + \epsilon_t, \qquad B_i := B_0 \cdot A_i,\ i = 1, \ldots, q \tag{93}$$

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, $i = 0, \ldots k$. 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 (91) and thus treated as known, the B_0 matrix will be completely identified by imposing $n_y(n_y+1)/2$ identifying constraints.

[22]There of course exist further solutions in the indeterminacy set – e.g. sunspot-free ones that are indexed by non-triangular $\tilde{M}$ coefficient matrices, and/or those emerging when $\tilde{A}^{-1}$ displays only one stable eigenvalue – for which this invariance property fails to hold. Which of the continuously infinite solutions prevails as the true DGP is ultimately an empirical question.

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.

A.5 Additional DGP: Leeper and Leith (2016)

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

$$
\begin{aligned}
g_t &= E_t[g_{t+1}] - \sigma(i_t - E_t[\pi_{t+1}]) + \omega_t^g && (94)\\
\pi_t &= \beta E_t[\pi_{t+1}] + \kappa g_t + \omega_t^\pi && (95)\\
i_t &= \varphi_\pi \pi_t + \varphi_g g_t + \omega_t^i && (96)\\
P_t^m &= -i_t + \beta\rho E_t[P_{t+1}^m] && (97)\\
b_t^m &= \beta^{-1} b_{t-1}^m + (\rho - 1)P_t^m + (1-\beta^{-1})s_t - \beta^{-1}\pi_t && (98)\\
\omega_t^j &= \rho_j \omega_{t-1}^j + \epsilon_t^j, \quad |\rho_j| < 1, \quad j = g, \pi, i && (99)\\
s_t &= \rho_s s_{t-1} + \epsilon_t^s, \quad |\rho_s < 1 && (100)
\end{aligned}
$$

where (97) is no-arbitrage condition linking bond prices to the one-period nominal interest rate and (98) 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 equlibria 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. $\varphi_\pi \in \left[0, 1 - \frac{1-\beta}{\kappa}\varphi_g\right)$.

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.

All the 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 the DNK model under scrutiny. 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.

Figures (10) to (12) 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 (10) and (11) reinforce our claim about the good approximating performance of Cholesky-VARs when applied to a DGP exhibiting a Cholesky-style recursive structure.

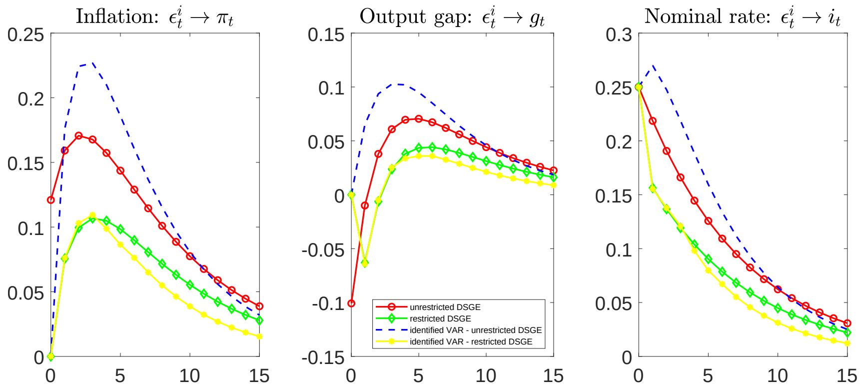

Figure 10: *Impulse response functions (IRFs) to a monetary shock for the Leeper and Leith (2016)'s model with* $\rho = 0$.

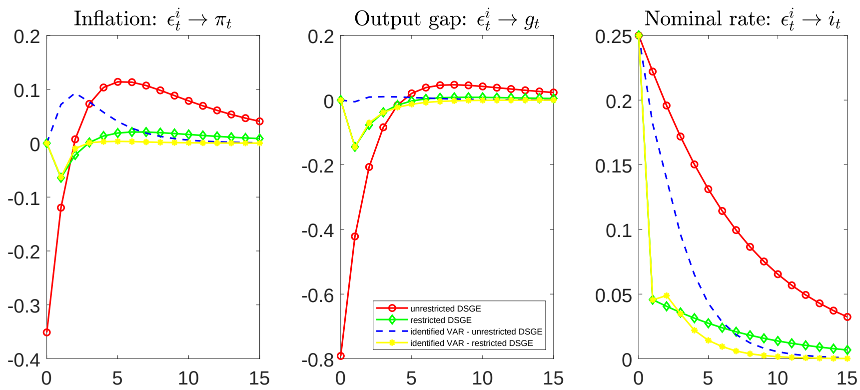

Figure 11: *Impulse response functions (IRFs) to a monetary shock for Leeper and Leith (2016)'s model with* $\rho = 0.95$.

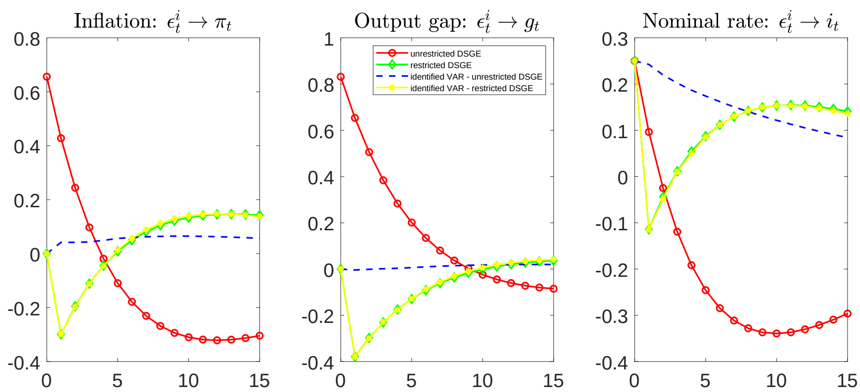

Figure 12: *Impulse response functions (IRFs) to a monetary shock for Leeper and Leith (2016)'s model with* $\rho = 0.95$ *and* $\rho_\pi = 0$.

Most remarkably, Figure (12) 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.[23]

[23] This observation also qualifies the results in Wolf (2020) as it points to the possibility of severe misidentification 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. We obtain similar findings (available on request) in the context of Benati and Surico (2009)'s hybrid NK framework, once information-based timing restriction are properly introduced in the model.