Exploring the Reliability of Policy Counterfactuals Based on Structural VARs∗

Luca Benati
University of Bern†

Abstract

Based on standard New Keynesian models, I explore the ability of policy counterfactuals based on the theoretical structural VAR representations of the models to replicate the impact on the reduced-form properties of the economy of changes in the DSGE model’s monetary rule. I show that SVAR-based counterfactuals may perform poorly, failing to reliably capture the authentic impact of changes in monetary rule in the underlying DSGE model. The problem appears to be comparatively more severe at the low frequencies, thus implying that it should be relevant for SVAR-based counterfactual analyses of historical episodes such as the Great Inflation and the Great Depression, characterized by persistent and drawn-out fluctuations in key macroeconomic variables.

Keywords: Lucas critique; structural VARs; policy counterfactuals; DSGE models; Taylor rules; monetary policy; Great Depression; Great Inflation; Great Moderation.

∗I wish to thank K. Adam, C. Favero, R. Farmer, C. Hellwig, R. Reis, and especially F. Canova for very helpful discussions; and seminar participants at Cambridge, Pompeu Fabra, PSE, Bern, Cass Business School, HEC Paris, Università di Pavia, and Banque de France for comments. Part of this material was presented at the EABCN conference ‘After the Crisis: A New Agenda for Business Cycle Research?’, at the Handbook of Monetary Economics conference held at the ECB, and at the 2010 Carlo Giannini conference held at the Banca d’Italia. Special thanks to my discussants—M. Wickens, L. Reichlin, and F. Ferroni, respectively—for helpful comments; and to H. Pesaran for spotting a math mistake in a previous version. Usual disclaimers apply.

†Department of Economics, University of Bern, Schanzeneckstrasse 1, CH-3001 Bern, Switzerland. Email: luca.benati@vwi.unibe.ch
1 Introduction

Since Sims (1980) introduced the VAR methodology into macroeconomics, monetary policy counterfactuals—in which the interest rate equation of the estimated structural VAR (henceforth, SVAR) for period I is imposed upon the estimated SVAR for period II—have been one of its most prominent applications.\footnote{These counterfactuals have been used, for example, by Sims (1998) to explore the role played by monetary policy in the Great Depression, and by Sims and Zha (2006) to assess the role played by monetary policy in causing/allowing the Great Inflation, and fostering the generalised fall in macroeconomic volatility associated with the Great Moderation.}

From the outset, it has been known that SVAR-based policy counterfactuals are in principle vulnerable to the Lucas critique, and they are therefore potentially unreliable. Thomas Sargent (1979), in particular pointed out that

\footnote{See, in particular, Sims (1982), Sargent (1984), and Sims (1986).}

\footnote{The reader may find this hard to believe, especially given the stature of the researchers involved. I urge the reader to go through all the papers of the Sargent-Sims debate: I was not able to find a single one in which the discussion/analysis was conducted conditional on a specific general equilibrium model.}

\[\text{[u]sers of [VARs] must recognize that the range of uses of these models is more limited than the range of uses that would be possessed by a truly structural simultaneous equations model. [...]}\]

One use to which [they] cannot be put is to evaluate the effect of policy interventions in the form of changes in the feedback rule governing a monetary or fiscal policy variable. \[\text{[...]}\]

The reason it is not appropriate is to be found in the dynamic economic theory alluded to above and described by Lucas and Sargent (1979). That body of theory delivers a set of cross-equations restrictions which imply that when one equation \[\text{[...]}\] describing a policy authority’s feedback rule changes, in general, all of the remaining equations will also change.

The first half of the 1980s saw an intense debate between Sargent and Sims on the reliability of SVAR-based policy counterfactuals.\footnote{The reader may find this hard to believe, especially given the stature of the researchers involved. I urge the reader to go through all the papers of the Sargent-Sims debate: I was not able to find a single one in which the discussion/analysis was conducted conditional on a specific general equilibrium model.} Around the mid-1980s, however, the debate was essentially abandoned, without any substantive agreement having been reached. A key reason for this, in my own view, is that the entire discussion was conducted ‘verbally’, that is, without any reference to specific structural general equilibrium models.\footnote{The reader may find this hard to believe, especially given the stature of the researchers involved. I urge the reader to go through all the papers of the Sargent-Sims debate: I was not able to find a single one in which the discussion/analysis was conducted conditional on a specific general equilibrium model.}

In this paper I attempt to provide some answers to the questions left ‘hanging over’ about three decades ago by the end of the debate between Sargent and Sims.

1.1 This paper: methodology and key results

Based on estimated standard New Keynesian models, I perform a systematic investigation of the reliability of SVAR-based policy counterfactuals, where by ‘reliability’
I mean ‘the ability of such counterfactuals to correctly capture the impact on the (reduced-form) properties of the economy of changes in the monetary rule within the New Keynesian model’.

The paper’s main results may be summarised as follows:

- SVAR-based counterfactuals appear, in general, as unreliable, exhibiting an inability to correctly capture the impact on the economy of changes in the monetary rule within the underlying DSGE model.

- The size of the errors made by SVAR-based policy counterfactuals—compared to the authentic, DSGE-based counterfactuals—is sometimes substantial, thus implying that they might well produce quite significantly distorted inference.

- Unreliability appears to be comparatively more severe at the low frequencies, thus implying that SVAR-based counterfactuals may fare especially badly in assessing the role played by monetary policy in causing phenomena such as the Great Inflation and the Great Depression, two episodes characterised by significant low-frequency fluctuations in inflation, and, in the case of the Depression, in output.

- Unreliability pertains not only to individual series’ characteristics—such as a series’ counterfactual path; its theoretical standard deviation; or its reduced-form innovations within a VAR context—but also to key aspects of the relationships among series, such as their unconditional correlation; the gain and coherence between them; and their lead-lag relationship as captured by either the phase angle or the delay.

- I show analytically that the problem (i) is a straightforward implication of the cross-equations restrictions imposed by rational expectations on a model’s structural solution—that is, it precisely originates from the shortcoming originally discussed by Sargent (1979); and (ii) it only disappears when the model’s solution is vector white noise.

- Finally, unreliability depends not only on the extent of the policy shift, but also—and crucially—on key structural characteristics of the economy, such as the extent of forward-, as opposed to backward-looking behavior.

1.2 Implications

These results have two main implications:

first, they suggest that the outcomes of SVAR-based policy counterfactuals should be taken with caution, as their informativeness for the issue at hand—e.g., understanding the role played by monetary policy in exacerbating the Great Depression,
causing the Great Inflation, or fostering the Great Moderation—is, principle, open to question.

Second—and more subtly—since the extent of reliability of SVAR-based counterfactuals crucially depends on unknown structural characteristics of the underlying data generation process, these results imply that reliability cannot simply be assumed, and can rather only be ascertained with a reasonable degree of confidence by estimating structural (DSGE) models. Eschewing estimation of structural macroeconomic models, and performing inference by imposing a minimal set of credible restrictions on the moving-average representation of the data is however the entire point of structural VAR analysis. As this paper shows, unfortunately, one important application of such methodology appears to suffer from a key logical problem, as, in general, its reliability can only be ascertained via structural (e.g., DSGE) estimation.

The paper is organised as follows. The next section briefly discusses the conceptual essence of the problem, stressing the non-equivalence between two alternative notions of policy counterfactuals: the authentic counterfactual (which is performed by switching the Taylor rules within the DSGE model), and the SVAR-based counterfactual (which is instead performed by switching the interest rate rules within the theoretical SVAR representation implied by the very same DSGE model). Section 3 provides several illustrations via numerical methods, based on three New Keynesian models with increasing extent of complexity, and conditional on grids of values for the parameters of the Taylor rule. Section 4 presents analytical illustrations of the problem at hand, and explores via numerical methods the role played by specific model features in generating these results. Section 5 provides a tentative assessment of the practical relevance of the problem, based on Smets and Wouters’ (2007) estimated medium-scale DSGE model of the U.S. economy.

2 The Problem in a Nutshell

2.1 The intuition

The essence of the problem can be succinctly described as follows. Consider a structural macroeconomic model—for the sake of the argument, a standard New Keynesian model—and assume (again, for the sake of the argument) that monetary policy follows the simple Taylor rule with smoothing

\[ R_t = \rho R_{t-1} + (1 - \rho)\left[ \phi_\pi \pi_t + \phi_y y_t \right] + \epsilon_{R,t} \]  

where \( R_t, \pi_t \) and \( y_t \) are the nominal rate, inflation and the output gap; \( \epsilon_{R,t} \) is a disturbance to the monetary rule; and \( \rho, \phi_\pi, \) and \( \phi_y \) are the smoothing coefficient, and the coefficients on inflation and the output gap, respectively.
Consider then two sets of parameters for the Taylor rule:

\[
\text{Taylorn} \equiv [\rho^1, \phi^1_x, \phi^1_y]^T \\
\text{Taylorn}^2 \equiv [\rho^2, \phi^2_x, \phi^2_y]^T
\]

with \( \text{Taylorn} \neq \text{Taylorn}^2 \). Together with the other structural parameters of the model, equation (1), and the equations describing the behaviour of the private sector, \( \text{Taylorn} \) and \( \text{Taylorn}^2 \) imply two different structures, with two different reduced-form VARMA representations, and therefore, as a logical corollary, two different structural VARMA representations, that is:

\[
\text{Taylorn} \implies \text{DSGE}^1 \implies \text{VARMA}^1 \implies \text{Structural VARMA}^1 \implies \text{MonetaryRule}^1 \\
\text{Taylorn}^2 \implies \text{DSGE}^2 \implies \text{VARMA}^2 \implies \text{Structural VARMA}^2 \implies \text{MonetaryRule}^2
\]

where \( \text{MonetaryRule}^1 \) and \( \text{MonetaryRule}^2 \) are the interest rate equations in the two structural VARMA representations, \( \text{Structural VARMA}^1 \) and \( \text{Structural VARMA}^2 \).

### 2.1.1 Two alternative notions of policy counterfactual

Switching \( \text{Taylorn}^1 \) and \( \text{Taylorn}^2 \) within the DSGE model is the *authentic policy counterfactual*, where the adjective ‘authentic’ simply comes from the fact that such a policy counterfactual is performed

- based on the authentic structure of the economy—the DSGE model—and
- based on the authentic monetary policy rule—the Taylor rule (1).

Switching \( \text{MonetaryRule}^1 \) and \( \text{MonetaryRule}^2 \), on the other hand, is the *SVARMA-based policy counterfactual*, that is, the counterfactual performed by switching the interest rate equations in the theoretical structural VARMA representations of the DSGE model generated conditional on the two Taylor rules.

The key issue, then—and the focus of this paper—is that *switching MonetaryRule* \(^1\) and \( \text{MonetaryRule}^2 \) *is not the same as switching Taylor* \(^1\) and \( \text{Taylor}^2 \), in terms of their impact on (the properties of) the economy. On the contrary, as this paper will show the difference is sometimes substantial.

### 2.2 A formal argument

Let the structural VARMA representation of a DSGE model’s solution be

\[
Y_t = B_1 Y_{t-1} + \ldots + B_p Y_{t-p} + A_0 \epsilon_t + A_1 \epsilon_{t-1} + \ldots + A_q \epsilon_{t-q}
\]

where \( Y_t \equiv [R_t, X_t]^T \) is an \( N \times 1 \) vector of endogenous variables, with \( R_t \) being the nominal interest rate and \( X_t \) being an \( (N-1) \times 1 \) vector of variables other than \( R_t \);
\( A_0 \) being the impact matrix of the structural shocks at zero; \( B_1, \ldots, B_p \) being the AR matrices of the VARMA; \( A_1, \ldots, A_q \) being the MA matrices; and \( \epsilon_t = A_0^{-1} u_t \) —where \( u_t \) is the \( N \times 1 \) vector collecting the reduced-form innovations—being a vector collecting the structural shocks. The vector \( \epsilon_t \) is defined as \( \epsilon_t = [\epsilon_{R,t}, \epsilon_{-R,t}]' \), where \( \epsilon_{R,t} \) is the monetary policy shock (that is, the shock to the Taylor rule), and \( \epsilon_{-R,t} \) is a vector collecting all the structural shocks other than \( \epsilon_{R,t} \). Let’s define \( \tilde{B}_0 \equiv A_0^{-1} B_1, \ldots, \tilde{B}_p \equiv A_0^{-1} B_p; \tilde{A}_1 \equiv A_0^{-1} A_1, \ldots, \tilde{A}_q \equiv A_0^{-1} A_q; \) and let’s partition \( \tilde{B}_0, \tilde{B}_1, \ldots, \tilde{B}_p, \tilde{A}_1, \ldots, \tilde{A}_q \) as

\[
\begin{bmatrix}
\tilde{B}_0(\theta_M, \theta^- M) \\
\tilde{B}_1(\theta_M, \theta^- M) \\
\tilde{A}_1(\theta_M, \theta^- M)
\end{bmatrix}
= \begin{bmatrix}
\tilde{B}_0^R(\theta_M, \theta^- M) \\
\tilde{B}_1^R(\theta_M, \theta^- M) \\
\tilde{A}_1^R(\theta_M, \theta^- M)
\end{bmatrix}, \ldots,
\begin{bmatrix}
\tilde{B}_p(\theta_M, \theta^- M) \\
\tilde{B}_p^R(\theta_M, \theta^- M) \\
\tilde{A}_p(\theta_M, \theta^- M) \\
\tilde{A}_p^R(\theta_M, \theta^- M)
\end{bmatrix}
\]

where \( \theta_M \) is a vector collecting the parameters of the DSGE model’s monetary policy rule—that is, within the present context, \( \theta_M \equiv [\rho, \phi_z, \phi_y] \); \( \theta^- M \) is a vector collecting the DSGE model’s other structural parameters; \( \tilde{B}_j^R(\theta_M, \theta^- M), j = 0, 1, \ldots, p, \) is the first row of \( \tilde{B}_j(\theta_M, \theta^- M), \) that is, the one corresponding to the monetary policy equation of the SVARMA (the fact that the monetary policy shock \( \epsilon_{R,t} \) is the first shock in \( \epsilon_t \) automatically identifies the first equation of the structural VARMA as the monetary policy equation; on this, see e.g. Arias, Caldara, and Rubio-Ramirez (2014)); \( \tilde{B}_j^R(\theta_M, \theta^- M) \) is a \((N-1) \times N\) matrix collecting the other equations of the SVAR representation of the DSGE model; and the \( \tilde{A}_j(\theta_M, \theta^- M), \tilde{A}_j^R(\theta_M, \theta^- M), \) and \( \tilde{A}_j^R(\theta_M, \theta^- M), j = 0, 1, \ldots, q, \) have an analogous interpretation. In (3) and (4) I have made explicit the functional dependence of all of the entries of the matrices \( \tilde{B}_j \) and \( \tilde{A}_j \) on \( \theta \equiv [\theta_M, \theta^- M] \): as it is well known, this is a straightforward implication of the cross-equations restrictions imposed by rational expectations on the solution of a general equilibrium model, and it therefore holds without any loss of generality.\(^4\) Since, throughout the entire paper, the non-policy parameters (the \( \theta^- M \)) are kept constant, in order to simplify the notation from now on I ignore them.

Consider now two alternative policy parameters’ vectors, \( \theta^1_M \) and \( \theta^2_M \), with \( \theta^1_M \neq \theta^2_M \), which imply the following two SVARMA representations

\[
\begin{align*}
\tilde{B}_0(\theta^1_M)Y_t &= \tilde{B}_1(\theta^1_M)Y_{t-1} + \ldots + \tilde{B}_p(\theta^1_M)Y_{t-p} + \epsilon_t + \tilde{A}_1(\theta^1_M)\epsilon_{t-1} + \ldots + \tilde{A}_q(\theta^1_M)\epsilon_{t-q} \\
\tilde{B}_0(\theta^2_M)Y_t &= \tilde{B}_1(\theta^2_M)Y_{t-1} + \ldots + \tilde{B}_p(\theta^2_M)Y_{t-p} + \epsilon_t + \tilde{A}_1(\theta^2_M)\epsilon_{t-1} + \ldots + \tilde{A}_q(\theta^2_M)\epsilon_{t-q}
\end{align*}
\]

\(^4\)In section 4 below we will show that these restrictions do not hold only in the extreme case in which the model solution is vector white noise, as in such case \( \theta^1_M \) drops out of \( \tilde{B}_j^R \) for all \( j \).
The policy counterfactual associated with imposing the SVARMA’s structural monetary rule\(^5\) for regime 2 onto the SVARMA for regime 1 produces the following structure: \(^6\)

\[
\begin{align*}
\begin{bmatrix}
\tilde{B}_0^R(\theta_M^2) \\
\tilde{B}_0^R(\theta_M^1)
\end{bmatrix} Y_t &= \begin{bmatrix}
\tilde{B}_1^R(\theta_M^2) \\
\tilde{B}_1^R(\theta_M^1)
\end{bmatrix} Y_{t-1} + \ldots + \begin{bmatrix}
\tilde{B}_p^R(\theta_M^2) \\
\tilde{B}_p^R(\theta_M^1)
\end{bmatrix} Y_{t-p} + \epsilon_t + \\
+ \begin{bmatrix}
\tilde{A}_1^R(\theta_M^2) \\
\tilde{A}_1^R(\theta_M^1)
\end{bmatrix} \epsilon_{t-1} + \ldots + \begin{bmatrix}
\tilde{A}_p^R(\theta_M^2) \\
\tilde{A}_p^R(\theta_M^1)
\end{bmatrix} \epsilon_{t-p}
\end{align*}
\]

Equation (8) shows that the SVARMA-based counterfactual can correctly capture the impact of the authentic, DSGE-based counterfactual only if the policy parameters do not appear in the non-policy equations of the SVARMA. As we will see in Section 4 below, this only happens if the model’s structural characteristics are such that its solution is vector white noise. In all other cases, the SVARMA-based counterfactual fails to correctly capture the impact of the authentic, DSGE-based counterfactual, as the SVARMA-based policy switch only affects the SVARMA’s interest rate equation, leaving all of the other equations unaffected. This is the essence of the problem originally discussed by Sargent (1979).

2.3 A notational issue: structural VARs and structural VARMA\(^s\)

Although in general DSGE models possess a VARMA representation—as opposed to a pure VAR one—in what follows I will almost exclusively talk about ‘VARs’ and ‘structural VARs’. This, however, is uniquely as a shorthand. So the reader should keep in mind that every time I refer to ‘VARs’ and ‘structural VARs’, what I am referring to, in fact, is ‘VARMAs’ and ‘structural VARMAs’.

Let’s now turn to a simple illustration of the problem at hand based on a single stochastic simulation.

\(^5\)The reader may have a legitimate question: ‘Why is it the case that the first equations of (5) and (6) are the SVAR’s monetary rules?’. (I wish to thank Hashem Pesaran for originally posing me this question.) The answer is that, given the postulated ordering of the shocks—with the monetary policy shock, \(\epsilon_{R,t}\), having been ordered first in the vector \(\epsilon_t\)—the first equation of the SVAR representation of the DSGE model gets automatically identified as the SVAR’s monetary rule. Indeed, if we ignore the structural shocks, the individual equations of the SVAR representation

\[
\tilde{B}_0(\theta_M) Y_t = \tilde{B}_1(\theta_M) Y_{t-1} + \ldots + \tilde{B}_p(\theta_M) Y_{t-p} + \epsilon_t
\]

are indistinguishable from one another: only the specific structural shock which is ‘appended’ to each individual equation characterizes it.

\(^6\)The alternative counterfactual is just symmetrical.
2.4 A simple example of a theoretical SVAR-based counterfactual

All of the SVAR-based monetary policy counterfactuals analyzed in this paper are theoretical SVAR-based counterfactuals, that is, they have been performed based on the theoretical structural VAR representation implied by the underlying DSGE model, conditional on a specific parameterization. As such, they therefore represent the theoretically correct counterfactual experiment, abstracting from issues such as identification, small-sample bias, etc. The key reason for ignoring these issues is that, although they are crucial in ‘real-world applications’ of SVAR-based counterfactuals, within the present context they would only introduce an element of confusion in the entire exercise. Just to provide a simple example, it is well-known, from the work of Canova and Pina (2005), that in DSGE models the impacts at $t=0$ of all structural shocks is in general different from zero, thus implying that in the theoretical SVAR representation of a DSGE model the impact matrix of the structural shocks at $t=0$ will have, in general, no zero elements. Canova and Pina’s (2005) entire point was that, given this, using inertial restrictions in order to identify (monetary policy) shocks will necessarily distort the inference. Within the present context, this means that if (say), based on the theoretical reduced-form VAR representation of a DSGE model, I identified monetary policy shocks based on the Cholesky factor of the VAR’s theoretical covariance matrix of reduced-form innovations (or any other inertial scheme), I would inextricably be mixing two sources of distortion: first, the distortion originating from using an incorrect identification scheme; second, the distortion originating from the inability of the SVAR-based counterfactual to exactly replicate the authentic DSGE-based counterfactual. By abstracting from issues such as identification, small sample bias, etc., in this paper I exclusively focus on the second problem. This means (e.g.) that, throughout the entire paper, all SVAR-based counterfactuals are performed based on the true impact matrix $A_0$ of the structural shocks at $t=0$, as implied by the underlying DSGE model.

---

7 As I discuss below, when the DSGE model I am considering does not possess a pure VAR representation, but rather a VARMA one, the theoretical counterfactuals have been performed based on the VAR($\infty$) representation implied by the VARMA, which I approximate with a large number of lags (typically, at least 100).

8 Using sign restrictions would not solve the problem, since (among other things) sign restrictions are biased, in the sense that, in general, the expectation of the structural impact matrix computed based on sign restrictions is different from the true impact matrix (this can be trivially shown via Monte Carlo).

9 I do want to stress this as strongly as possible because, during presentations of this work at seminars and conferences, a surprisingly common misunderstanding was that I am here identifying monetary shocks via Cholesky, so that, based on Canova and Pina’s (2005) logic, it is not surprising that the resulting counterfactuals performs poorly. This misunderstanding likely originates from the fact that many papers in the literature do indeed use Cholesky to identify monetary shocks: this is however not what I am doing, as here everything is based on the true impact matrix $A_0$, as defined by the DSGE model.
In order to further clarify what I do in this paper, let me provide a simple illustration of a theoretical SVAR-based counterfactual based on one of the model I will use in what follows. The model is the one estimated by Lubik and Schorfheide (2004), which is described by the following equations: \(^{10}\)

\[
y_t = y_{t+1|t} - \tau (R_t - \pi_{t+1|t}) + g_t \\
\pi_t = \beta \pi_{t+1|t} + \kappa [y_t - z_t] \\
R_t = \rho R_{t-1} + (1 - \rho) [\phi_x \pi_t + \phi_y (y_t - z_t)] + \epsilon_{R,t}
\]

where the notation is obvious, \(R_t, \pi_t,\) and \(y_t\) are the nominal interest rate, inflation, and the output gap, respectively; \(g_t\) and \(z_t\) are AR(1) disturbances, \(g_t = \rho_s g_{t-1} + \tilde{g}_t,\) and \(z_t = \rho_z z_{t-1} + \tilde{z}_t;\) and \(\epsilon_{R,t}\) is white noise. It can be easily shown that the model has a reduced-form VAR(1) representation, which implies the following SVAR representation

\[
\begin{bmatrix}
R_t \\
\pi_t \\
y_t
\end{bmatrix}
= \begin{bmatrix}
B_{11} & B_{12} & B_{13} \\
B_{21} & B_{22} & B_{23} \\
B_{31} & B_{32} & B_{33}
\end{bmatrix}
\begin{bmatrix}
R_{t-1} \\
\pi_{t-1} \\
y_{t-1}
\end{bmatrix}
+ \begin{bmatrix}
A_{011} & A_{012} & A_{013} \\
A_{021} & A_{022} & A_{023} \\
A_{031} & A_{032} & A_{033}
\end{bmatrix}
\begin{bmatrix}
\epsilon_{R,t} \\
\tilde{g}_t \\
\tilde{z}_t
\end{bmatrix}
\]

(12)

where \(A_0\) is the impact matrix of the structural shocks. By defining the inverse of \(A_0\) as

\[
\hat{A}_0 = \begin{bmatrix}
\hat{A}_{011} & \hat{A}_{012} & \hat{A}_{013} \\
\hat{A}_{021} & \hat{A}_{022} & \hat{A}_{023} \\
\hat{A}_{031} & \hat{A}_{032} & \hat{A}_{033}
\end{bmatrix} = A_0^{-1}
\]

(13)
equation (12) therefore becomes

\[
\begin{bmatrix}
A_{011} & \hat{A}_{012} & \hat{A}_{013} \\
A_{021} & \hat{A}_{022} & \hat{A}_{023} \\
A_{031} & \hat{A}_{032} & \hat{A}_{033}
\end{bmatrix}
\begin{bmatrix}
R_t \\
\pi_t \\
y_t
\end{bmatrix}
= \begin{bmatrix}
\hat{B}_{11} & \hat{B}_{12} & \hat{B}_{13} \\
\hat{B}_{21} & \hat{B}_{22} & \hat{B}_{23} \\
\hat{B}_{31} & \hat{B}_{32} & \hat{B}_{33}
\end{bmatrix}
\begin{bmatrix}
R_{t-1} \\
\pi_{t-1} \\
y_{t-1}
\end{bmatrix}
+ \begin{bmatrix}
\epsilon_{R,t} \\
\tilde{g}_t \\
\tilde{z}_t
\end{bmatrix}
\]

(14)

where \(\hat{B} = \hat{A}_0 B = A_0^{-1} B.\)

Now that the structural shocks have been ‘separated’, so that each row has its own shock, the equations in (14) can be given a structural interpretation. In particular, the fact that the monetary policy shock shows up in the first equation in (14) identifies this equation as the SVAR’s monetary policy rule. Based on (14), the SVAR-based monetary policy counterfactual is therefore performed as follows.

Consider two alternative configurations of the policy parameters in (9)-(11), ‘A’ and ‘B’. In particular, assume, for the sake of the argument, that the two parameters’

\(^{10}\)In equations (9)-(11) I slightly changed Lubik and Schorfheide’s notation in order to put it in line with the notation I use in the rest of the paper.
implies that, the theoretical matrix formed in this way, based on the relevant theoretical objects—
that is: the theoretical properties of \( \tilde{\phi}^\pi \) (identifying small-sample issues, etc.) have been eliminated at the outset.

The theoretical SVAR-based counterfactual in which I impose the monetary rule \( B \) in the SVAR \( A \) is obtained by replacing, in the first row of (15), \( \tilde{A}_0^{11}, \tilde{A}^{12}, \tilde{A}^{13} \) with \( \tilde{A}_0^{B,11}, \tilde{A}_0^{B,12}, \tilde{A}_0^{B,13} \) and \( \tilde{B}_1^{A}, \tilde{B}_2^{A}, \tilde{B}_3^{A} \) with \( \tilde{B}_1^{B}, \tilde{B}_2^{B}, \tilde{B}_3^{B} \). Then, the resulting counterfactual SVAR, \( \tilde{A}_0^{C} Y_t = B^C Y_{t-1} + \epsilon_t \) (where \( 'C' \) stands for ‘counterfactual’) immediately gives us the representation for \( Y_t, Y_t = [\tilde{A}_0^{C}]^{-1} B^C Y_{t-1} + [\tilde{A}_0^{C}]^{-1} \epsilon_t = B^C Y_{t-1} + A_0^C \epsilon_t \). Based on this representation it is then possible to explore the theoretical properties of \( Y_t \) it implies, and to compare them to the theoretical properties implied by either the original representation \( A \), or—and this is going to be the focus of this paper—by the representation \( B \). In particular, if the SVAR-based policy counterfactual worked perfectly—in the sense of being able to perfectly mimic the impact on the economy of changing \( \phi_\pi \) from \( \phi_\pi^A \) to \( \phi_\pi^B \) in the underlying DSGE model (9)-(11)—then the counterfactual representation \( Y_t = B^C Y_{t-1} + A_0^C \epsilon_t \) would be identical to the representation \( Y_t = B^B Y_{t-1} + A_0^B \epsilon_t \). The entire point of the paper is that (i) this is never the case, and (ii) sometimes, the difference between the two representations is substantial.

In what follows all of the theoretical SVAR-based counterfactuals will be performed in this way, based on the relevant theoretical objects—that is: the theoretical matrix \( A_0 \) and the theoretical VAR’s coefficients matrices, \( B_1, B_2, \ldots, B_p \) (and, in the case of VARMAs, the MA coefficients matrices, \( A_1, A_2, \ldots, A_q \)), where \( p \) is the lag length (and \( q \) is the MA order), conditional on the DSGE model’s structural parameters, and alternative configurations of the parameters of the Taylor rule. This implies that, by construction, the only reason why the SVAR-based counterfactual cannot exactly replicate the authentic, DSGE-based counterfactual is because of the problem originally pointed out by Sargent (1979), since all other sources of distortion (identification, small-sample issues, etc.) have been eliminated at the outset.
2.5 A straightforward illustration based on a single stochastic simulation

Figure 1 shows results from a single stochastic simulation in which a standard New Keynesian model is fed the same set of structural shocks conditional on two alternative monetary policy rules, a ‘good’ (that is: comparatively more aggressively counter-inflationary) one, and a ‘bad’ (that is: comparatively less aggressively counter-inflationary) one. Results for the three variables of interest are reported in blue, for the ‘good’ policy regime, and in black, for the ‘bad’ policy regime, respectively. The authentic (that is: DSGE-based) policy counterfactual involves switching the two Taylor rules within the DSGE model, and then ‘rerunning history’ conditional on the same set of structural shocks: by definition, the outcome of such a switch implies switching the black and blue lines, so that what was ‘bad’ becomes ‘good’, and what was ‘good’ becomes ‘bad’. The SVAR-based policy counterfactual from the ‘bad’ to the ‘good’ regime, on the other hand, involves imposing the interest rate rule of the theoretical SVAR representation of the DSGE model conditional on the ‘good’ policy regime upon the theoretical SVAR for the ‘bad’ policy regime (exactly as described in the previous sub-section), and then ‘rerunning history’ based on the original structural shocks. The result from such exercise is shown, for either of the three variables, in red. By definition, if the SVAR-based counterfactual worked perfectly fine, the red lines would be identical to the blue lines. As the figure shows, however, this is definitely not the case: on the contrary, the SVAR-based counterfactual clearly fails to capture the truth as defined by the experiment we designed, with the red lines for inflation and the output gap, in particular, being remarkably close to the black lines (that is: to the ‘bad’ policy regime) rather than to the blue lines, as they should be; as for the interest rate, the red lines are basically ‘all over the place’, thus highlighting, once again, the unreliability—within the context of this simple experiment—of the SVAR-based counterfactual.

An obvious objection to these results is that they are based on a single stochastic simulation. Section 3, which presents results based on numerical methods, shows that the problem is a general one.

2.6 Why uniquely focusing on this particular type of counterfactual

This paper’s exclusive focus on a very specific type of policy counterfactual, involving changes in the parameters of the DSGE model’s monetary rule, may appear at first

11Since the issue discussed in this paper is a strictly conceptual one, details on the specific characteristics of the model used in this stochastic simulation are, in principle, irrelevant. To be precise, however, the model is the one estimated by Benati (2008b) for the post-WWII United States (estimates are reported in his Table XII), which is described by equations (1) and (17)-(18) below. The ‘good’ monetary policy is the one associated with Benati’s (2008) benchmark estimates, whereas the ‘bad’ one is obtained by setting $\rho=\phi_z=0$. 

11
sight as exceedingly narrow. In fact, it is not. *First*, although, for reasons of space, I ignore the analogous exercise involving changes in the parameters of a fiscal policy rule, this could be easily performed along the same lines as here. As a matter of logic, it appears as unlikely that whereas, as I show, SVAR-based monetary policy counterfactuals are often problematic, fiscal policy counterfactuals may possess a significantly superior performance. *Second*, a type of SVAR-based counterfactual often found in the literature involves ‘killing off’ identified monetary policy shocks. Since this counterfactual does not involve any manipulation of the SVAR’s monetary rule, it is not associated with any change in the VAR’s coefficient matrices. As a result, this counterfactual works perfectly well. A *third* type of counterfactual, less frequent, but sometimes found in the literature, involves imposing a path for the monetary policy rate alternative to the historical one. I do not consider this type of counterfactual for a very simple reason. A meaningful assessment of the reliability of any counterfactual can only be done conditional on a specific structural (DSGE) model. From a DSGE model’s perspective, an alternative path for the interest rate can result from either a particular sequence of monetary policy shocks, or from a change in the parameters of the monetary policy rule. In the former case, as I just mentioned, the SVAR-based counterfactual works perfectly fine. In the latter case, on the other hand, the analysis of this paper applies. Since any alternative path for the monetary policy rate results from either of the two possibilities (or from a combination of them), there is no need to separately analyze the reliability of this particular type of SVAR-based policy counterfactual.

### 2.7 ‘Once-and-forever’ versus recurrent change

An important point which is not tackled in the current version of the paper—but which I will address in the next version—is the following. As originally stressed by Sims in his debate with Sargent in the early 1980s, the Lucas critique, as usually formulated, pertains to unanticipated, ‘once-and-forever’ policy changes. The very fact that policy has changed in the past, however, implies that it can change in the future, so that economic agents should be thought as possessing an entire probability distribution over a set of possible future policy regimes. One way to formalize this is by modelling the DSGE model’s policy rules as Markov-switching processes, along the lines, e.g., of the recent work of Francesco Bianchi and Leonardo Melosi.

Under this respect, the following important point ought however to be stressed. For Sims’ argument to truly ‘bite’—so that, when modelling policy as a Markov process, the impact of the Lucas critique is not substantial—it ought to be the case that the transition matrix governing the evolution of the policy process is of a particular type. To fix ideas, let’s assume that the monetary rule switches between two regimes, and consider the following transition matrices:

\[
T_1 = \begin{bmatrix}
0.5 & 0.5 \\
0.5 & 0.5
\end{bmatrix} \quad
T_2 = \begin{bmatrix}
1-\epsilon & \epsilon \\
0 & 1
\end{bmatrix}
\]
with $\epsilon$ a small positive number in the neighborhood of zero. The transition matrices $T_1$ and $T_2$ encode two extreme, polar cases. Under $T_1$, agents’ expectation of the future is independent of the current state of the economy, with the result that the impact of any change in policy is minimised, because it only affects period $t$, whereas it has no impact on expectations about future policy regimes. This is an extreme example of Sims’ point. On the other hand, $T_2$ is very close to the notion of an unanticipated and permanent change in regime, so that the impact of the change in policy is maximised: this is in line with what I do in this paper. The key issue, for our purposes, is therefore which of the two cases is closer to reality. If reality is closer to $T_2$, then all of the results in this paper apply directly. If, on the other hand, the way policy regimes evolve is closer to the description provided by $T_1$, then this paper’s analysis is largely irrelevant. The evidence produced by Bianchi (2013) suggests that the world is likely closer to $T_2$ than to $T_1$: in particular, the modal estimates of the diagonal elements of the transition matrix of the Markov process for the Taylor rule are equal to 0.9861 and 0.9345, respectively, with the corresponding 90 per cent-coverage intervals from the posterior distribution being equal to [0.8746; 0.9925] and [0.8286; 0.9644], respectively. This points towards a substantial extent of persistence of the estimated policy regimes (e.g., the half-life of the first regime is slightly more than 12 years), and suggests that the results produced by an analysis, such as the one in the present paper, in which policy changes are modelled as ‘once-and-forever’ may not be too far away from the truth.

3 Illustrations Based on Numerical Methods

In this section I explore the reliability of SVAR-based counterfactuals based on three New Keynesian models, and conditional on grids of values for two parameters in the Taylor rule, the smoothing coefficient and the long-run coefficient on inflation. I start by exploring the ability of SVAR-based counterfactuals to correctly recover the impact of the authentic, DSGE-based counterfactuals on individual series’ characteristics (e.g., a series’ theoretical standard deviation). I then turn to relationships among variables, such as the one between inflation and the output gap.

Three general findings emerge from this analysis:

First, irrespective of the specific model I use, SVAR-based counterfactuals appear incapable of reliably capturing the macroeconomic impact of the authentic counterfactual. This holds true for both individual series’ characteristics, and the relationship among them.

Second, the problem appears to be especially severe at the low frequencies, and less so at the business-cycle frequencies, thus implying that SVAR-based counterfactuals might fare especially badly in assessing the role played by monetary policy in causing phenomena such as (e.g.) the Great Inflation or the Great Depression, which had both been characterised by prolonged and persistent fluctuations in the series of interest.

Third, the magnitude of the problem appears, in general, as non-negligible.
3.1 Three models

I consider the following three standard New Keynesian models, characterised by an increasing extent of complexity.

The first model is the one estimated by Lubik and Schorfheide (2004), which is described by equations (9)-(11). I calibrate the model based on Lubik and Schorfheide’s (2004) mean estimates for the post-1982 period as found in their Table 3.

The second model is the standard forward- and backward-looking model described by

\[
y_t = \gamma y_{t+1|t} + (1 - \gamma) y_{t-1} - \sigma (R_t - \pi_{t+1|t}) + \epsilon_{y,t}
\]

\[
\pi_t = \frac{\beta}{1 + \alpha \beta} \pi_{t+1|t} + \frac{\alpha}{1 + \alpha \beta} \pi_{t-1} + \kappa y_t + \epsilon_{\pi,t}
\]

where \(\gamma\) is the forward-looking component in the intertemporal IS curve, \(\alpha\) is price setters’ extent of indexation to past inflation, and everything else is the same as before. The model is closed with the monetary rule (1), and it is calibrated based on Benati’s (2008) modal estimates for the post-WWII United States as reported in his Table XII.

The third model is the one estimated by Smets and Wouters (2007) for the post-WWII United States, which I calibrate based on their modal estimates (see their tables 1A and 1B for the full-sample estimates, and table 5 for the sub-sample estimates).

3.2 Evidence for individual variables

3.2.1 Macroeconomic volatility

Exploring the role played by (improved) monetary policy in fostering the generalized fall in macroeconomic volatility associated with the Great Moderation has been, in recent years, one of the most prominent applications of SVAR-based policy counterfactuals. In this section I therefore investigate to which extent such counterfactuals can reliably capture the impact on economic volatility of changes in the parameters of the monetary policy rule in the underlying DSGE model. Although this section is narrowly focussed on individual series’ volatilities, analogous results on the reliability of SVAR-based policy counterfactuals pertain to alternative univariate properties of individual series, such as their persistence. These additional results are not reported here partly for reasons of space, and partly because the literature has mostly focused on the impact of changes in the SVAR’s monetary policy rule on macroeconomic volatility, but they are available upon request.

The experiment Figures 2-4 show, for either of the three models, results from the following experiment. For either model I compute its theoretical VAR and struc-
tural VAR\(^{12}\) representations conditional on the previously discussed benchmark estimates,\(^{13}\) which I label as VAR\(^B\) and SVAR\(^B\) respectively. The two representations imply certain benchmark values for the series’ theoretical standard deviations, which I collect in a vector labeled as STDs\(^B\). In line with the notation used in Section 2, I label the benchmark Taylor rule, and the benchmark interest rate equation in SVAR\(^B\), as Taylor\(^B\) and MonetaryRule\(^B\), respectively. I then consider grids of values for \(\rho\), from 0.4 to 0.95, and for \(\phi\), from 0.25 to 2.5 for the first two models, and for \(\rho\), from 0.41 to 0.91, and for \(\phi\), from 0.13 to 3.03 for Smets and Wouters’ model. On the other hand, I keep the other parameter(s) in the Taylor rule (\(\phi\), and, in the case of Smets and Wouters’ model, \(\Delta\phi\), that is: the coefficient on output growth) at the values implied by the benchmark estimates. For each combination of values of \(\rho\) and \(\phi\) in the grids, I solve the DSGE model,\(^{14}\) and I compute its theoretical VAR and SVAR representations, which I call VAR\(^A\) and SVAR\(^A\) respectively (where \(A\) stands for ‘alternative’), and the associated vector of implied theoretical standard deviations, STDs\(^A\). I label the alternative Taylor rule and the interest rate equation in SVAR\(^A\) as Taylor\(^A\) and MonetaryRule\(^A\), respectively.

By definition, switching Taylor\(^B\) and Taylor\(^A\) within the DSGE model (that is: performing the authentic counterfactual) inverts the two vectors of theoretical standard deviations, STDs\(^B\) and STDs\(^A\). If the SVAR-based counterfactual worked perfectly fine, we should therefore be able to obtain exactly the same result by switching MonetaryRule\(^B\) and MonetaryRule\(^A\). As I will now show, this is actually not the case. Let SVAR\(^C\)—where \(C\) stands for ‘counterfactual’—be the SVAR we obtain by imposing MonetaryRule\(^B\) withing SVAR\(^A\) (that is, we take away MonetaryRule\(^A\) and

---

\(^{12}\)To be precise, both model (9)-(11), and model (17), (18), (11), have a pure VAR representation. The model of Smets and Wouters (2007), on the other hand, does not admit a pure VAR representation, but rather a VARMA one. When I work with Smets and Wouters’ model I therefore work its VAR\(^\infty\) representation, which I approximate with a VAR with 100 lags (at lags greater than 100 the elements of the AR matrices of the VAR are of an order of magnitude smaller than \(10^{-8}\), and they can therefore be safely ignored).

\(^{13}\)For the model of Smets and Wouters (2007), I take as benchmark estimates the full-sample modal estimates as reported in their Tables 1A and 1B.

\(^{14}\)Given the wide ranges of values I consider for \(\rho\) and \(\phi\), some parameters’ combinations imply indeterminacy of the model solution. In these cases, I solve the model as in Lubik and Schorfheide (2004), picking the solution that they label as ‘continuity’. Further, in order to make the present exercise as transparent as possible, I set the standard deviation of the sunspot shock equal to zero. An important point to stress is that, in this way, I am essentially ‘stacking the cards against myself’, since (i) the presence of sunspot shocks under indeterminacy is, in principle, perfectly legitimate (on the contrary: their absence is open to question, and should be regarded as an extreme assumption); and (ii) sunspot shocks would inject additional volatility into the economy, and by ‘blowing up’ the diagonal elements of the covariance matrix of the theoretical VAR representation of the model, they would distort the results of the SVAR-based counterfactual. My choice of excluding sunspot shocks is motivated by the goal of making the results as transparent as possible (if I had also included sunspot shocks, it would have been impossible to determine the extent to which the results would have been driven by the presence of sunspots, as opposed to being due to the specific problem studied herein).
we replace it with MonetaryRule\textsuperscript{B}), and let VAR\textsuperscript{C} be its associated reduced-form VAR.\textsuperscript{15} VAR\textsuperscript{C} implies a vector of theoretical standard deviations for the series of interest, which I label as STDs\textsuperscript{C}. If the SVAR-based counterfactual were able to perfectly mimic the authentic DSGE-based counterfactual, for each possible combination of alternative values of \( \rho \) and \( \phi_{\pi} \) in the grids, we would have that STDs\textsuperscript{C}=STDs\textsuperscript{B}, so that for each individual variable \( i \) it would uniformly be STDs\textsuperscript{C}/STDs\textsuperscript{B}=1. On the other hand, the extent to which the SVAR-based counterfactual fails to replicate the impact of the authentic DSGE-based counterfactual is captured by how much, for each series \( i \), the ratio STDs\textsuperscript{C}/STDs\textsuperscript{B} deviates from one.

Figures 2-4 show, based on either of the three models, and for either of the model’s endogenous variables, the value taken by the ratio STDs\textsuperscript{C}/STDs\textsuperscript{B} as a function of alternative combinations of values of \( \rho \) and \( \phi_{\pi} \) in the grids. As the figures clearly illustrate, SVAR-based counterfactuals fail to replicate the outcome of the DSGE-based counterfactuals: based on either model, and for either series, the ratio STDs\textsuperscript{C}/STDs\textsuperscript{B} is, in general, different from one—sometimes quite markedly so—and it is (trivially) very close to one only for combinations of \( \rho \) and \( \phi_{\pi} \) which are sufficiently close to the benchmark estimates.

Further, the magnitude of the error made by SVAR-based counterfactuals is typically non-negligible, and it is rather often quite substantial. Focusing, e.g., on Figure 3, reporting results for the standard backward- and forward-looking New Keynesian model, for both inflation and the interest rate the counterfactual standard deviation is, for some combination of alternative values of \( \rho \) and \( \phi_{\pi} \), \textit{50 to 60 per cent higher} than it should be. Results for the model of Lubik and Schorfheide (2004) are even worse, although they should probably be given less weight due to the extreme, purely forward-looking nature of the model. Finally, results based on the model of Smets and Wouters (2007)—which should be given more weight, in consideration of the comparatively more sophisticated nature of the model, and of its medium-scale dimension—are comparatively bad for output, consumption, and the Federal Funds rate, whereas they are less distorted for variables such as investment, inflation, and the nominal wage. Since exploring the impact of changes in the monetary policy rule on output dynamics and volatility has been one of the most prominent applications of SVAR-based counterfactuals, it is worth spending a few words on the first panel of Figure 4. Smets and Wouters’ full-sample modal estimates for the two parameters are \( \phi_{\pi}=2.03 \) and \( \rho=0.81 \). As the figure shows, keeping \( \rho \) around the estimated value,\textsuperscript{16} and decreasing \( \phi_{\pi} \) to smaller and smaller values leads to a progressive deterioration of the performance of the SVAR-based counterfactual, which for values of \( \phi_{\pi} \) below

\textsuperscript{15} Qualitatively the same results are obtained based on the alternative SVAR-based counterfactual, in which we impose MonetaryRule\textsuperscript{A} within SVAR\textsuperscript{B}.

\textsuperscript{16} I focus on the experiment in which I change the value of \( \phi_{\pi} \), and I keep instead \( \rho \) at the value estimated by Smets and Wouters (2007), because evidence of changes in \( \rho \) from the pre-Volcker period to the period following the end of the Volcker stabilization is weak compared to the corresponding evidence of changes in \( \phi_{\pi} \).
1 reaches about 30 per cent (meaning that the standard deviation of counterfactual output generated by the SVAR-based counterfactual is 30 per cent greater than the corresponding standard deviation of counterfactual output generated by the DSGE-based counterfactual). Although Smets and Wouters’ estimates of $\phi_\pi$ for their two sub-samples (1966Q1–1979Q2 and 1984Q1–2004Q4) are very close (being equal to 1.65 and 1.77, respectively), several other papers, from Clarida, Gali, and Gertler (2000) to Lubik and Schorfheide (2004), have produced evidence of significant shifts in the coefficient on (expected) inflation in the FED’s monetary policy rule from the pre-Volcker period to the period following the end of the Volcker stabilization. As a result, the less-than-satisfactory performance of SVAR-based counterfactuals associated with values of $\phi_\pi$ below 1 should not be dismissed out of hand as empirically irrelevant.

**Implications**  
As previously mentioned, in recent years one of the most prominent applications of SVAR-based counterfactuals has been exploring the role played by monetary policy in fostering the generalised fall in macroeconomic volatility associated with the Great Moderation. Under this respect, the results shown in Figures 2-4 suggest that some caution in interpreting the results produced by these counterfactuals is warranted. The simple standard deviations of U.S. real GDP growth had indeed been equal to 4.3 per cent over the period 1960Q1-1983Q4, whereas it fell to 2.8 per cent over the period 1984Q1-2008Q2, corresponding to a decrease of -34.9 per cent. As we discussed in the previous paragraph, about 30 per cent might be regarded as a plausible upper bound for the distortion produced by the SVAR-based counterfactual associated to the two periods, compared to the DSGE-based counterfactual, for the standard deviation of output. Further, the direction of the distortion leads precisely to underestimating the impact of the change in $\phi_\pi$ on the standard deviation of output. As a result, it is not entirely implausible to envisage a situation in which monetary policy did indeed significantly contribute to the fall in output volatility from the first to the second period, but the distortions associated with the SVAR-based counterfactual leads to underestimating its effect, thus suggesting that the impact has instead been modest.29

---

17 Real GDP growth has been computed as 400 times the log-difference of GDPC96 (‘Real Gross Domestic Product, 3 Decimal, U.S. Department of Commerce: Bureau of Economic Analysis’), from FRED II at the St. Louis FED’s website.

18 In order to prevent results from being distorted by the increase in macroeconomic volatility associated with the financial crisis, I am here only considering the period up to the quarter immediately following the collapse of Lehman Brothers.

19 To be precise, the literature on the Great Moderation has consistently explored the role played by monetary policy *via* counterfactual simulations: I will illustrate the unreliability of such simulations in Section 5 below.
3.2.2 Results based on Fourier analysis

We now turn to results based on Fourier analysis, exploring the theoretical cross-spectral statistics between each individual variable as implied by the benchmark VAR, and the same variable as implied by the counterfactual VAR (that is, VAR^C). If the SVAR-based counterfactual worked fine, for each individual series i, and for each possible combination of alternative values of \( \rho \) and \( \phi_x \) in the grids, the theoretical gain and coherence between the series as implied by VAR^B and by VAR^C would be uniformly equal to one, whereas the theoretical phase angle and delay would be uniformly equal to zero.

Figures 5-7, 8-10, and 11-13 show, for each individual series, the average gain, coherence, and delay based on the three New Keynesian models.

Two main findings emerge from figures 5-13:

- consistent with the results discussed in the previous sub-section, the performance of SVAR-based counterfactuals is typically less-than-optimal, and it is often quite poor. In particular, whereas the coherence is, in general, quite high, and close to one, for most combinations of alternative values of \( \rho \) and \( \phi_x \), the gain is often quite off the mark. This implies that whereas the explanatory power of the counterfactual series for the benchmark series (or vice versa) is almost uniformly high, what the SVAR-based counterfactual misses, sometimes quite badly, is the proportionality (or scale) between the two series. This is in line with the results of the previous sub-section, where we saw how the SVAR-based counterfactual tends to miss the series’ volatilities. Finally, as Figures 7, 10, and 13 show, the SVAR-based counterfactual also introduces, in many instances, a clear phase shift between the benchmark and the counterfactual series, so that, in general, the counterfactual series is either leading or lagging the benchmark series.

- In general the magnitudes of the errors made by SVAR-based counterfactuals appear to be comparatively larger at the low frequencies, rather than at the business-cycle frequencies (this is especially clear for the gain and the delay statistics). This is a crucial point, because some of the phenomena investigated via this type of counterfactual had been characterised by prolonged and persistent fluctuations in the series of interest—that is, fluctuations pertaining precisely to the low frequencies. This is the case, for example, for the dramatic output contraction and deflation associated with the Great Depression, and for the prolonged and persistent inflation outburst associated with the Great Inflation.

---

20 For each frequency \( \omega \), the delay—which is measured in time units (e.g., quarters)—is defined as the ratio between phase angle and frequency (see e.g. Sargent (1987)).
21 I chose to show the delay, rather than the phase angle, because being expressed in quarters, rather than radians, it is easier to interpret. Results for the phase angle are however available upon request.
Focusing once again on Smets and Wouters’ model, whose results should be regarded as more plausible and informative than those based on the other two, simpler models, the first row of Figure 11 clearly shows how the SVAR-based counterfactual produces counterfactual series which are, at the low frequencies, a sometimes quite significantly ‘scaled up’ or ‘scaled down’ version of the benchmark series. This is especially apparent for output, investment, hours, and the FED Funds rate, whereas the distortion appears smaller for the remaining series. A comparison, for either series, between the results reported in the first and second row also clearly illustrates how the problem is uniformly more serious at the low frequencies than it is at the business-cycle frequencies. For output, for example, the average gain between the counterfactual and the benchmark series ranges between 0.736 and 1.689 at the low frequencies, whereas the corresponding values for the business-cycle frequency band are 0.706 and 0.882. For the Federal Funds rate the extent of comparative distortion at the low frequencies is even larger, with the average gain ranging between 0.447 and 1.700 at the low frequencies, and between 0.797 and 0.981 at the business-cycle frequencies. For other series the extent of comparative distortion is smaller (for inflation, for example, the gain ranges between 0.722 and 0.930 at the low frequencies, and between 0.897 and 0.969 at the low frequencies), but the overall pattern is quite clear. Although I do not have any explanation for why this should be the case, a comparatively greater extent of distortion at the low frequencies appears indeed to be a reasonably robust property of SVAR-based monetary policy counterfactuals, at least within the class of models considered herein.

A comparison between Figures 11 and 12 also provides a clear illustration of the second, previously mentioned point: although for the gain between the counterfactual and the actual series the extent of distortion is sometimes quite significant, especially at the low frequencies, the distortion appears to be smaller for the coherence. For output, for example, the average coherence between the counterfactual and the benchmark series ranges between 0.729 and 0.862 at the low frequencies, and between 0.715 and 0.837 at the business-cycle frequencies, whereas for investment it ranges between 0.942 and 0.970 at the low frequencies, and between 0.924 and 0.940 and the business-cycle frequencies.

Finally, Figure 13 shows how the time-shift between the counterfactual and the benchmark series introduced by the SVAR-based policy counterfactual is sometimes quite substantial. Focusing on the low frequencies, counterfactual investment leads the benchmark series by between 5 and 6 quarters, whereas for inflation the extent of the distortion is measured in years, with counterfactual inflation leading the benchmark series by about three years.

3.3 Evidence for macroeconomic relationships

Let’s now turn to bivariate macroeconomic relationships. For reasons of space, in this sub-section I only briefly discuss results based on the backard- and forward-
looking New Keynesian model. Results based on the other two models are however qualitatively in line with those discussed herein, and they are available upon request.

3.3.1 Unconditional correlations

Figure 14 shows, for each combination of alternative values of ρ and ϕₚ, the differences between the bivariate unconditional correlations implied by the counterfactual and benchmark VARs. In order to correctly interpret the information contained in the figure, it is important to keep in mind that unconditional correlations are bounded, by construction, between -1 and 1. If the SVAR-based counterfactual worked fine, such differences would uniformly be equal to zero: as the figure shows, however, this is not the case, with the SVAR-based counterfactual distorting key macroeconomic relationships (such as that between the nominal interest rate and inflation, or between inflation and the output gap), sometimes by non-negligible amounts.

3.3.2 Cross-spectral statistics extracted from the benchmark and the counterfactual VARs

Figures 15-17 show, for each combination of alternative values of ρ and ϕₚ, the differences between the average gain, coherence, and delay as implied by the counterfactual and benchmark VARs. As in Section 3.2.2, results are shown for both the low and the business-cycle frequencies. Once again, if the SVAR-based counterfactual worked fine, all of the objects reported in the figures would uniformly be equal to zero. Conceptually in line with the evidence discussed up until now, however,

first, the results produced by the SVAR-based policy counterfactual appear as, in general, potentially problematic, introducing a sometimes quite significant distortion in key macroeconomic relationships. This is the case, e.g., of the relationship between inflation and the nominal rate, for which—conceptually in line with the evidence reported in Figures 5, 8, and 11 for the two individual series—the gain between the series implied by the counterfactual VAR appears, for a wide range of parameters’ combinations, as quite different from that implied by the benchmark VAR. In particular, the extent of the distortion introduced by the SVAR-based counterfactual appears as sizeable not only at the low frequencies, but also at the business-cycle frequencies. On the other hand, distortions appear as less sizeable for the other two bivariate relationships, between inflation and the output gap, and between the interest rate and the output gap.

Second, conceptually in line with what we saw in Section 3.2.2, the evidence reported in Figure 16 suggests that SVAR-based counterfactuals distort bivariate macroeconomic relationships significantly less in terms of the coherence between the series. This implies that the explanatory power of one variable for the other remains largely undistorted by the SVAR-based counterfactual, whereas, as discussed in the previous paragraph, the strength of the relationship between the series is much more likely to get distorted, sometimes quite significantly so.
Third, results for the delay are qualitatively in line with those for the gain, with a non-negligible time-shift being sometimes introduced between the two series, and with the extent of the distortion being comparatively greater at the low than at the business-cycle frequencies.

4 Exploring the Impact of Individual Model Features

All of the results discussed up until now have been obtained conditional on specific parameterizations for the three DSGE models we have used, that is: conditional on (e.g.) specific extents of (i) forward- and backward-looking behaviour in the intertemporal IS and Phillips curves; (ii) serial correlation of the structural disturbances; etc.. This raises an obvious question: ‘Other things equal, how does the reliability of SVAR-based policy counterfactuals change with changes in some of the DSGE model’s key structural features, such as the extent of forward-lookingness of the intertemporal IS and Phillips curves?’ I tackle this issue next.

4.1 A purely forward-looking model

Consider the following model:

\[
\begin{align*}
R_t &= \rho R_{t-1} + (1 - \rho) \phi_\pi \pi_t + \phi_y y_t + \varepsilon_{R,t} \\
\pi_t &= \beta \pi_{t+1} + \kappa y_t + \epsilon_{\pi,t} \\
y_t &= y_{t+1} - \sigma (R_t - \pi_{t+1}) + \epsilon_{y,t}
\end{align*}
\]

(19) (20) (21)

with \( \varepsilon_{R,t} \sim \text{WN}(0, \sigma_R^2) \), \( \epsilon_{\pi,t} = \rho_{\pi} \epsilon_{\pi,t-1} + \tilde{\epsilon}_{\pi,t} \), and \( \epsilon_{y,t} = \rho_y \epsilon_{y,t-1} + \tilde{\epsilon}_{y,t} \).

4.1.1 White noise shocks and no interest rate smoothing

Assuming no reaction to the output gap on the part of the central bank, no interest rate smoothing, and serially uncorrelated shocks—that is, setting \( \phi_y = \rho = \rho_\pi = \rho_y = 0 \)—under determinacy model (19)-(21) has the following solution

\[
\begin{bmatrix}
R_t \\
\pi_t \\
y_t \\
y_t
\end{bmatrix} =
\begin{bmatrix}
1 & 1 - \rho & \kappa \phi_\pi \\
1 & \kappa \phi_\pi & \kappa \\
-\sigma & -\sigma & -\sigma \\
1 & 1 & 1
\end{bmatrix}
\begin{bmatrix}
\phi_\pi \\
\kappa \\
-\sigma \\
1
\end{bmatrix}
\begin{bmatrix}
\varepsilon_{R,t} \\
\varepsilon_{\pi,t} \\
\varepsilon_{y,t} \\
\epsilon_t
\end{bmatrix}
\]

(22)

with the system exhibiting no dynamics because (i) the model is purely forward-looking, and (ii) all the shocks are serially uncorrelated. Going from (22) to the
structural VAR representation of the model requires inverting the impact matrix $A_0$. After some algebra, we obtain:

$$
\begin{bmatrix}
1 & -\phi \pi & 0 \\
0 & 1 & -\kappa \\
\sigma & 0 & 1
\end{bmatrix}
A_0^{-1}

\begin{bmatrix}
R_t \\
\pi_t \\
y_t
\end{bmatrix}
=
\begin{bmatrix}
\bar{\epsilon}_{R,t} \\
\bar{\epsilon}_{\pi,t} \\
\bar{\epsilon}_{y,t}
\end{bmatrix}
$$

(23)

Equation (23) exhibits a crucial characteristic: the policy parameter, $\phi$, does not appear in the equations of the model’s SVAR representation other than the interest rate rule. As a consequence, based on our discussion of Section 2.2, we should logically expect the SVAR-based counterfactual to work perfectly. As the first row of Figure 18 shows, under determinacy this is indeed the case. Figure 18 shows the ratios between the series’ theoretical standard deviations implied by the counterfactual VAR (VAR$^C$, in the notation used in the previous sections) and the benchmark theoretical standard deviations implied by the benchmark VAR (VAR$^B$), where the benchmark is defined based on the mean estimates for the post-1982 period reported in Lubik and Schorfheide’s (2004) Table 3. (So to be clear, the exercise we are performing herein is exactly the same as we performed in Section 3.2.1.) In the first row of Figure 18 all of the parameters have been set equal to Lubik and Schorfheide’s estimates except for the autocorrelation of the shocks and the interest rate smoothing parameter, which have all been set to zero; in the second row only $\rho$ has been set to zero; and in the third row only the extent of autocorrelation of the shocks has been set to zero.

Several findings emerge from the first row of Figure 18. In particular, with white noise shocks and no interest rate smoothing, the SVAR-based counterfactual works perfectly—as expected—within the determinacy region, where the model’s solution is vector white noise. Under indeterminacy, on the other hand, the model’s solution is not vector white noise any longer, since—as shown by Lubik and Schorfheide (2003, 2004)—it depends on an additional unobserved and serially correlated state variable, so that now $\phi$ does not disappear from the SVAR’s equations other than the interest rate one. As a consequence, under indeterminacy the SVAR-based counterfactual fails, although by a very limited extent.

4.1.2 Autocorrelated shocks and no interest rate smoothing

Let’s now relax the extreme assumptions we have held so far under a single dimension. Specifically, whereas we still assume that $\phi_y=\rho=\rho_*=0$, we let the autocorrelation coefficient of the IS curve shock to be non zero, that is $\rho_y \neq 0$. After some algebra, it can be shown that, under determinacy, the model’s solution for the variables other

---


23See e.g. Lubik and Schorfheide (2004, equation 34, page 201).
than $R_t$—which is what matters for the present purposes—is given by

$$\begin{bmatrix} \pi_t \\ y_t \end{bmatrix} = \frac{1}{(1 + \kappa \sigma \phi_{\pi})} \begin{bmatrix} -\kappa \sigma & 1 \\ -\sigma & -\kappa \phi_{\pi} \end{bmatrix} \begin{bmatrix} \tilde{\epsilon}_{R,t} \\ \tilde{\epsilon}_{\pi,t} \end{bmatrix} + J(\phi_{\pi}) \begin{bmatrix} (1 - \lambda_1 \rho_y)^{-1} \Gamma_{12}(\phi_{\pi}) \\ (1 - \lambda_2 \rho_y)^{-1} \Gamma_{22}(\phi_{\pi}) \end{bmatrix} \epsilon_{y,t}$$

(24)

where $\lambda_1$ and $\lambda_2$ are the two roots of the characteristic polynomial of the relevant matrix in the forward-looking rational expectations solution of the model for $\pi_t$ and $y_t$, with

$$\lambda_{1,2} = (1 + \kappa \tau \phi_{\pi}) \left( 1 + \beta \kappa \phi_{\pi} \right) \pm \sqrt{(1 - \beta)^2 + \kappa^2 + \kappa \tau (2(1 + \beta) - 4\beta \phi_{\pi})}$$

(25)

the matrix $J(\phi_{\pi})$ collects the two eigenvectors associated with $\lambda_1$ and $\lambda_2$, and is given by

$$J(\phi_{\pi}) = \begin{bmatrix} \kappa \\ \lambda_1 (1 + \kappa \sigma \phi_{\pi}) - \beta - \kappa \sigma \\ \lambda_2 (1 + \kappa \sigma \phi_{\pi}) - \beta - \kappa \sigma \end{bmatrix}$$

(26)

and $\Gamma_{12}(\phi_{\pi})$ and $\Gamma_{22}(\phi_{\pi})$ are equal to

$$\Gamma_{12}(\phi_{\pi}) = \lambda_2 (1 + \kappa \sigma \phi_{\pi}) - \beta - \kappa \sigma - 1$$

(27)

$$\Gamma_{22}(\phi_{\pi}) = -\lambda_1 (1 + \kappa \sigma \phi_{\pi}) + \beta + \kappa \sigma + 1$$

(28)

From (24)-(28) we immediately have that

$$\pi_{t+1} = \rho_y \pi_t + \rho_y \kappa \sigma \tilde{\epsilon}_{R,t} \frac{(1 + \kappa \sigma \phi_{\pi})}{(1 + \kappa \sigma \phi_{\pi})}$$

(29)

$$y_{t+1} = \rho_y y_t + \rho_y \sigma \tilde{\epsilon}_{R,t} \frac{\phi_{\pi} \tilde{\epsilon}_{\pi,t}}{(1 + \kappa \sigma \phi_{\pi})}$$

(30)

so that,

- since $\rho_y \neq 0$, $\pi_{t+1}$ and $y_{t+1}$ on the right-hand side of (20)-(21) will be different from zero, and will therefore not drop out of the model’s solution; and

- crucially—as it clearly emerges from (29)-(30)—both $\pi_{t+1}$ and $y_{t+1}$ depend on the policy parameter, $\phi_{\pi}$. In turn, this implies that $\phi_{\pi}$ enters the equations for $\pi_t$ and $y_t$ in the SVAR representation of the model, which is the key necessary and sufficient condition for the outcomes of SVAR-based counterfactuals to deviate from those of DSGE-based counterfactuals.

Indeed, as the second row of Figure 18 shows, with autocorrelated shocks the SVAR-based counterfactual fails, and (trivially) it only works reasonably well when the alternative value of $\phi_{\pi}$ is close to the benchmark.

These results have been generated by setting both $\rho_{\pi}$ and $\rho_y$ to the values estimated by Lubik and Schorfheide (2004).

24
4.1.3 White noise shocks and interest rate smoothing

Finally, the bottom row of Figure 8 shows results for the case in which \( \rho \) has been set to the (non-zero) value estimated by Lubik and Schorfheide (2004), whereas \( \rho_x \) and \( \rho_y \) have been set to zero. Once again, the SVAR-based policy counterfactual fails to reliably capture the impact of the authentic, DSGE-based counterfactual. The explanation, once again, has to do with the fact that under all circumstances in which the model’s solution is not vector white noise, \( \pi_{t+1|t} \) and \( y_{t+1|t} \) on the right-hand side of (20)-(21) do not drop out of the model’s solution, with the result that the cross-equations restrictions implied by rational expectations cause the policy parameter to appear in all the equations of the SVAR representation of the model.

4.2 A backward- and forward-looking model

The fact that the crucial issue here is the (un)forecastability of \( \pi_t \) and \( y_t \) suggests that the problem should also appear in the presence of backward-looking components in the IS and Phillips curves. As Figure 19 shows, this is indeed the case. The results reported in the figure have been generated based on the standard backward- and forward-looking New Keynesian model (1), (17), (18). The model has been calibrated based on Benati’s (2008) modal estimates for the post-WWII United States as reported in his Table XII for all parameters except (i) the autoregressive parameters in the shocks’ processes, which, for the sake of transparency, have been set to zero;\(^{25}\) and (ii) \( \alpha \) and \( \gamma \), for which I have considered three sets of values:

(i) \( [\alpha \gamma]'=[0 1]' \), which implies that the IS and Phillips curves are purely forward-looking;

(ii) \( [\alpha \gamma]'=[0.5 0.5]' \), which implies that they are partly forward- and partly backward-looking;

(iii) \( [\alpha \gamma]'=[0.9 0.1]' \), which implies that they are very backward-looking.

For each combination of values of \( \alpha \) and \( \gamma \) I have performed this paper’s standard exercise conditional on grids of values for \( \rho \) and \( \phi_\pi \) as before. For each point in the grid, the benchmark Taylor rule is characterised by a value of \( \phi_\pi \) equal to Benati’s (2008) modal estimate, and by a value of \( \rho \) equal to the value taken by \( \rho \) in that point. So the results reported in Figure 19:

(i) are based on a set of benchmark values for \( \rho \), and

(ii) uniquely depend on the difference between the value taken by \( \phi_\pi \) and its benchmark value.

\(^{25}\)I say ‘for the sake of transparency’ for the following reason. By setting the autocorrelation of the shocks to zero, the extent of persistence of \( \pi_t \) and \( y_t \) uniquely depends on the values taken by the backward-looking components in the intertemporal IS and Phillips curve, so that interpreting the results is straightforward. On the other hand, if the autocorrelation of the shocks were different from zero persistence of \( \pi_t \) and \( y_t \) would not uniquely depend on such backward-looking components, and interpreting the results would therefore not be straightforward.
The reason for doing this is in order to explore the impact of $\phi_n$ on the reliability of the SVAR-based counterfactual conditional on several alternative benchmark values of $\rho$. As the figure shows,

- if the IS and Phillips curves are purely forward-looking, the problem is clearly apparent under indeterminacy, whereas under determinacy it only appears for comparatively high values of $\rho$ (this is especially apparent for the output gap). Consistent with the previous analysis based on the (modified) model of Lubik and Schorfheide (2004), if $\rho=0$, under determinacy the counterfactual works perfectly.

- If the IS and Phillips curves are not purely forward-looking, however, the problem is clearly always there, as $\pi_{t+1|t}$ and $y_{t+1|t}$ are not equal to zero, thus causing the policy parameter to appear in the equations for $\pi_t$ and $y_t$ in the SVAR representation of the model.

4.3 Implications

Two things ought to be stressed here.

First, a key theme emerging from the previous discussion is that the extent of (un)reliability of SVAR-based policy counterfactuals crucially hinges on the forecastability of inflation and the output gap (more generally, on the forecastability of the model’s endogenous variables other than the interest rate). On the face of it, this does not bode well for the reliability of SVAR-based policy counterfactuals, because

(i) although, as I have extensively discussed in Benati (2008b), high inflation persistence has historically been the exception, rather than the rule—with inflation having been uniformly very highly persistent only around the time of the Great Inflation episode—output gap estimates are instead typically characterized by a very high extent of persistence. Given the well-known, positive relationship between a series’ persistence and its extent of forecastability, this implies the we should logically expect SVAR-based policy counterfactuals to potentially perform poorly as a simple consequence of an easily verifiable and robust univariate property of a key macroeconomic time series. More generally, most macroeconomic time series are well known for being typically quite persistent, which just reinforces the point.

---

26 See, e.g. the Congressional Budget Office’s estimate of the U.S. output gap, the Office for Budget Responsibility’s estimate of the U.K. output gap, or the European Central Bank’s estimate of the Euro area output gap.

27 See e.g. Granger and Newbold (1986) and Barsky (1987). In general, the relationship between persistence and forecastability is one-for-one only within a univariate context, but it is still very strong even within a multivariate context.

28 To be fair, the output gap is unobserved. However, my point here is that all output gap estimates I am aware of can easily be verified to be very highly persistent ...
(ii) Given the link between a series’ persistence and its extent of forecastability, and the fact that both the Great Depression and the Great Inflation had been characterized by highly persistent fluctuations in inflation, whereas the former episode has also exhibited a very highly persistent fluctuation in output, this provides another reason why SVAR-based policy counterfactuals should logically be expected to fare especially poorly when applied to these two historical episodes.

Second, in practice the extent of forward- and backward-lookingness of the IS and Phillips curves (more generally, of the DSGE model’s structural equations) is unknown. This logically implies that reliability of SVAR-based counterfactuals cannot possibly be assumed, and can rather only be ascertained with a reasonable degree of confidence by estimating structural (DSGE) models. Eschewing estimation of structural macroeconomic models, and performing inference by imposing a minimal set of credible restrictions on the moving-average representation of the data is however the entire point of structural VAR analysis. As these results show, unfortunately, one important application of this methodology suffers from a key logical problem, as, in general, its reliability can only be ascertained via structural estimation.

5 How Relevant Is the Problem in Practice?

What is the practical relevance of the problem discussed in the present work? Specifically, what is the likely size of the error incurred by a researcher when performing a SVAR-based policy counterfactual, where the error is defined as the difference between the outcome of the SVAR-based counterfactual, and the outcome of the authentic counterfactual the researcher would have performed had (s)he known the true model of the economy? Providing a precise answer to this question is obviously impossible, as this would require knowledge of the true data generation process. A necessarily limited and tentative answer can however be provided (i) conditional on a specific estimated DSGE model, and (ii) for a specific counterfactual (e.g., ‘bringing Alan Greenspan back in the 1970s’).

In this section I therefore explore this issue based on Smets and Wouters’ (2007) model, and conditional on the sub-sample modal estimates reported in their Table 5 for the periods 1966Q1-1979Q2 and 1984Q1-2004Q4. Based on such estimates I ‘re-run history’ for the two sub-periods exactly as I did, based on a single stochastic simulation, in Section 2.2, the only difference being that there I performed the counterfactuals based on simulated data, whereas here I do it based on the actual historical data.29 Specifically,

- I perform the DSGE-based monetary policy counterfactuals by switching across sub-periods the values of the parameters of the Taylor rule, and then re-running

29 Specifically, I use Smets and Wouters’ (2007) original series, as found at the website of the American Economic Review.
history conditional on either period’s original set of structural shocks.\textsuperscript{30}

- I perform the SVAR-based monetary policy counterfactuals by switching across sub-periods the monetary policy rules in the theoretical structural VARMA\textsuperscript{31} representations implied by the two periods’ DSGE models, and then re-running history conditional on either period’s original set of structural shocks.

Figure 20 reports results from the counterfactual in which I switch the three key parameters of the monetary policy rule\textsuperscript{32} ($\rho$, $\phi_x$, and $\phi_y$) across periods. Specifically, the figure reports, for either period, the true series (in black), together with the counterfactual series produced by the DSGE-based and the SVAR-based counterfactuals (in blue and red, respectively). In order to correctly interpret the results, it is important to keep in mind that, exactly as in Section 2.2, if the SVAR-based counterfactual worked perfectly fine, the red lines would always be identical to the blue lines for every series and period.

Two main findings emerge from the figure:

(i) for four variables (inflation, nominal wage growth, the FED Funds rate, and hours) the DSGE-based counterfactual has virtually no impact on either series, and the SVAR-based counterfactual correctly captures this.

(ii) For the remaining three series (output, consumption, and investment) DSGE-based counterfactuals do indeed have non-negligible impacts on either series, but the corresponding SVAR-based counterfactuals typically fail to reliably capture such impacts. This holds for both types of counterfactuals (that is: both when I impose the first period’s monetary rule onto the second period’s estimated models, and \textit{vice versa}).

These results show, by example, that

\textit{first}, SVAR-based policy counterfactuals may perform poorly not only within simple, three-equations New Keynesian models, but also within realistic environments, such as Smets and Wouters’ (2007) estimated medium-scale model of the U.S. economy; and

\textit{second}, there does not seem to be a clear, specific pattern to the good or bad performance of SVAR-based counterfactuals. For example, as Figure 20 shows, they work well for four series, but work quite poorly for the remaining three, and there does not seem to be any obvious criteria which sets apart the series for which the SVAR-based counterfactual works well from those for which it performs poorly.\textsuperscript{33}

\textsuperscript{30}The two periods’ structural shocks are identified based on either period’s estimated structural model.

\textsuperscript{31}As previously pointed out, Smets and Wouters’ model does not possess a pure VAR representation, but rather a VARMA one. Once again, in practice I work with the VAR(∞) representation implied by the VARMA, which I approximate with a finite-order VAR with 100 lags.

\textsuperscript{32}The monetary policy rule estimated by Smets and Wouters (2007) also features a coefficient on output growth. In performing all of the counterfactuals reported in this section, I always keep this coefficient at its original sub-sample estimate.

\textsuperscript{33}In fact, the three series for which SVAR-based counterfactuals perform poorly are all national...
This suggests that, in general, the possibility that SVAR-based policy counterfactuals may perform quite poorly should be regarded as a realistic possibility.

6 The Consequences of the Possible Presence of Sunspot Shocks Under Indeterminacy

Finally, a further, important issue to be stressed is the following. All of the counterfactual exercises performed in this paper have been generated based on models in which, as previously pointed out, I allowed in principle for one-dimensional indeterminacy. What I did not allow, on the other hand, was for sunspot shocks under indeterminacy. The reason for this is straightforward: if I had allowed for sunspots under indeterminacy, I would have run into an identification problem. With $N$ reduced-form residuals from the VAR, and $N+1$ structural shocks, there would have been no way to identify the structural disturbances within the SVAR framework, which would have automatically put SVAR-based counterfactuals at a disadvantage, compared to the DSGE-based counterfactuals. In order to give SVARs their best shot, I therefore ruled out sunspots from the outset of the exercise. From a conceptual point of view, however, it is difficult to justify ruling out altogether sunspots under indeterminacy, as this is essentially a ‘corner solution’, and it is therefore more reasonable to assume that, under that regime, sunpots may play some role (that is, that their standard deviation is non-zero). If that’s the case, however, this is going to create two problems to SVAR analysis.

First, as I just mentioned, an identification problem, in the sense that it is impossible for the researcher to correctly identify all of the shocks.34 This implies that the identified structural shocks under indeterminacy will be unavoidably ‘contaminated’ by the sunspots.

Second, it will (further) distort the results of the SVAR-based counterfactual compared with those of the authentic, DSGE-based one. To fix ideas, suppose that, for the post-WWII U.S., we identify, in line with Clarida et al. (2000), indeterminacy for the Great Inflation period, and determinacy for the later period. This automatically implies that, when imposing the Taylor rule for the later period into the DSGE for the first period, one of the implications of such counterfactual will be to ‘kill off’ account components, but this may well just be a coincidence for this specific model and counterfactuals.

34 Quite obviously, this is based on the assumption that the number of structural disturbances other than the sunspot shock is equal to the number of observed variables, so that under determinacy the structural shocks are exactly recoverable. Although in practice this assumption may or may not be satisfied, it is important to stress that, within the present context, this is just an operational assumption: If I had assumed that structural shocks cannot be recovered even under determinacy (because the number of structural shocks other than the sunspot is greater than the number of observable variables), then it would not be possible to meaningfully compare SVAR- and DSGE-based counterfactuals under any circumstance).
the sunspots, thus automatically decreasing, *ceteris paribus*, macroeconomic volatility across the board. When performing instead the SVAR-based counterfactual, on the other hand, this—by the very logic of the exercise—*will not happen*, with the result that such counterfactual will necessarily understate the stabilising impact of the change in monetary policy.

7 Conclusions

Based on standard New Keynesian models I have shown that policy counterfactuals based on the theoretical structural VAR representations of the models fail to reliably capture the impact of changes in the parameters of the Taylor rule on the (reduced-form) properties of the economy. Based on estimated models for the Great Inflation and the most recent period, I have shown that, as a practical matter, the problem appears to be non-negligible. These results imply that the outcomes of SVAR-based policy counterfactuals should be regarded with caution, as their informativeness for the specific issue at hand—e.g., understanding the role played by monetary policy in exacerbating the Great Depression, causing the Great Inflation, or fostering the Great Moderation—is, in principle, open to question. Finally, I have argued that SVAR-based policy counterfactuals suffer from a crucial logical shortcoming: given that their reliability crucially depends on unknown structural characteristics of the underlying data generation process, reliability cannot simply be assumed, and can instead only be ascertained with a reasonable degree of confidence by estimating structural (DSGE) models.

References


Figure 1  A simple illustration based on a single stochastic simulation: results for the ‘bad’ and the ‘good’ policy regimes, and from the SVAR-based policy counterfactual from ‘bad’ to ‘good’
Figure 2  Ratios between the series' theoretical standard deviations from the SVAR-based policy counterfactual and the series' benchmark standard deviations (based on the model of Lubik and Schorfheide, 2004)
Figure 3  Ratios between the series’ theoretical standard deviations from the SVAR-based policy counterfactual and the series’ benchmark standard deviations (based on the model of Benati, 2008)
Figure 4  Ratios between the series’ theoretical standard deviations from the SVAR-based policy counterfactual and the series’ benchmark standard deviations (based on the model of Smets and Wouters, 2007)
Figure 5  Average gain between the series as implied by the benchmark VAR and the same series as implied by VAR produced by the SVAR-based policy counterfactual (based on the model of Lubik and Schorfheide, 2004)
Figure 6  Average coherence between the series as implied by the benchmark VAR and the same series as implied by VAR produced by the SVAR-based policy counterfactual (based on the model of Lubik and Schorfheide, 2004)
Figure 7: Average delay (in quarters) between the series as implied by the benchmark VAR and the same series as implied by VAR produced by the SVAR-based policy counterfactual (based on the model of Lubik and Schorfheide, 2004).
Figure 8 Average gain between the series as implied by the benchmark VAR and the same series as implied by VAR produced by the SVAR-based policy counterfactual (based on the model of Benati, 2008)
Figure 9  Average coherence between the series as implied by the benchmark VAR and the same series as implied by VAR produced by the SVAR-based policy counterfactual (based on the model of Benati, 2008)
Figure 10  Average delay (in quarters) between the series as implied by the benchmark VAR and the same series as implied by VAR produced by the SVAR-based policy counterfactual (based on the model of Benati, 2008)
Figure 11  Average gain between the series as implied by the benchmark VAR and the same series as implied by the VAR produced by the SVAR-based policy counterfactual (based on the model of Smets and Wouters, 2007)
Figure 12  Average coherence between the series as implied by the benchmark VAR and the same series as implied by the VAR produced by the SVAR-based policy counterfactual (based on the model of Smets and Wouters, 2007)
Figure 13  Average delay between the series as implied by the benchmark VAR and the same series as implied by the VAR produced by the SVAR-based policy counterfactual (based on the model of Smets and Wouters, 2007)
Figure 14 Exploring the distortions induced in macroeconomic relationships: difference between counterfactual and benchmark unconditional correlations (based on the model of Benati, 2008)
Figure 15 Exploring the distortions induced in macroeconomic relationships: difference between the counterfactual and benchmark gain (based on the model of Benati, 2008)
Figure 16 Exploring the distortions induced in macroeconomic relationships: difference between the counterfactual and benchmark coherence (based on the model of Benati, 2008)
Figure 17  Exploring the distortions induced in macroeconomic relationships: difference between the counterfactual and benchmark delay (based on the model of Benati, 2008)
Figure 18  Exploring the role played by the serial correlation of the shocks and by interest rate smoothing (based on the model of Lubik and Schorfheide, 2004)
Figure 19 Exploring the role played by the model’s extent of forward-lookingness (based on the standard New Keynesian backward- and forward-looking model with white noise shocks)
Figure 20  Rerunning U.S. post-WWII macroeconomic history taking the Smets and Wouters (2007) model as the truth: true series, and DSGE-based and SVAR-based theoretical counterfactual series (based on Smets and Wouters’ estimates for the 2 periods)