Testing for Spurious Dynamics in Structural Models with Applications to Monetary Policy *

Mikhail Dmitriev[1] Manoj Atolia[2]

FSU FSU

August 2021

Abstract

We propose a universal and straightforward test for validating assumptions in the structural models. Structural models impose a causal structure, take data as an input, and then produce exact structural parameters. We simulate the new data while breaking the original causal structure. We then feed the model the simulated data and then see whether it produces different results. If its conclusions are the same, then the models' implications are not sensitive to the underlying data, and the model fails the test. We then apply our test to the models analyzing monetary policy. We find out that simple SVARs successfully pass the test and can be used to identify monetary policy effects. On the other hand, DSGE models estimated via full-information methods such as Smets and Wouters (2007) fail the test and potentially force their conclusions on the data.

Keywords: VARs; SVARs; DSGE, monetary policy.

JEL Classification Numbers: C68, E44, E61.

*We would like to thank Kevin Cruz for superior research assitance, we also thank for valuable comments Yuriy Gorodnichenko, Alp Simsek, Veincenzo Quadrini, Zheng Liu, Jonathan Heathcote, Emil Verner, Urban Jermann, Simon Gichrist, Enrique Mendoza, Kartik Athreya, Jonathan Kreamer, Paul Beaumont, Xiaohong Chen, Anna Mikusheva. We also thank the seminar participants at Florida State University and Asian Meeting of Econometric Society.

[1]Corresponding author. Department of Economics. 255 Bellamy Building, Tallahassee 32306. Email: mdmitriev@fsu.edu. Web: mikhaildmitriev.org

[2]Department of Economics. 288A Bellamy Building, Tallahassee 32306. Email: matolia@fsu.edu. Web: https://myweb.fsu.edu/matolia/

Introduction

In the ideal world, it is easy to find the causal effect of policies. We need to simulate the world many times, generate arbitrary policy changes, see the results, and average them. Unfortunately, it is typically not possible since we live only in one actual world. Thus, economists often use structural models to make assumptions about the causal relationships between variables and then ensure that such models are consistent with the data. After that, we can see the effect of the policy under consideration in the model and then assume that this policy will have a similar impact on the world in the actual data. However, this exercise poses certain risks. For example, the models can be misspecified and have incorrect causal relationships but can still be consistent with the data. Dynamic general equilibrium models are at risk of being misspecified and fitting the data, as they allow for many shocks and match very persistent data. Thus, how can one check if the model is accurate when it is consistent with the data? In this paper, we propose such a test.

The idea is to simulate alternative data where policy under consideration is irrelevant and then feed this data to the model. If the model correctly detects that the policy does not affect other variables in the simulated data, we hold the model as accurate and effective. However, suppose the model insists that the policy under consideration has a strong effect on other variables. In that case, we conclude that the model is misspecified, forcing its policy conclusions on the data. We then consider two applications of the test for monetary policy, where we investigate whether monetary policy has a casual effect on the rest of the economy.

In the first application, we consider a simple recursive structural VAR exercise proposed by Stock and Watson (2001), where they investigate the effect of the Fed funds rate changes on unemployment and inflation. The authors estimate a recursive structural VAR with three variables and compute an impulse response to Fed funds rate shock. They find that an increase in the Fed funds rate shock increases unemployment and lowers inflation. Our goal is to see what the Stock and Watson (2001) model says if we feed it the simulated data whereby construction monetary policy does not affect the economy.

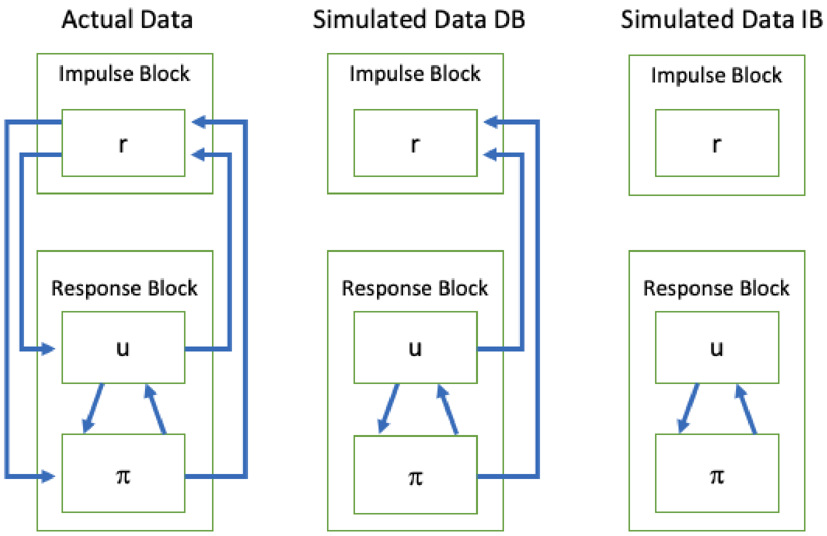

Figure 1: The Diagram of Causal Effects Between Interest Rate r, Unemployment u, and Inflation π in the Actual and Simulated Data

We split the data from Stock and Watson (2001) into an impulse block consisting of the interest rate and a We split the data from Stock and Watson (2001) into an impulse block consisting of the interest rate and a response block, which includes two other variables: inflation and unemployment. The diagrams in Figure 1 allow visualizing the relationship between these three variables. The left chart in Figure 1 reflects the actual data. Arrows show the causal effect of one variable on others, and we allow for all possible causal relationships. In the central chart, two left arrows from the interest rate to unemployment and inflation are missing. Their absence reflects that monetary policy does not have a causal effect on inflation and unemployment. However, the reverse is not valid, and inflation and unemployment still can have a causal impact on unemployment. We denote this setup as simulated data with dependent blocks or simply DB. Finally, the diagram on the right reflects the case where we remove all causal links between the blocks. In this case, the interest rate is independent of inflation and unemployment, while unemployment and inflation can have causal effects on each other as they stay within the block. We denote this environment as a simulation with independent blocks or simply IB.

We begin with the simulation of dependent blocks, displayed on the central chart in Figure 1. We create a simulated response block by reproducing the relationship between the variables within the actual data response block. In particular, we match with the actual data the volatilities of simulated inflation and simulated unemployment and their correlations with each other and their lags. We then simulate the impulse block, and we allow it to depend on the simulated response block. Thus, we denote this specification as dependent blocks or simply DB. For example, the interest rate can depend on simulated unemployment and simulated inflation, similar to the data and consistently with the Taylor rule. Also, we match covariance and autocorrelation properties of the actual interest rate. The simulated data's key feature is that monetary policy does not affect output and inflation. We run the simulation 300 times and then estimate the original SVAR model on each case's simulated data.

In the simulated data, we find that an unanticipated increase in the fed funds rate leads to a mean zero response of unemployment and inflation. This result is not surprising since the simulated Fed funds rate has no causal effect on simulated unemployment by construction. Moreover, the decrease in inflation in the actual data in response to the fed fund rate shock in the actual data significantly differs from the simulations. The increase in unemployment obtained in the actual data stays within the confidence interval for the simulated data response. Nevertheless, it is close to the boundaries and has a correct sign. Overall, the effect of the interest rate on inflation and unemployment in the actual data differs from the simulations. Thus, the model is capable of differentiating between two sets of data-generating assumptions.

We then run a more conservative simulation with independent blocks, displayed on the right chart in Figure 1. In the IB case, a simulated interest rate by construction does not react to simulated output and inflation and does not affect macroeconomic data. We still match covariance and autocorrelation properties within the impulse and response blocks and remove links between them. Recursive SVAR estimation still performs well, and our results are largely similar to the estimation with dependent blocks described above.

As a result, we conclude that the SVAR recursive structure used by Stock and Watson (2001) allows differentiating between alternative causal structures and can be successfully used for testing the effects of monetary policy.

In the second application, we test a traditional dynamic general equilibrium model developed by Smets and Wouters (2007). They estimate a medium-scale New Keynesian model on the US macroeconomic data and determine that such a model has good predictive power and can be used to assess the monetary policy's effect. We split the data from Smets and Wouters (2007) into the impulse block, which consists of the interest rate, and the response block, which includes six other variables: log hours worked, as well as the growth rate of GDP, consumption, investment, real wage, and GDP deflator. We simulate the response block by reproducing the covariance and autocorrelation of the actual response block. We then move to the simulation of the impulse block. We allow the simulated interest rate to depend on simulated output and simulated inflation similar to the data and match covariance and autocorrelation properties of the actual interest rate. The simulated data's key feature is the absence of any effect of the simulated monetary policy on any of the six other variables. We then run a Bayesian estimation on the simulated data and find that monetary policy strongly affects output and inflation, even though the opposite is true. The impulse response obtained under actual data are indistinguishable from the results obtained under simulated data.

We then run a more conservative IB test, where the impulse and response blocks are entirely independent of each other. The simulated interest rate does not react to simulated output and inflation and does not affect any of the six macroeconomic series. We still match covariance and autocorrelation properties within the impulse and response blocks and remove links between them. The result stays the same. The estimated Bayesian model implies strong effects of the monetary policy.

Why is it the case that the Smets and Wouters (2007) find strong monetary policy effects even when it is irrelevant? In reality, monetary policy might fail for many reasons, money neutrality, failure of the central bank to communicate, problems with the transmission of lower fed funds rate into lower mortgage rates, and lack of desire from households and firms to borrow and spend. In the model, the monetary policy fails to have a substantial effect only if prices are flexible. However, the actual volatility of prices is small; and the model concludes that the monetary policy has strong effects for stable observed inflation even when the central banks can do little to help the economy. Are there any other indicators that the exercise is failing? If the monetary policy is irrelevant, but the effect of monetary shocks is assumed to be strong, we still should observe these shocks' small role in variance decomposition, which is precisely the case in Smets and Wouters (2007). In their paper, monetary policy shocks explain only a small percentage of output variation despite having a potentially strong effect. Thus, variance decomposition plays a complementary role to our test in detecting false positives.

We do not interpret our results as evidence against full-information methods of estimation of structural models. However, we object to the excessive focus of such models and methods on consistency with the data. Our exercise points that it is crucial to distinguish between the cases when the model fails and when it succeeds. The traditional New Keynesian model allows very few possible monetary policy scenarios to have weak effects. However, it often allows a large number of possibilities when it can match the data. Consequently, it fits the data even when if its policy conclusions are incorrect. Regarding monetary policy, newer generations of models should focus on the transmission mechanism, including commercial banks' excess reserves, collateral requirements, etc. In this case, the monetary policy cannot stimulate the economy either because of the transmission mechanism or flexible prices.

Related Literature

Our paper relates to Chari et al. (2005), who criticize SVAR conclusions based on their conclusions drawn from the data simulated under the real business cycle theory. In a certain way, we generalize their approach by applying it to any structural methodology and practically any null hypothesis. Our key advantage is that we do not need to tinker with particular structural assumptions and details of the model; we simulate alternative inputs. In terms of SVARs, we relate to the debate on the usefulness of SVARs by Cooley and Dwyer (1998) and Christiano et al. (2007). Our advantage is that we generalize this debate by being agnostic about the particular nature of the models.

Concerning dynamic general equilibrium models and a New Keynesian model, our work relates to the criticism of New Keynesian models by Chari et al. (2008), and its estimation by Christiano et al. (2005) and others. We differ from this literature by not taking any particular stance for or against any specific model. Instead, we simulate alternative data by tweaking the assumptions about the policy under consideration, feed it to the model, and then investigate whether conclusions of the model are consistent with the assumptions used in the simulation of the data.

The General Methodology

Let's consider a stationary dataset y_t consisting of two blocks of variables y_t^1 and y_t^2 so that $y_t = [y_t^1, y_t^2]$. We are interested in the causal links from variables in block y_t^1 on y_t^2. The set y_t^1 can include the fed funds rate, government purchases, capital requirements, or other variables. We are going to refer to these variables as a group by impulse block. We denote the remaining variables in y_t^2 as the response block. The response block typically includes variables such as GDP, consumption, investment, trade balance, and others. Let us also assume that a model f implies some causal relationship between the two blocks. Formally, the model can be represented as a system of equations $f(y_t^1, y_t^2, \theta, e_t) = 0$, where θ is a set of models parameters, and e_t is a set of exogenous variables, such as measurement error or possibly other exogenous shocks. Such a model can be a structural DSGE model, SVAR model, or even a reduced form linear regression that potentially links two blocks. Under the null hypothesis $\theta = \theta_0$, there is no causal relationship between the two blocks. In the typical estimation exercise, we identify a mean/mode value of θ^* and the confidence interval $[\bar{\theta}, \underline{\theta}]$. If this interval excludes θ_0, then the null hypothesis is rejected, and we assume that the causal effect from y_t^1 to y_t^2 is consistent with the data.

Our goal is to evaluate the model f. , We simulate dataset $\tilde{y}_t$, also consisting of the impulse block $\tilde{y}_t^1$ and the response block $\tilde{y}_t^2$. The simulated data should be similar to the actual data with all respect but the impulse block's effect on the response block. In particular, **our null hypothesis is that the impulse block does not affect the response block.**

We do not generate the model under θ_0. Our simulation is model-free, agnostic, and relies only on the properties of data y_t. We build the data $\tilde{y}_t^1$ and $\tilde{y}_t^2$, and by construction $\tilde{y}_t^1$ has no causal effect on $\tilde{y}_t^2$. We run the simulation many times, and in each iteration i, we obtain a value $\tilde{\theta}_i$ from the estimation of the original f model on the new simulated data. Having a set of $\tilde{\theta}_i$ allows building a new confidence interval for the estimated parameters $[\bar{\tilde{\theta}}; \underline{\tilde{\theta}}]$. We now check whether the value θ^* obtained in the estimation of the model on the actual data belong to the new confidence interval $\theta^* \in [\bar{\tilde{\theta}}; \underline{\tilde{\theta}}]$. If it does, we claim that the model cannot identify causal links in the data as it suffers from false positives or type I error. We also check whether the null hypothesis θ_0 belongs to the interval $\theta_0 \in [\bar{\tilde{\theta}}; \underline{\tilde{\theta}}]$. If it does not, the model has too tight restrictions that force rejecting the null hypothesis.

To proceed with the estimation, we do not need to understand the functional relationships in the model f. All we need for the model is to spit out the particular set of parameters θ^* once we feed the data to the model. After that, we simulate the data many times, estimate the model each time i and obtain the value of $\tilde{\theta}_i$. Having the set of $\tilde{\theta}_i$ allows us to obtain the necessary confidence interval $[\bar{\tilde{\theta}}; \underline{\tilde{\theta}}]$ and check whether the original value θ^* is a part of this interval. Following the steps described below, we construct the simulated data consistent with the null.

1. *Estimate a regression for the response block in the actual data:*

$$y_t^2 = \hat{\Lambda}_2 y_{t-1}^2 + e_t^2. \tag{1}$$

Then after obtaining the residual term e_t^2, construct the variance V_2:

$$\hat{V}_2 = \frac{1}{T-k} \sum e_t^2 (e_t^2)'. \tag{2}$$

Build η_2 using Cholesky decomposition so that

$$\eta_2 \eta_2' = \hat{V}_2. \tag{3}$$

Matrix η_2 allows having a similar variance of the residual in the simulations relative to the actual data. In this step, we evaluate the relationships within the response block. In the next step, we are going to match them in the simulated data.

2. *Build the simulated series for the response block*

$$\tilde{y}_t^2 = \hat{\Lambda}_2 \tilde{y}_{t-1}^2 + \eta_2 \epsilon_t^2, \tag{4}$$

where ϵ_t^2 are standard white noises, independent across time and variables.

In this step, we simulate the response block y_t^2 in equation (4) to not be affected by the impulse block. It depends only on exogenous shocks, as well as its own lagged values.

3. *Estimate the relationship in the actual data for the impulse block.*

We allow for two possibilities here. First, while we assume that the impulse block does not have a causal effect on the response block, we allow the impulse block to depend on the response block. We refer to this case as dependent blocks or DB. For the second possibility, we simulate the impulse block to be completely independent of the response block. We refer to this case as independent blocks or IB. The mathematical relationships for both DB and IB cases are outlined below.

(a) *The impulse block values can depend on contemporaneous response block values (DB):*

$$\text{Dependent Blocks}: \quad y_t^1 = \hat{\Lambda}_{11} y_{t-1}^1 + \hat{\Lambda}_{12} y_t^2 + e_t^1. \tag{5a}$$

In this case, the variables in the simulated block y_t^1 depend not only on their lagged values y_{t-1}^1 and exogenous shocks e_t^1, but contemporaneous variables from the response block y_t^2 as well.

(b) *The impulse block is completely independent of the response block (IB):*

$$\text{Independent Blocks}: \quad y_t^1 = \hat{\Lambda}_{11} y_{t-1}^1 + e_t^1. \tag{5b}$$

In this case, the variables in the simulated block depend only on their lagged values y_{t-1}^1 and exogenous shocks e_t^1.

We then run the regression (5a) or (5b) on the actual data and obtain the residuals e_t^1, In the next step, we construct a variance of e_t^1:

$$\hat{V}_1 = \frac{1}{T-k} \sum e_t^1 (e_t^1)'. \tag{6}$$

We then decompose the variance $\hat{V}_1$ using Cholesky decomposition into the product of η_1 and its transpose so that

$$\eta_1 \eta_1' = \hat{V}_1. \tag{7}$$

4. *Build the simulated series for the impulse block as dependent block (DB) in (8a) [1] or as an independent block (IB) in (8b) so that*

$$\tilde{y}_t^1 = \hat{\Lambda}_{11} \tilde{y}_{t-1}^1 + \hat{\Lambda}_{12} \tilde{y}_t^2 + \eta_1 \epsilon_t^1, \tag{8a}$$

$$\tilde{y}_t^1 = \hat{\Lambda}_{11} \tilde{y}_{t-1}^1 + \eta_1 \epsilon_t^1, \tag{8b}$$

where ϵ_t^1 is the standard white noise, independent across time.

5. *In each simulation i, apply the model f to the simulated data, estimate structural parameters θ_i and determine whether shocks to $\tilde{y}_t^1$ cause the effect on $\tilde{y}_t^2$. If the impact is significant, then the model generates a "false positive" or type one error.*

[1] We credit Jonathan Kreamer for suggesting it during one of the Macro workshops at FSU.

6. *We repeat the steps above for N times and establish the mean effect of the impulse block on the response block* $\bar{\tilde{\theta}} = \sum_{i=1}^{N} \tilde{\theta}_i$ *and the confidence intervals* $[\bar{\tilde{\theta}}; \underline{\tilde{\theta}}]$.

7. *Determine whether the value of estimated parameters on the actual data* θ^* *is a part of the confidence interval* $[\bar{\tilde{\theta}}; \underline{\tilde{\theta}}]$. *If it is, then the model is unable to identify the causal relationship in the data. Also, determine whether the parameter* θ_0*, under which there is no relationship between the* $\tilde{y}_t^1$ *and* $\tilde{y}_t^2$ *according to the structural model, is a part of the interval* $[\bar{\tilde{\theta}}; \underline{\tilde{\theta}}]$. *If it is not, then the restrictions of the model might be too tight.*

Application of the steps listed above does not require a deep understanding of the model tested. The model takes the raw data as input and then generates a specific set of parameters. The methodology proposed here modifies the input data without touching the mechanics of the model. It then compares the model's conclusions with the assumptions that guide the simulation of inputs. However, one of the consequences is that the methodology is silent on the particular flaws of the model if the latter generates type one error.

One way we partly address the degree of misspecification is by having two versions of the test. In a less conservative version of the test, we allow the impulse block to depend on the response block while not the other way around. In this case, having some interdependency between impulse and response blocks allows it easier for the model to generate the results that the impulse block can have a causality on the response block. Thus, we have a more conservative version of the test; we break all links between the impulse and the response block in the simulated data and then examine whether the model generates corresponding conclusions. If the model still insists on the causal relationship between the blocks, the case of "false reject" is more severe.

We consider applying the methodology to the structural VAR estimation and Bayesian DSGE estimation in the following sections.

Example 1. A Toy VAR Model from Stock and Watson (2001).

Summary of the Original Exercise

The canonical structural vector autoregression model considers the relationship between economic activity and monetary policy. In the classical example described below, we consider the relationship between inflation π_t, unemployment u_t and a Fed funds rate R_t following Stock and Watson (2001). In particular, we have a recursive VAR ordered as 1) inflation, 2) the unemployment rate, and 3) the interest rate:

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + A_3 Y_{t-3} + A_4 Y_{t-4} + C u_t, \tag{9}$$

where Y_t is a vector of $[\pi_t, u_t, R_t]$, and C is a lower triangular matrix, while u_t is a vector of uncorrelated exogenous disturbances. Consistently with Stock and Watson (2001), our VAR allows for four lags of each variable and covers the quarterly data from 1960:I-2000:IV.

The particular recursive arrangements of the variables in a VAR are not arbitrary. In the original paper, the authors run a Granger-causality test and show inflation predicting unemployment. Still, the fed funds rate does not help with predicting unemployment. Unemployment helps to predict inflation, while the Fed funds rate does not improve the forecast of inflation. Finally, both inflation and the unemployment rate improve the prediction of the Fed funds rate. These results are consistent with the chosen recursive arrangement. The shock to inflation affects both the contemporaneous Fed funds rate and unemployment. The unemployment shock affects the current inflation rate, and the Fed funds rate shock affects both unemployment and inflation with a lag.

In Figure 2 below, we replicate the original exercise by generating the impulse response after estimating a recursive structural VAR:

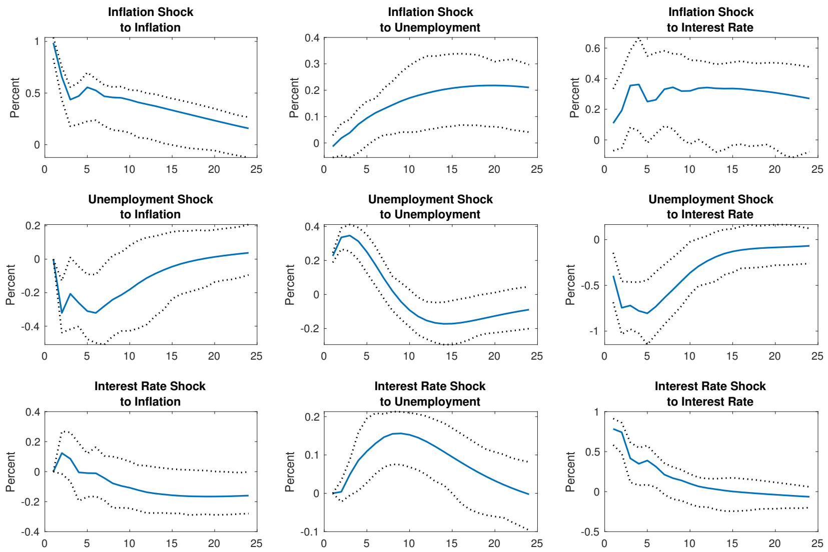

Figure 2: IRF with SVAR Estimation on the Actual Data

These results are intuitive. We interpret inflation shock as a negative supply shock, reported in the first row of Table 1. As expected, this shock leads to an increase in inflation, unemployment, and interest rate. On the other hand, an increase in unemployment reflects a negative demand shock. As expected, the second row of Table 1 demonstrates that an increase in unemployment leads to lower inflation, higher unemployment, and lower interest rate. The last three subplots in the third row describe the effect of a rise in the interest rate. As a negative demand shock, it leads to lower inflation, higher unemployment, and higher interest rate. Overall, these results are intuitive and do not raise any suspicions about the statistical procedure.

We are primarily interested in the third row, which addresses the effect of monetary policy on the economy. In the next section, we simulate the new data assuming that the monetary policy is irrelevant.

Simulating the Alternative Data

In this section, we simulate the data under the assumption that monetary policy does not affect the economy. We want the simulated data to be similar to the actual data in all aspects except the effect of monetary policy. While in the actual data, we are agnostic about monetary policy's impact, we know that monetary policy is irrelevant in the simulation. More formally, we state the null hypothesis below:

Null Hypothesis 1 *Monetary policy does not affect the economy.*

Consistently with the general methodology, we separate the actual data into two blocks. The impulse block y_t^1 consists only from the Fed funds rate or $y_t^1 = [R_t]$. The response block y_t^2 consists of both inflation and unemployment or $y_t^2 = [\pi_t, u_t]$. We are then ready to apply the general methodology.

1. We estimate the AR(1) process for the response block following (1), consisting of inflation and unemployment. Our goal here is to preserve the relationship between inflation, unemployment, and their lagged values in the simulations.

2. Here, we build the simulated values of inflation and unemployment using pseudo-random generators according to equation (4).

3. We estimate the relationship in the actual data between the Fed funds rate, its lags, and other variables. Consistently with the general methodology, we allow for two possibilities. According to the first dependent or DB case described in (5a), we allow the fed funds rate to depend on contemporaneous unemployment and inflation. In the second independent block or IB case, we allow the fed funds rate to be completely independent of other variables and follow an AR(1) process according to equation (5b).

4. We simulate the data for the Fed funds rate for both cases. In the first DB case, where the interest rate depends on inflation and unemployment, the data is simulated according to (8a). In the second IB case, the fed funds rate is simulated according to (8b).

5. We apply recursive VAR estimation to the simulated Fed funds rate, inflation, and unemployment and build the impulse response function in each iteration.

6. We repeat the procedure N=300 times and establish the mean impulse response for the identified shocks in the simulated data and the confidence intervals for the simulated impulse responses.

7. We compare the impulse response from the actual data with the confidence intervals of the simulated data impulse responses.

We then proceed to the simulation with dependent blocks displayed in Figure 3. In this environment, the fed funds rate does not have any causal effect on the unemployment rate or inflation. The reverse is not valid, and the unemployment rate and inflation might have a causal impact on the Fed funds rate. Our central focus is the third row of the figure. We find that a positive interest rate shock has no effect on unemployment or inflation in the simulated data. This result is entirely consistent with the data generating process, where the interest rate has no causal impact on neither inflation nor unemployment. This is not the case in the actual data. The increase in the interest rate leads to lower inflation after ten quarters in a statistically significant way relative to the simulated data. While the interest rate also leads to a higher response of unemployment in the actual data, the response stays within the confidence interval of the simulated. Overall, the effect of interest rate shock on unemployment and inflation in the actual data differs from the simulated data.

The first row of figure three also shows the startling difference between simulated and actual data. Inflation shock seems to move inflation, unemployment, and interest rate stronger than in the simulated data in a statistically significant way. Moreover, higher inflation shock causes higher interest rates in the actual data, which is consistent with the economic intuition. In the simulated data, however, a positive inflation shock leads to lower interest rates. Thus, we can conclude that the inflation shock has sharply different effects relative to the simulated data in the actual data.

The second row of figure three demonstrates the effect of the unemployment or demand shock. Again, we see the sharp difference between actual and simulated data. While in the actual data, higher unemployment leads to lower inflation, in the simulated data, we see inflation increasing with higher unemployment. Moreover, while higher unemployment leads to a lower interest rate in the actual data, in the simulations, we observe higher interest rates in response to unemployment. As before, the results in the simulations seem to be inconsistent with economic intuition. In all diagrams for the second row in the figure, we observe statistically significant differences between simulated and actual data.

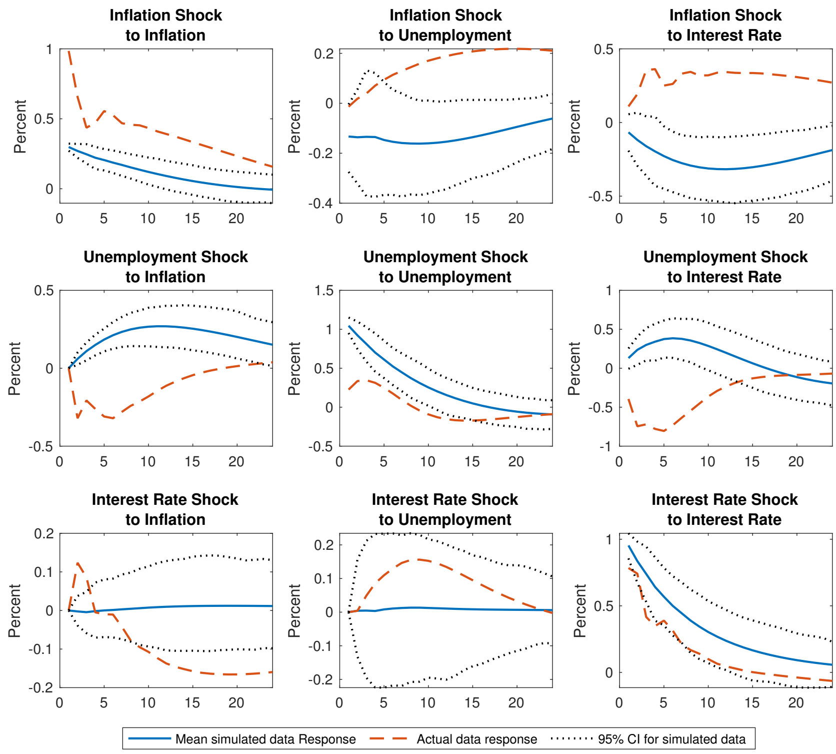

Figure 3: SVAR Estimation for the Simulated Data with Dependent Blocks

Figure 4 displays the impulse response functions for estimating the recursive structural vector autoregression for the simulated data with independent blocks. We preserve the relationship between inflation and unemployment in the simulated data since these variables both enter the response block. On the other hand, simulated Fed funds rate data is entirely orthogonal to the simulated inflation and simulated unemployment. Overall, the simulated Fed funds rate's persistence and volatility are similar to the actual data, while its relationships with simulated unemployment and inflation are not.

In Figure 4, we are primarily interested in the third row, which shows how a positive shock to the fed funds rate affects itself, unemployment, and inflation. As expected, an increase in the interest rate leads to a mean zero response in unemployment and inflation. Like in figure 3, the response of inflation in the actual data is lower than in the simulated data in a statistically significant way. Also, similar to figure 3, a higher interest rate leads to higher unemployment, which is consistent with economic intuition, but the results stay within the simulated data's confidence interval.

Like in figure 3, higher inflation shock leads to higher unemployment in the actual data and lower unemployment in the simulated data, delivering statistically different results. Moreover, positive inflation shock leads to a mean zero response in the interest rate, which is consistent with the data generating process, as the interest rate is entirely orthogonal to inflation and unemployment. In contrast, in the actual data, higher inflation leads to a higher interest rate in a statistically significant way relative to the simulated data, which is also consistent with economic intuition.

Finally, we turn our look to the second row, which shows the effect of an unemployment shock. Higher unemployment leads to lower inflation in the actual data and higher unemployment in the simulated data, similar to the results displayed in figure three. This difference is statistically significant. As expected, an unemployment shock on average leads to a zero mean effect on the interest rate, as the latter is entirely independent of unemployment. In the actual data, however, a positive shock to unemployment leads to a lower interest rate. As before, this result is statistically significant and consistent with economic intuition.

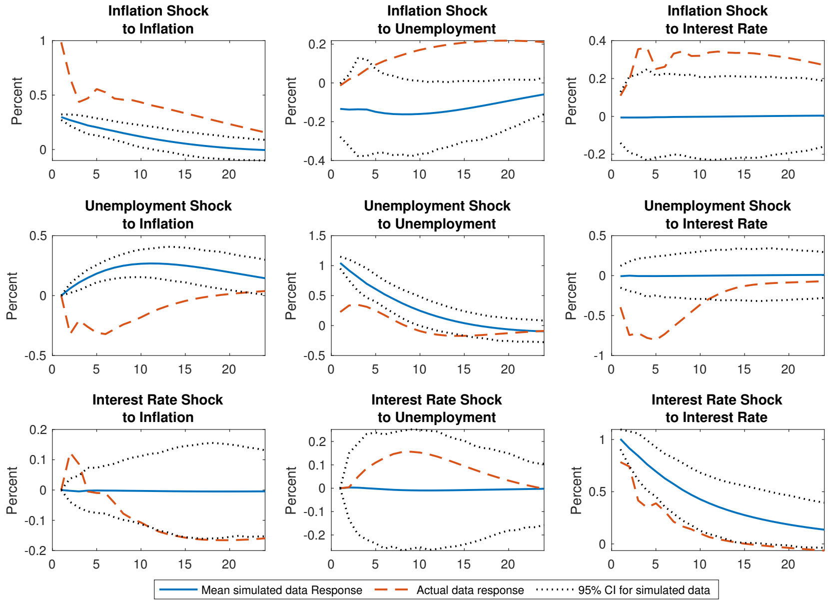

Figure 4: IRF with SVAR Estimation on the Simulated Data with Independent Blocks

To summarize, we find the results for the SVAR estimation to be strongly positive. Simple lag restrictions allow differentiating between simulated and actual data for all three shocks. In almost all cases, the difference between the responses in simulated and actual data is statistically significant. Moreover, the results in the simulated data seem to be often inconsistent with the economic intuition, which makes sense as we simulate the alternative realities.

The success of the SVAR estimation seems to be partly driven by its model-free agnostic assumptions. The methodology intrinsically allows monetary policy to have both positive and negative effects, both strong and weak. Having very mild restrictions and strong results allows differentiating the cases in the actual data relative to counterfactual scenarios.

Example 2. Smets and Wouters (2007).

While in the previous section, we analyze the problem of false positives arising in structural vector autoregressions, this section looks into the estimated dynamic general equilibrium model developed by Smets and Wouters (2007).

Summary of the Original Exercise

In this section, we apply the test to the bayesian estimation of the New Keynesian DSGE model developed by Smets and Wouters (2007). The authors demonstrate that the New Keynesian model with price and wage rigidities can successfully match the data in the paper.

They estimate the model using seven economic time series: the log difference of real GDP, real consumption, real investment, real wage, log hours worked, the log difference of the GDP deflator, and the Fed funds rate. We denote these series by $\triangle y_t, \triangle c_t, \triangle i_t, \triangle w_t, l_t, \pi_t$, and r_t.

The model also has seven exogenous processes to match the data: total factor productivity ϵ_t^a, investment-specific technology ϵ_t^i, risk premium ϵ_t^b, exogenous spending ϵ_t^g, price mark-up ϵ_t^p, wage mark-up ϵ_t^w, and monetary policy shocks ϵ_t^r. The behavior of these shocks follows the seven equations below:

$$\epsilon_t^g = \rho_g \epsilon_{t-1}^g + \eta_t^g + \rho_{ga} \eta_t^a, \tag{10}$$

$$\epsilon_t^b = \rho_b \epsilon_{t-1}^b + \eta_t^b, \tag{11}$$

$$\epsilon_t^i = \rho_i \epsilon_{t-1}^i + \eta_t^i, \tag{12}$$

$$\epsilon_t^a = \rho_a \epsilon_{t-1}^a + \eta_t^a, \tag{13}$$

$$\epsilon_t^p = \rho_p \epsilon_{t-1}^p + \eta_t^p - \mu_p \eta_{t-1}^p, \tag{14}$$

$$\epsilon_t^w = \rho_w \epsilon_{t-1}^w + \eta_t^w - \mu_w \eta_{t-1}^w, \tag{15}$$

$$\epsilon_t^r = \rho_r \epsilon_{t-1}^r + \eta_t^r. \tag{16}$$

Shocks $\epsilon_t^b, \epsilon_t^i, \epsilon_t^a, \epsilon_t^r$ follow AR(1) processes, and shocks $\epsilon_t^g, \epsilon_t^p, \epsilon_t^w$ follow ARMA(1,1) process. These processes are characterized by ten parameters of persistence $\rho_g, \rho_b, \rho_i, \rho_a, \rho_p, \rho_w, \rho_r,\ \rho_{ga}, \mu_p, \mu_w$ and volatilities $\sigma_a, \sigma_b, \sigma_g, \sigma_l, \sigma_r, \sigma_p, \sigma_w$. All of these parameters are going to be estimated and discussed further below.

Having the same number of shocks as variables allows the model exactly to match the data. While this certainly gives advantages in estimation, it also makes it harder for the model to fail. Regardless of how misspecified is the endogenous structure and policy implications, the model perfectly fits the data.

The endogenous structure of the model is more complex and is described below in 14 equations (17)-(30). There are also 14 endogenous variables: output y_t, consumption c_t, investment i_t, value of capital q_t, capital services used in production k_t^s, physical capital k_t, capital utilization z_t, the rental capital r_t^k, price mark-up μ_t^p, inflation π_t, wage mark-up μ_t^w, the real wage w_t, hours worked l_t, and interest rate r_t.

$$y_t = c_y c_t + i_y i_t + z_y z_t + \epsilon_t^g \tag{17}$$

$$c_t = c_1 c_{t-1} + (1-c_1)E_t c_{t+1} + c_2(L_t - E_t l_{t+1}) - c_3(r_t - E_t\pi_{t+1} + \epsilon_t^b) \tag{18}$$

$$i_t = i_1 i_{t-1} + (1-i_1)E_t i_{t+1} + i_2 q_t + \epsilon_t^i \tag{19}$$

$$q_t = q_1 E_t q_{t+1} + (1-q_1)E_t r_{t+1}^k - (r_t - E_t\pi_{t+1} + \epsilon_t^b) \tag{20}$$

$$y_t = \phi_p(\alpha k_t^s + (1-\alpha)l_t + \epsilon_t^a) \tag{21}$$

$$k_t^s = k_{t-1} + z_t \tag{22}$$

$$z_t = z_1 r_t^k \tag{23}$$

$$k_t = k_1 k_{t-1} + (1-k_1)i_t + k_2\epsilon_t^i \tag{24}$$

$$\mu_t^p = mpl_t - w_t = \alpha(k_t^s - l_t) + \epsilon_t^a - w_t \tag{25}$$

$$\pi_t = \pi_1\pi_{t-1} + \pi_2 E_t\pi_{t+1} - \pi_3\mu_t^p + \epsilon_t^p \tag{26}$$

$$r_t^k = -(k_t - l_t) + w_t \tag{27}$$

$$\mu_t^w = w_t - mrs_t = w_t - (\sigma_l l_t + \frac{1}{1-\lambda/y}(c_t - \lambda/\gamma c_{t-1})) \tag{28}$$

$$w_t = w_1 w_{t-1} + (1-w_1)(E_t w_{t+1} + E_t\pi_{t+1}) - w_2\pi_t + w_3\pi_{t-1} - w_4\mu_t^w + \epsilon_t^w \tag{29}$$

$$r_t = \rho r_{t-1} + (1-\rho)\{r_\pi\pi_t + r_Y(y_t - y_t^p)\} + r_{\triangle y}[(y_t - y_t^p) - (y_{t-1} - y_{t-1}^p)] + \epsilon_t^r \tag{30}$$

Equation (17) defines goods market clearing. Investment-output share i_y is defined by the following relationship $i_y = (\gamma - 1 + \delta)k_y$, where γ is the steady-state growth, and k_y is the steady-state output ratio. Capital utilization costs $z_y = R_*^k k_y$, where R_*^k is the steady-state rental rate of capital. Government share g_y is defined exgenously, and consumption share follows $c_y = 1 - i_y - g_y$.

In the Euler equation (18), $c_1 = (\lambda/\gamma)/(1+\lambda/\gamma)$, $c_2 = [(\sigma_c - 1)(W_*^h L_*/C^*)]/[\sigma_c(1+\lambda/\gamma)]$, and $c_3 = (1-\lambda/\gamma)/[(1+\lambda/\gamma)\sigma_c]$. The deep parameter λ reflects the degree of habit formation. Variables W_*^h, L_*, and C^* correspond to the steady state values of wages, hours, and consumption. This equation shows that the current consumption is affected by the past consumption, future expected consumption, expected growth in hours, real interest rate, and the exogenous shock ϵ_t^b.

In the equation describing investment dynamics (19), $i_1 = 1/(1+\beta\gamma^{1-\sigma_c})$, $i_2 = [1/(1+\beta\gamma^{1-\sigma_c})\gamma^2\varphi]$, ϕ is the steady-state elasticity of the capital adjustment cost function, and β is the household discount factor. Investments depend on their past and expected future value, the cost of capital q_t, and the investment-specific technology shock ϵ_t^i.

In the equation for the value of capital (20), $q_1 = \beta\gamma^{-\sigma_c}(1-\delta) = [(1-\delta)/(R_*^k) + (1-\delta)]$. The value of capital increaes with its expected future value, as well as expected real rental rate on capital. It declines in response to higher real interest rate and risk premium shock ϵ_t^b.

In the aggregate production function (21) output y_t increases with higher capital k_t^s and labor services l_t, and total factor productivity ϵ_t^a.

Capital services in equation (22) increase with the capital installed in the previous period k_{t-1} and its utilization z_t. Utilization of capital in equation (23) increases with the rental return r_t^k. Its sensitivity $z_1 = (1-\psi)/\psi$, and ψ is a function of the elasticity of the utilization adjustment cost, normalized to be between zero and one. In the equation for capital accumulation (24), $k_1 = (1-\delta)/\gamma$ and $k_2 = (1-(1-\delta/\gamma))(1+\beta\gamma^{(1-\sigma_c)})\gamma\varphi$. Capital here increases with previous capital, flow of investment, and investment-specific technology shock.

Price markup μ_t^p in equation (25) depends on the capital-labor ratio, real wage w_t, and total factor productivity ϵ_t^a. Inflation π_t in the Phillips curve (26) depends positively on the past, expected future inflation, and markup shock ϵ_t^p. It decreases with the current price markup μ_t^p. Here $\pi_1 = \iota_p/(1+\beta\gamma^{1-\sigma_c}\iota_p)$, $\pi_2 = \beta\gamma^{1-\sigma_c}/(1+\beta\gamma^{1-\sigma_c}\iota_p)$, and $\pi_3 = 1/(1+\beta\gamma^{1-\sigma_c})[(1-\beta\gamma^{1-\sigma_c}\xi_p)(1-\xi_p)/\xi_p((\phi_p-1)\epsilon_p+1)]$.

The rental rate of capital r_t^k in (27) decreases with capital-labor ratio $k_t - l_t$ and increases with real wage w_t. Wage markup follows (28), where it depends on the difference between the real wage and the rate of substitution between consumption and leisure. In this equation, σ_l is the elasticity of labor supply with respect to the real wage.

Behavior of real wages is described in (29), where $w_1 = 1/(1+\beta\gamma^{1-\sigma_c})$, $w_2 = (1+\beta\gamma^{1-\sigma_c}\iota_w)/(1+\beta\gamma^{1-\sigma_c})$, $w_3 = \iota_w/(1+\beta\gamma^{1-\sigma_c})$, and $w_4 = 1/(1+\beta\gamma^{1-\sigma_c})[(1-\beta\gamma^{1-\sigma_c}\xi_w)(1-\xi_w)/(\xi_w((\phi_w-1)\epsilon_w+1))]$. The real wage depends on its past and expected values, wage markup, expected, current, and past inflation, and wage-markup shock ϵ_t^w.

Monetary policy in equation (30) follows a Taylor rule, where the interest rate r_t reacts to its past value, inflation, output gap, and output gap change.

Having the model description completed, we are ready to proceed with its evaluation.

Introducing the Null and Building the Simulated Data

As before, we decompose the data into two blocks: the impulse block and the response block. The impulse block consists of only the interest rate or algebraically $y_t^1 = [r_t]$, and six other variables compose the response block, or mathematically $y_t^2 = [\triangle y_t, \triangle c_t, \triangle i_t, \triangle w_t, l_t, \pi_t]$. We construct the new simulated data using the methodology described above and assume the null hypothesis to hold.

Null Hypothesis 2 *Monetary policy does not affect the economy.*

According to the null, the movement of the interest rate should not have any effect on the data. However, the null still allows the rest of the economy to affect the monetary policy. Thus, we are not ruling out existing monetary policy rules, where the interest rate is affected by GDP growth and inflation. Below we outline the application of the general methodology to the problem.

1. In this step, we treat the original quarterly data on the interest rate, GDP growth, consumption growth, investment growth, wage growth, hours

worked, inflation as stationary. Thus, these data do not require any further modifications or filtering. More formally, $x_t^1 = y_t^1$ and $x_t^2 = y_t^2$.

2. We estimate the AR(1) process for the response block following (1), consisting of GDP growth, consumption growth, investment growth, wage growth, hours worked, and inflation. Our goal here is to preserve the relationship between the variables and their lags within the response block.

3. Here, we build the simulated values of GDP growth, consumption growth, investment growth, wage growth, hours worked, and inflation using pseudo-random generators according to equation (4).

4. Similarly to the previous section, we estimate the relationship in the fed funds rate's actual data. As before, we allow for two possibilities. In the DB case, we allow the fed funds rate to depend on contemporaneous GDP growth and inflation according to (5a). In the second IB case, we allow the fed funds rate to be independent of six other series and follow and AR(1) process according to equation (5b).

5. We simulate the data for the Fed fund rate for both cases. For dependent blocks or DB case, the interest rate depends on inflation and gdp growth, and the data is simulated according to (8a), where $\hat{\Lambda}_{12} = \begin{bmatrix} \lambda_{12}^y & 0 & 0 & 0 & 0 & 0 & \lambda_{12}^\pi \end{bmatrix}$. Thus, we allow interest rate to depend on contemporaneous inflation and output following Taylor rule. For independent blocks or IB case, Fed fund rate is simulated according to (8b).

6. The simulated raw data for both the impulse block $\tilde{x}_t^1$ and the response block $\tilde{x}_t^2$ is identical to the stationary counterparts $\tilde{y}_t^1$ and $\tilde{y}_t^2$ since the raw data in the paper are already differenced out.

7. We apply the Bayesian DSGE estimation to the simulated Fed fund rate, GDP growth, consumption growth, investment growth, wage growth, hours

worked, and inflation.

8. We repeat the procedure N=300 times and establish the mean impulse response for the identified shocks in the simulated data and the confidence intervals for the simulated data estimation.

After building the simulated data, we are ready to apply the model to these data and evaluate its conclusions regarding the relationship between the impulse and the response block.

Applying the Model to the Simulated Data

We simulate the series for 1000 periods and then use the last 232 observations, making the simulated data consistent with the authors' actual data. We run the simulations 300 times, and in each of the cases, estimate the model and obtain the mode. We then build the confidence interval for modes.

We report the results in Table 1 and Table 2 given below. In both tables, column 2 describes the mode value of the estimation performed on the actual data using the Metropolis-Hasting (MH) algorithm. Columns 3-5 report the mean value and the confidence interval for the parameters in the simulated data, where the impulse block can depend on the response block. Notice, in these columns, the quantiles are computed from sorting modes within 300 simulations. In other words, five % quantile for simulated posterior group corresponds to a mode of estimated parameters for a particular simulation. Finally, columns 6-8 correspond to the simulated IB case, where the impulse block is orthogonal from the response block.

Table 1: Prior and Posterior Distribution of Structural Parameters

	Actual Mode	Simulated Mode DB			Simulated Mode IB		
	Mode	Mean	95%	5%	Mean	95%	5%
φ	4.97	4.69	6.97	2.83	4.89	6.99	2.94
σ_c	1.32	1.03	1.62	0.60	0.93	1.40	0.55
h	0.73	0.67	0.86	0.40	0.70	0.87	0.44
ξ_w	0.72	0.76	0.85	0.63	0.76	0.86	0.62
σ_l	1.59	1.65	2.36	0.66	1.66	2.40	0.64
ξ_p	0.63	0.54	0.67	0.48	0.55	0.71	0.48
ι_w	0.57	0.61	0.74	0.48	0.61	0.74	0.47
ι_p	0.23	0.29	0.38	0.19	0.28	0.37	0.18
ψ	0.44	0.37	0.64	0.08	0.35	0.62	0.07
Φ	1.69	1.64	1.76	1.51	1.62	1.74	1.51
r_π	2.01	1.77	1.93	1.52	1.75	1.96	1.46
ρ	0.81	0.88	0.91	0.83	0.88	0.91	0.83
r_y	0.08	0.11	0.19	0.02	0.11	0.21	0.02
$r_{\triangle y}$	0.21	0.13	0.17	0.08	0.13	0.17	0.07
$\bar{\pi}$	0.60	0.63	0.83	0.47	0.63	0.85	0.47
$100(\beta^{-1}-1)$	0.17	0.18	0.21	0.16	0.18	0.21	0.17
$\bar{l}$	0.55	0.06	2.37	-3.20	-0.14	2.16	-3.33
$\bar{\gamma}$	0.45	0.49	0.59	0.35	0.49	0.60	0.36
$\bar{\alpha}$	0.30	0.29	0.35	0.23	0.30	0.36	0.23

Table 2: Posterior Distribution of Shock Processes

	Actual Mode	Simulated Mode DB			Simulated Mode IB		
	Mode	Mean	95 %	5%	Mean	95 %	5 %
σ_a	0.43	0.48	0.52	0.43	0.48	0.52	0.43
σ_b	0.23	0.19	0.37	0.10	0.18	0.38	0.10
σ_g	0.55	0.69	0.72	0.61	0.68	0.72	0.61
σ_l	0.41	0.49	0.67	0.30	0.48	0.70	0.28
σ_r	0.24	0.25	0.28	0.22	0.27	0.30	0.23
σ_p	0.14	0.20	0.24	0.14	0.20	0.24	0.14
σ_w	0.24	0.26	0.30	0.22	0.26	0.30	0.22
ρ_a	0.96	0.96	0.99	0.88	0.95	0.99	0.87
ρ_b	0.20	0.68	0.90	0.15	0.70	0.91	0.15
ρ_g	0.97	0.96	0.99	0.91	0.96	0.99	0.90
ρ_l	0.77	0.69	0.96	0.46	0.68	0.96	0.44
ρ_r	0.14	0.19	0.30	0.09	0.24	0.34	0.11
ρ_p	0.92	0.90	0.98	0.73	0.90	0.99	0.70
ρ_w	0.97	0.88	0.98	0.63	0.87	0.98	0.61
μ_p	0.75	0.63	0.82	0.39	0.62	0.83	0.37
μ_w	0.87	0.78	0.92	0.53	0.78	0.92	0.53
ρ_{ga}	0.60	0.64	0.79	0.47	0.65	0.80	0.47

We are ready to discuss the results displayed in Table 1 and Table 2. Table 1 shows that most of the structural parameters not relevant for monetary policy are similar between the actual and simulated data. When we look at the sensitivity of the interest rate to inflation r_π, its value 2.03 is higher for the actual data than for the simulated data with dependent blocks, which varies between 1.41 and 1.89, however it falls within the 90 percent interval for the simulation with independent blocks. For the sensitivity of the interest rate to output growth $r_{\triangle y}$, its value in the actual data 0.22 is higher than the value within 90 % confidence intervals for both the simulation with independent blocks (0.13;0.21) and dependent blocks. On the other hand, the sensitivity of interest rate to output gap r_y for actual data is well within the 90 % confidence interval for both simulations with dependent and independent blocks. When we look at the monetary policy's persistency ρ, its value is 0.81 in the actual data than the confidence interval (0.87; 0.93) for the simulated data with dependent blocks and the simulated data with independent blocks (0.84;0.92). Overall, the results in Table 1 suggest that the interest rate reacts somewhat more to inflation and output growth than in the simulated data. However, we do not observe big economic differences between the estimated results on simulated and actual data despite having a statistical difference.

In Table 2, the standard deviation of the monetary shock averaged across the simulated data is 0.27, which is equal to the upper 95 % bound for the estimation performed on the actual data, which is equal to 0.24. The mode of autocorrelation for the monetary shock in the simulated data is 0.13 with dependent blocks and 0.22 in the independent blocks, which is higher than the mode of the estimated value of 0.12 from the actual data. Overall, monetary shocks appear to be more volatile and persistent in the simulated data than in the actual data, pointing towards a more significant monetary shocks role. This result is surprising since, in the simulated data, monetary policy is entirely irrelevant.

Impulse responses for the monetary policy in the simulated data are given below in Figure 5 for dependent blocks and in Figure 6 for independent blocks. The solid line in each figure corresponds to the impulse response generated from the estimation performed on the actual data, where the interest rate goes up by one standard deviation. We compose the dashed line by averaging three hundred impulse responses in the simulated data, computing each impulse response for the mode of parameters.

Both Figure 5 and Figure 6 demonstrate that the interest rate's effect on average increases when the estimation is performed on the simulated data for a comparable shock to monetary policy. This finding is surprising because monetary policy does not affect macroeconomic data in the simulation. Nevertheless, it is consistent with the fact that monetary shocks have higher standard deviation and persistence in simulated data. The estimation performed on the actual data implies that an increase in the interest rate by 0.24 percentage points leads to a decline in output by 0.2 percent. However, a rise in interest rate by 0.27 percent leads to a reduction in output by 0.35 percent for the simulated data in Figure 5. These results strongly imply that the model cannot distinguish the estimation on the actual data from the simulated cases, where the monetary policy does not affect the real economy.

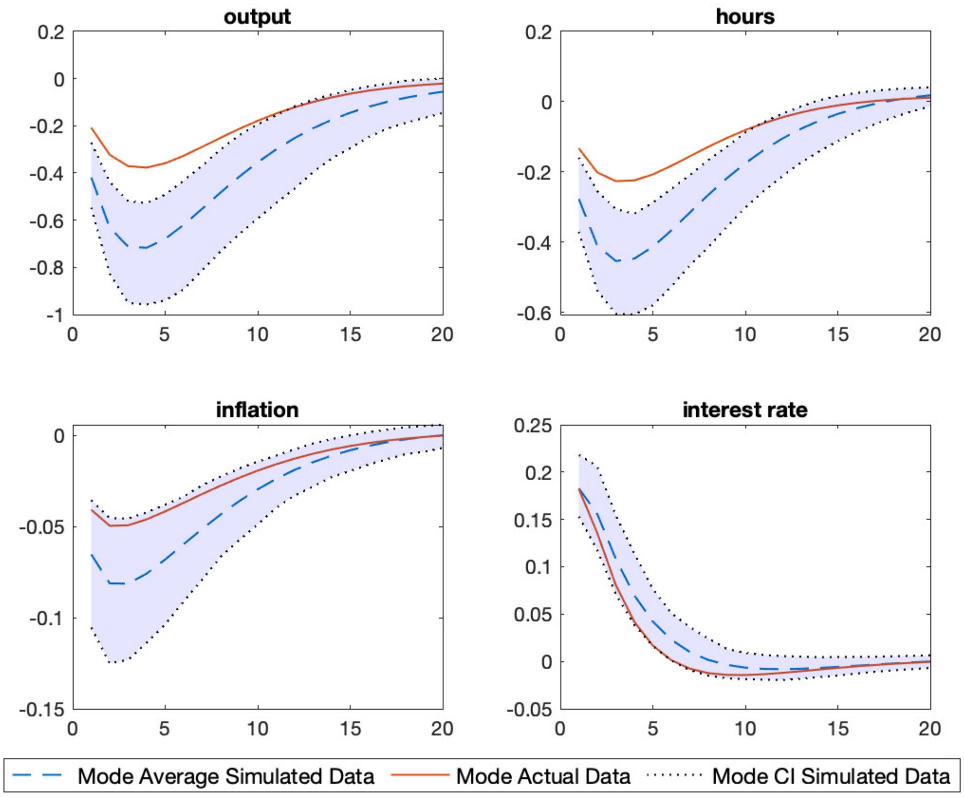

Figure 5: Impulse Response Functions To A Monetary Policy Shock in The Actual and Simulated Data with Dependent Blocks

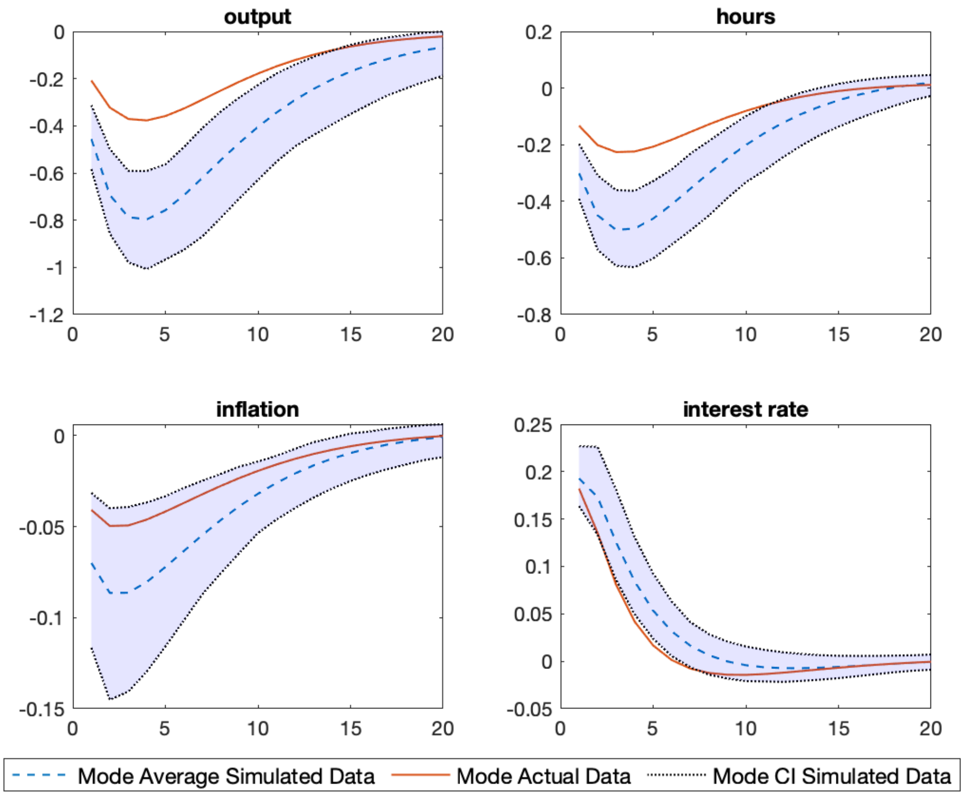

Figure 6: Impulse Response Functions To A Monetary Policy Shock in The Actual and Simulated Data with Independent Blocks

We then shift our focus to the variance decomposition results displayed in Table 3 and Table 4. Table 3 shows the variance decomposition of output Y, labor L, inflation π, and interest rate r for the seven exogenous processes described above. Our primary focus in this table is the last column, which corresponds to the portion of the variance explained by monetary shocks ϵ_r. We see that monetary shocks explain a smaller percentage of output, labor, and inflation in the actual data relative to simulated data with dependent or independent blocks. This result is intuitive since these shocks have lower volatility and persistence relative to the simulation. The only exception to this rule is the interest rate itself, where shocks in the actual data explain 12 percent of its total variation. In comparison, in the estimation with dependent blocks, they explain 11 percent of the variation.

Table 3: Variance Decomposition

	Temp. TFP ϵ^a	Risk Premium ϵ^b	Ex. spending ϵ^g	Inv. Specific ϵ^i	Price Markup ϵ_p	Wage Markup ϵ^w	Mon. Pol. ϵ_r
Y actual	25.49	1.45	3.02	13.77	8.52	45.40	2.35
L actual	1.87	2.29	8.49	13.80	5.97	64.56	3.02
π actual	2.30	0.50	0.66	5.13	30.64	56.74	4.04
r actual	6.96	6.81	2.81	28.50	7.61	34.25	13.07
Y simulated	38.16	10.81	4.28	13.99	14.83	10.65	7.27
L simulated	6.38	19.34	13.24	17.35	11.97	19.49	12.23
π simulated	3.47	5.26	0.38	2.29	58.30	23.64	6.67
r simulated	4.57	40.72	1.50	9.43	15.31	13.33	15.14
Y simulated IB	35.77	11.58	4.48	13.06	14.49	11.47	9.14
L simulated IB	6.03	19.92	12.65	16.07	10.91	19.56	14.85
π simulated IB	3.39	4.50	0.29	1.81	58.64	23.80	7.56
r simulated IB	4.26	40.53	1.12	6.78	14.91	14.31	18.10

Table 3 displays the mean contribution of monetary shocks across 300 simulations, and Table 4 shows the confidence interval of contribution to the variance of monetary shocks across different simulations. These results strongly suggest no statistical difference between the contribution to the variance of monetary shocks in the actual and simulated data. For example, while monetary shocks explain 12 percent of the interest rate movement in the actual data, the 90 percent confidence interval for the simulated data with dependent blocks is between 3 percent and 28 percent, and between 4 and 37 percent with independent blocks. The results in Table 4 also strongly suggest no statistical difference between monetary shocks' role in actual and simulated data.

Table 4: Contribution of Monetary Shocks in Variance Decomposition of Aggregate Variables

	Actual	Simulated Mean DB			Simulated Mean IB		
	Mode	Mode	95%	5%	Mode	95%	5%
y	2.35	7.27	17.77	1.98	9.14	19.20	2.19
l	3.02	12.23	25.47	4.11	14.85	28.32	4.16
π	4.04	6.67	13.53	2.15	7.56	15.81	2.01
r	13.07	15.14	29.85	5.17	18.10	35.74	6.63

Interpreting the Results for Smets and Wouters (2007) Model

Introducing the test for Smets and Wouters (2007) model suggests two sets of results. First, when we construct the data in a way that makes monetary policy irrelevant, Bayesian estimation indicates that monetary policy still has strong effects. Thus, the model cannot reject the null that monetary policy does not affect the economy. Therefore, it generates false positives and type I error.

Second, monetary policy shocks play a minor role in both the simulated and actual data. This finding partly suggests that the real effect of monetary policy might be small. Why? Because if the model assumes strong monetary policy effects regardless of whether it is true in the data, then the estimation should generate small shocks for monetary policy if the latter is not essential. In the exercise above, we show both facts. First, the model enforces monetary policy to be influential on the data. Second, monetary policy shocks' role is smaller in the actual data than in the simulated data.

To be precise, we do not prove that monetary policy is not essential. Instead, we show that the Smets and Wouters (2007) model rather assumes than proves that monetary policy has strong effects. Since the microfounded model's primary goal is to help the policy analysis, we have found our result to provide a severe limitation for conclusions from Smets and Wouters (2007).

Conclusion

We propose a universal and straightforward test for validating assumptions in the structural models. Since the data in economics are limited, persistent, and are often impossible to replicate, there is a risk that structural models can be misspecified despite being consistent with the data. Structural models impose a causal structure on economic variables, take the actual data as an input, and then produce the parameters that make the model consistent with the data. We simulate the data under an alternative causal relationship, feed these simulated data to the model, and then see whether the model can produce different conclusions from the simulated data.

We apply the test to the recursive structural VAR estimation of monetary policy by Stock and Watson (2001). We simulate the data under the assumption that monetary policy is irrelevant and run the simulated data's estimation procedure. We find that the model can successfully differentiate between simulated and actual data.

We then apply the test to the estimation and data by Smets and Wouters (2007), which is estimated via full-information Bayesian method. As before, we simulate the data assuming that monetary policy is irrelevant and then feed these simulated data to the model. To our surprise, the model concluded that monetary policy has a strong effect in the simulations contrary to the data generating process assumptions. Thus, we have concluded that the Smets and Wouters (2007) is not suitable for policy analysis.

In terms of direct implications, our results suggest the surprising strength of indirect inference methods, such as matching up impulse responses between the model and the data ala Christiano et al. (2005) relative to full-information estimation. Nevertheless, we have tested our methodology on only two models, and further research is required to establish whether a particular estimation method has an advantage.

Our results point the economic models should develop stronger mechanisms of measuring out-of-sample performance and focus less on fitting and replicating the data. While in data science, validating results plays a central role; many economic models still focus on the performance in sample, which creates the danger of overfitting misspecified models. We partly address this problem in this project, but new research and testing of other models using our methodology are required.

References

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (2005). A critique of structural VARs using real business cycle theory. Technical report.

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (2008). New Keynesian models: not yet useful for policy analysis. Technical report.

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

Christiano, L. J., M. Eichenbaum, and R. Vigfusson (2007, May). *Assessing Structural VARs*, pp. 1–106. MIT Press.

Cooley, T. F. and M. Dwyer (1998). Business cycle analysis without much theory A look at structural VARs. *Journal of Econometrics 83*(1-2), 57–88.

Smets, F. and R. Wouters (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review 97*(3), 586–606.

Stock, J. H. and M. W. Watson (2001, December). Vector autoregressions. *Journal of Economic Perspectives 15*(4), 101–115.