Identifying Demand and Supply Shocks*

Marco Brianti† Fang Li‡

August 30, 2023

Abstract

We used a new identification strategy to identify the aggregate demand and the aggregate supply shock on real U.S. data. The simulation study and the robust test show that this identification strategy works well: (1) the estimated implied impulse response functions are consistent with the model-implied impulse response functions; (2) the estimated shocks and actuals shocks are very highly correlated. The empirical results indicate that GDP is less persistent than inflation in response to the demand shock, whereas GDP is more persistent than inflation to the supply shock. Also, the demand shock is a more vital driver of boom-bust cycles than the supply shock.

JEL classification:
Keywords: demand shocks, supply shocks, SVAR

*

†Department of Economics, University of Alberta, 8-14 Tory Building, Canada T6G 2H4; mail:
‡Department of Economics, University of Alberta, 8-14 Tory Building, Canada T6G 2H4; mail:

1 Introduction

Identifying structural shocks, such as the aggregate demand (AD) and aggregate supply (AS) shocks, is a crucial topic in macroeconomics (Blanchard and Quah, 1988; Gali, 1992). Given that both output (and unemployment rate) and inflation are influenced by the (aggregate) demand and supply shocks, these shocks are essential to explain the business cycles. According to the AD-AS analysis framework, it is clear that the demand shock and supply shock have different impacts on output and inflation: a positive demand shock imposes a positive impact on both the output and inflation, whereas a positive supply shock imposes a positive impact on the output and the negative impact on the inflation. In this paper, we apply a new identification strategy proposed by Brianti (2021) to identify the aggregate demand and supply shocks using this information. In this paper, we define the aggregate demand shock as the shock that shifts the AD curve and the aggregate supply shock as the shock that moves the AS curve.

Generally, two popular methods are widely used to identify the (aggregate) demand and supply shocks. One is the sign restriction (Faust, 1998; Uhlig, 1998; Fry and Pagan, 2011). Sign restriction relies on the qualitative (directional) information of the response to identify shocks. It repeatedly draws the impact matrix (or rotation) and discards the inconsistent draws until the number of simulations is satisfied. Baumeister and Hamilton (2015) and Furlanetto et al. (2019) used the sign restriction to identify the demand and supply shocks in a specific market. Furlanetto et al. (2019) identifies six shocks in a six-variable SVAR model by imposing sign restrictions for every shock. Specifying signs for every shock excludes the possibility of multiple shocks problem. But this method is very time-consuming in a large VAR system. The identification of their six-variable system took almost one week to complete[1]. Also, sign restriction is not a point estimate but a set estimate. We can only estimate the bound of impulse responses even if we have a huge amount of data, which gives less precision. Normally, researchers present the median impulse responses with the confidence interval to illustrate how variables react to shocks. According to Inoue and Kilian (2022), the median could sometimes be very misleading and cannot be trusted to capture the shape of the impulse response function.

[1] They are using the Bayesian sign restriction. That is, the estimation depends on the drawn rotation and the reduced-form parameters drawn from the posterior distribution. Drawing from the posterior distribution is not the time consuming part.

The other famous method to identify the demand and supply shock is proposed by Blanchard and Quah (1988). The starting point of their identification strategy is the demand-side shock does not have a long-term effect on the output, but the supply-side shock permanently impacts the output[2]. This is referred to as "long run identification" because the identification strategy excludes the long-term impact of a specific variable. Actually, Blanchard and Quah's method depends on a specific representation of the variable. However, economic theory does not provide sufficient structure to choose among many moving-average representations associated with an estimated VAR (Lippi and Reichlin, 1993).

Our method, the generalized penalty function approach, takes the directional information as in the sign restriction to construct a target function. We pick the impact matrix by maximizing this target function. That means our method does not depend on any decomposition of the economic series or variables. In addition, it is less time-consuming than the sign restriction, and we can use bootstrap to get the confidence interval. Depending on the information we used to construct the target function, our method still subject to the multiple shock problem, our method provides a upper bound of the impulse responses estimate.

The rest of the paper is organized as follows. Section 2 provides the detail about the identification methodology used in the paper. Section 3 provides tests of the reliability and stability of the methodology. We use a simple New Keynesian model to simulate data and use our method to identify shocks in the simulated data. We also conduct some robustness tests on the method. Section 4 applies the methodology on U.S. aggregate data to identify demand and supply shocks. Section 5 concludes the paper.

2 Methodology

Our identification strategy follows Brianti (2021), which identifies the macroeconomic shocks by maximizing a specific target function. Without loss of generality, consider a n-variable dynamic system consists of various macroeconomic variables, $X_t = [Y_t, \pi_t, V_t']'$. X_t is a $n \times 1$ endogenous variable vector, Y_t is the output, π_t is the inflation, and V_t represents the rest $(n-2)$ endogenous variables. Here we separate the output and inflation from the rest macro variables because our identification strategy relies on the information

[2] The application of our identification method on the real U.S. data also confirms this pattern.

on the output and inflation. And we will use the terms demand shock and supply shock instead of aggregate demand shock and supply shock in the rest of the paper for simplicity.

The reduced-form VAR (we suppress the mean) is,

$$X_t = B(L)X_{t-1} + u_t$$

where $B(L) = B_1 L + B_2 L^2 + \cdots + B_p L^p$ and B_i is a $n \times n$ matrix. u_t is the reduced-form innovations and $E(u_t u_t^{'}) = \Sigma$, is the variance-covariance of the $u_t = [u_t^Y, u_t^\pi, \cdots, u_t^n]^{'}$. The objective of the identification strategy is to identify structural shocks of interest: the demand shock (ϵ_t^D) and the supply shock (ϵ_t^S) from the reduced-form innovations with a proper structural impact matrix A, such that $A\epsilon_t = u_t$, and $E(\epsilon_t \epsilon_t^{'}) = I$ and $\epsilon_t = [\epsilon_t^D, \epsilon_t^S, \cdots]$. Given this structure, we have $\Sigma = AA^{'}$. Following the structural VAR literature, we assume that the shocks ϵ are orthogonal to each other with unit variance, i.e., I is an identity matrix. Without any restrictions, there will be infinite numbers of possible candidates for matrix A. That is, suppose Q is an *arbitrary* orthogonal matrix and let $w_t = AQ\epsilon_t$, then $E(w_t w_t^{'}) = AA^{'} = \Sigma$.

In order to identify the demand and supply shocks, according to the AD-AS model, we have the following information at hand for identification: the demand shock, ϵ_t^D, imposes a positive impact on both output, Y_t, and inflation, π_t, and the supply shock, ϵ_t^S, imposes a positive impact on the output, Y_t, and a negative impact on the inflation, π_t. Table 1 summarizes this qualitative information:

Table 1: The Summary of Restriction

Parameter	Demand shock, ϵ_t^D	Supply shock, ϵ_t^S
output, Y_t	positive	positive
inflation, π_t	positive	negative

With these information in hand, here defines the econometric procedure.

Definition 1 *Decompose $A = CD$ where C is the Cholesky decomposition of Σ and $D = [d_D, d_S, ...]$ is an orthogonal matrix where:*

1. *Column vector d_D is the solution of the following problem,*

$$d_D = \arg\max_d \{\delta^* e_1 Cd + (1-\delta^*) e_2 Cd, \text{ subject to } d^{'} d = 1\}; \tag{1}$$

2. *Column vector d_S is the solution of the following problem,*

$$d_S = \arg\max_{d} \{\delta^* e_1 Cd - (1-\delta^*) e_2 Cd, \text{ subject to } d' d = 1\}; \tag{2}$$

3. *$\delta^* \in (0,1)$ is such that $d_D' d_S = 0$.*

Please note that e_j is the $j-$th row of a $n \times n$ identity matrix, we use e_j to pick out the $j-$th element of a column matrix. Cd_D and Cd_S represent the impact of a standard deviation demand shock ϵ_t^D and a standard deviation supply shock ϵ_t^S respectively, and δ^* is a scalar that takes a real value strictly between zero and one. In the following part, we mathematically show that δ^* is determined solely by σ_1 and σ_2. It governs the weights between two parts of information in identifying a shock. With the above construction of impact matrix A, we know $AA' = CDD'C' = CC' = \Sigma$ satisfies $A\epsilon_t = u_t$ and then the demand and supply shocks are identified.

Given ϵ_t^D imposes a positive impact on both Y_t and π_t, the response of Y_t and π_t to ϵ_t^D should be *increasing* functions of the size of the demand shock, ϵ_t^D. So, in Problem (1), the demand shock ϵ_t^D is identified by maximizing a function that is increasing in the (contemporaneous) response of the output Y_t ($e_1 Cd_D$) and increasing in the (contemporaneous) response of the inflation π_t ($e_2 Cd_D$). And δ^* governs the weights of these two pieces of information. Given ϵ_t^S imposes a positive (negative) impact on Y_t (π_t), the response of output (inflation) to the supply shock ϵ_t^S should be a *increasing* (*decreasing*) function of the size of the supply shock, ϵ_t^S. So, in Problem (2), the supply shock is identified by maximizing a function that is increasing in the (contemporaneous) response of the output Y_t ($e_1 Cd_S$) and decreasing in the (contemporaneous) response of the inflation π_t ($e_2 Cd_S$). The parameter δ^* governs the weights of these two pieces of information. In addition, the constraint $d_D d_D' = 1$, $d_S d_S' = 1$ are conditions that guarantees the matrix D is an orthogonal matrix along with the last condition $d_D d_S' = 0$ by choosing a proper value of δ^*. This procedure identifies one shock controlling the other shock by jointly identifying two shocks together. Logically, the name indicates the method should be minimizing the penalty function. Given the fact that minimizing the penalty function is equivalent to maximizing the negative of the penalty function, we call it the generalized penalty function, or the target function.

Actually, as long as the D matrix is orthogonal and C is the Cholesky decomposition of Σ, $A = CD$ would work as the impact matrix (Uhlig, 2005). For the rest shocks, they impose a different impact on the output Y_t and inflation π_t. Our method identifies the demand shock ϵ_t^D and the supply shock ϵ_t^D without controlling for other structural shocks. This makes sense. The demand and supply shocks are not necessarily structural shocks but a mixture of several shocks. Since we define the demand shock as the shock that shifts the AD curve and the supply shock as the shock that moves the AS curve, any shock that has the same effect on the output and inflation as summarized in Table 1 should be our target. Also, according to this identification strategy, we can find the most suitable shock vectors for the first or the second part of the information at the econometrician's hand. Without further information on the other shocks, we should believe the demand and the supply shock are identified. We pick the shocks that are the most decisive response of the variables (Uhlig, 2005). If there exists another shock, ϵ_t^K, imposes the same directional restrictions on the output and inflation with the demand shock (or the supply shock). Then our method identifies the upper bound of the demand shock (or the supply shock) since we leave the rest variables unconstrained.

Third, our methodology is close to the sign restriction (Faust, 1998; Rubio-Ramirez et al., 2010; Fry and Pagan, 2011). But our procedure provides the point estimation of the impulse response functions instead of the confidence interval, and it is more computationally efficient than sign restriction. Our methodology used not only the directional information (just like sign restriction) but also use the magnitude information. In Furlanetto et al. (2019), they estimated a six variable system using sign restriction, and the process took almost one week to finish. In addition, unlike Caldara et al. (2016), who used Uhlig's method to identify financial and uncertainty shocks, our procedure does not need a specific ordering assumption since we simultaneously identify two shocks instead of sequentially identifying one shock after the other one. Given that we use quantitative information to simultaneously identify the demand and supply shocks, our methodology is still valid even if the inflation π_t negatively responds to demand shock ϵ_t^D and supply shock ϵ_t^S but is more negative to the supply shock ϵ_t^S. In this case, the target function corresponding to the demand shock ϵ_t^D should also be an increasing function in the response of the inflation π_t compared to the target function corresponding to the supply shock ϵ_t^S.

In the above procedures, the imposed response functions are the instantaneous response to the shocks. We could extend the penalty function to include several periods of the impulse response function, not only the contemporaneous impact. But we only focus on the contemporaneous impact since the directional information is well accepted and the most theory grounded. According to the AD-AS framework, we are being agnostic about the subsequent response of output and inflation to the demand and supply shocks. Lastly, our method uses a unique weighting parameter in Problem (1) and Problem (2). That leaves one problem in this econometric procedure: the existence and uniqueness of δ^*. Following Brianti (2021), here we present the following proposition.

Proposition 1 *If $|Corr(u_t^Y, u_t^\pi)| \neq 1$, then the solution δ^*, d_D, and d_S exists and unique:*

$$\delta^* = \frac{\sigma_2}{\sigma_1 + \sigma_2}$$

The proof of this proposition can be find in Appendix. Given the δ is the weighting in the first term in both problems, and the larger σ_2 compared to σ_1 (smaller θ^*) puts a higher weight on the first term (larger δ^*). That means if one of the reduced-form innovations has a larger variance, the estimation of that reduced-form innovation is less precise, and we put less weight on that term, i.e., we depend less on that part of the information in both problems. This proposition indicates that our method is always valid if the two reduced-form innovations are not perfectly positively correlated or perfectly negatively correlated. At least one of the reduced-form innovations carries some information that the other one does not, this is what needs to achieve identification. Mathematically, in the above bivariate case, if $|\text{Corr}(u_t^Y, u_t^\pi)| = 1$, then $\exists\, a,\, b$ such that $\Pr(u_t^Y = a + bu_t^\pi) = 1$. It is almost surely that there is a linear relation between u_t^Y and u_t^π, which implies $\text{rank}(u_t) = 1$. Given $A = CD$ and C is the Cholesky decomposition of Σ and D is orthogonal, then A must be invertible and $\epsilon_t = A^{-1}u_t$. We know $\text{rank}(\epsilon_t) \leq \min\{\text{rank}(A^{-1}), \text{rank}(u_t)\} = 1$, i.e., ϵ_t^D and ϵ_t^S are not orthogonal to each other. This contradicts the orthogonality requirements for shocks.

If we do not restrict the δ in Problem (1) and Problem (2) to be the same, there will be infinite numbers of solutions given $|\text{Corr}(u_t^Y, u_t^\pi)| \neq 1$. That means our method turns to be set-identified, not point identified. If we present the median impulse response, it is possible that the median impulse responses are misleading (Inoue and Kilian, 2022). Inoue and Kilian (2022) provided a way of picking the most probable impulses responses according to the posterior kernel. We need to add Bayesian part in order to adopt their method. Since the core of the paper is to show the method, we are still using the frequentism method[3].

[3] Even in the Bayesian setting, our method is still less time-consuming comparing to sign restriction.

Details can be also found in Appendix.

Algorithm 1 Identifying Procedure

Input:

The initial value for δ, $\delta^0 = 0$;
The estimated covariance matrix of the reduced-form VAR, Σ;
The step length, η;

Output:

An orthogonal matrix, D;

1: Cholesky decompose $\Sigma = CC'$;
2: **while** $d_1 d_2' \neq 0$ **do**
3: $\quad d_1 = \arg\max_d \{(1-\delta^n)e_1 Cd + \delta^n e_2 Cd, \text{ subject to } dd' = 1\}$;
4: $\quad d_2 = \arg\max_d \{(1-\delta^n)e_1 Cd - \delta^n e_2 Cd, \text{ subject to } dd' = 1\}$;
5: $\quad \delta^{n+1} \leftarrow \delta^n + \eta$;
6: **end while**
7: Fill in the rest columns of matrix $D = [d_1, d_2, \tilde{D}]$ to make it orthogonal;

This econometric procedure searches the optimal δ on the $[0,1]$ grid. Starting with a specific value of δ, we construct the target functions in Problem (1) and in Problem (2). Then we maximize each target function to find the optimal shock vector, d_D and d_S, under the normalization condition. Next, we check if these two shock vectors are orthogonal to each other: if not, we search for the next δ value by adding a small number (step length) to the current δ, and otherwise, we stop and return these two shock vectors. The proposition 1 guarantees that δ^* exists and it is unique. So the algorithm works. For the rest column vector in the matrix D, we don't care about them, we just randomly select the rest columns to make matrix D orthogonal[4] (most commercial software packages can do that). That is, we are agnostic about the rest shocks. The procedure can be summarized as the following algorithm 1.

In order to test the reliability of the econometric procedure, we simulate data from a simple New Keynesian model described in the Section 3 and use this econometric strategy to identify demand and supply shocks on U.S. data in the Section 4.

[4] To solve this mathematically, it is equivalent to find an orthonormal basis of null space of $[d_1, d_2]'$. Suppose a and b are two vectors in the orthonormal basis, then: (i) orthonormal basis implies $a'b = b'a = 0$ and $a'a = b'b = 1$; and (ii) basis of the null space of $[d_1, d_2]'$ implies $[d_1, d_2]'a = [d_1, d_2]'b = [0, 0]'$.

3 Monte Carlo Simulation

The natural setting to test the reliability of this strategy is the simplest NK model, where we only focus on the supply and demand shocks. This baseline model has several shocks, including the supply and the demand shock. Given the information that the demand shock imposes a positive impact on the output and inflation, whereas the supply shock imposes a positive impact on output and a negative impact on inflation (this can be verified by looking at the model-implied impulse responses), we can use the strategy described in the Section 2 to identify the demand and supply shock from the simulated data. Once we get the impulse response from the simulated data, we compare them with the model-implied impulse responses.

3.1 Model

This section describes a simple DSGE model with nominal frictions (Rotemberg, 1982). We want to use this model to simulate the data and test the reliability of the identification strategy.

The economy is populated by (i) a continuum of homogeneous utility-maximizing households that choose consumption C_t and leisure $1 - N_t$; (ii) a continuum of value-maximizing firms $i \in [0, 1]$ that makes pricing and production decision to maximize the discounted present value of dividends in the current and future periods; (iii) a retailer that aggregates the differentiated goods produced by each firm; (iv) a central bank that is responsible for setting the monetary policy, the nominal risk-free rate.

3.1.1 Households

The model contains a continuum of homogeneous households that consumes the composite consumption goods C_t. The preference of households are defined over the consumption (of the composite consumption goods) and the leisure with a separable CRRA utility function as follows

$$\mathbb{E}_t \sum_{s=0}^{\infty} \beta^s b_{t+s} \left[\frac{C_{t+s}^{1-\sigma_c}}{1-\sigma_c} + \psi \frac{(1 - N_{t+s})^{1-\sigma_n}}{1-\sigma_n} \right]$$

where $\beta \in [0, 1]$ is the discount factor that governs the household patient level. Parameter σ_c is the relative risk aversion coefficient of the household, governs the risk attitude of the household, and σ_n is the Frisch elasticity of labor. ψ measures the relative importance of consumption and leisure in generating utility. b_t is the variable that governs the preference shock (Smets and Wouters, 2003). This is a shock to the discount factor and hence affects the inter-temporal substitution between the household's current and future consumption. From the Euler equation (4), given $\sigma_c > 1$, a large preference shock increases b_{t+1}, and then the next period consumption also increases accordingly. So we can consider this preference shock as a demand shock.

In addition, the household maximizes the present value of utility subject to the following budget constraint

$$C_t + \frac{B_t}{P_t} = \frac{W_t}{P_t}N_t + (1 - i_{t-1})\frac{B_{t-1}}{P_t} + d_t \tag{3}$$

This budget constraint is expressed in real terms. W_t is the nominal wage rate for labor, and we can define $w_t = W_t/P_t$ as the real wage rate for the labor. i_{t-1} is the nominal net interest rate set by the central bank on a previous risk-free bond B_{t-1}. d_t represents a series of real transfers from firms to the household, we assume that the household owns firms, and thus can be considered as the dividend.

Optimal conditions include the inter-temporal Euler equation for risk-free bonds

$$1 = \beta\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\sigma_c}\frac{b_{t+1}}{b_t}\frac{1+i_t}{\pi_{t+1}}\right] \tag{4}$$

and the intra-temporal condition for the labor supply

$$w_t = C_t^{\sigma_c}\psi(1 - N_t)^{-\sigma_n}$$

In the Euler equation, $\pi_{t+1} = P_{t+1}/P_t$ is the overall price inflation in period $t + 1$.

3.1.2 Firms

The model contains a continuum of monopolistically competitive firms indexed by $i \in [0, 1]$ producing a differentiated variety of goods with the following production function

$$Y_{i,t} = A_t N_{i,t}^{1-\alpha}$$

A_t is the (total) factor productivity. In addition, $\alpha \in (0, 1]$ governs the degree of decreasing return to scale of labor input $N_{i,t}$. If $\alpha = 0$, the production function turns to constant return to scale production function.

Firms make pricing decisions $P_{i,t}$ after the current period total factor productivity shock is realized. By knowing the price, the firm decides how much labor $N_{i,t}$ to hire and how much product $Y_{i,t}$ to produce. The firm then decides on the transfer to the household by the following constraint

$$d_{i,t} = \frac{P_{i,t}}{P_t}Y_{i,t} + \frac{B_{i,t}}{P_t} - w_t N_{i,t} - (1+i_{t-1})\frac{B_{i,t-1}}{P_t} - \frac{\gamma_p}{2}\left(\frac{P_{i,t}}{P_{i,t-1}} - \pi_{ss}\right)^2 Y_t \tag{5}$$

This condition contains the nominal rigidity expressed by the price adjustment cost a la Rotemberg (1982). If the firm decides to choose a price that changes a lot from its previous price, the firm needs to pay a real term cost. That means firms will not change the price a lot in order to avoid the incurred cost and generate the price sticky. In addition, π_{ss} is the steady-state value of inflation.

The firm's objective is to maximize the expected present value of dividends,

$$\max_{\{P_{i,t+k}, N_{i,t+k}, Y_{i,t+k}, B_{i,t+k}\}} \mathbb{E}_t \left[\sum_{k=0}^{\infty} m_{t,t+k} d_{i,t+k}\right]$$

where $m_{t,t+k}$ represents the stochastic discount factor set by the household, subject to the firm's budget constraint, the production function, and demand for $Y_{i,t}$ introduced in the following subsection.

3.1.3 Retailer

In the economy, there is a retailer that takes $Y_{i,t}$ produced by each firm as input to produce the composite commodity, which is directly consumed by the household. The retailer minimizes the total cost of producing the composite commodity,

$$\max_{Y_{i,t}} \int_0^1 P_{i,t} Y_{i,t} di$$

subject to the Dixit and Stiglitz (1977) aggregator,

$$Y_t \leq \left[\int_0^1 Y_{i,t}^{\frac{\eta-1}{\eta}} di\right]^{\frac{\eta}{\eta-1}}$$

The parameter η represents the elasticity of substitution between any two differentiated good $Y_{i,t}$ and $Y_{j,t}$. Defining the Lagrangian multiplier of the retailer's problem as the average of all the $P_{i,t}$, the price level, we have the condition for the demand of $Y_{i,t}$

$$Y_{i,t} = \left(\frac{P_{i,t}}{P_t}\right)^{-\eta} Y_t$$

The demand for $Y_{i,t}$ is a decreasing function in $P_{i,t}$ and an increasing function in P_t and Y_t.

3.1.4 Closing the model

The central bank sets the nominal interest rate i_t following a standard Taylor rule,

$$(1 + i_t) = (1 + i_{t-1})^{\rho_i} \left[(1 + i_{ss}) \left(\frac{\pi_t}{\pi_{ss}}\right)^{\psi_\pi}\right]^{1-\rho_i}$$

where i_{ss} and π_{ss} represent the steady-state values of nominal interest rate i_t and inflation π_t. The parameter ρ_i, ψ_π governs the degree of nominal interest rate persistent (interest rate smoothing) and how much the central bank wants to stabilize the inflation, Taylor coefficient.

In addition, because the price adjustment cost exists, the market clear condition for output is

$$Y_t = C_t + \frac{\gamma_p}{2} \left(\frac{P_t}{P_{t-1}} - \pi_{ss}\right)^2 Y_t$$

This condition can be derived from the household budget constraint (3) and the firm's constraint on dividends (5). We replace individual price with the aggregate price level, this is because, in the equilibrium, every firm sets the same price.

Lastly, we have two shocks in the model, total factor productivity shock to future total factor productivity A_t, preference shock to the variable that governs the inter-temporal preference variable b_t. The corresponding law of motion are,

$$\ln A_t = \rho_A \ln A_{t-1} + \sigma^A \epsilon_t^A$$
$$\ln b_t = \rho_b \ln b_{t-1} + \sigma^b \epsilon_t^b$$

where ϵ_t^A, ϵ_t^b are a technology shock that can be considered as the supply shock, and a preference shock that can be considered as the demand shock. In addition, ρ_A and ρ_b govern the persistent of the above four process, and σ^A and σ^b govern the variance of the four shocks.

3.2 Calibration

In this section, we numerically solve the model to see the model-implied impulse response functions and use these impulse response functions as the benchmark to check how well the methodology works. We solve the model to a first-order approximation of the policy functions to get the state-space representation of the model: all the jump variables are the linear function of the state variables. Then we hit the economy with one standard deviation of each shock to get the impulse response function of that shock. The impulse response functions are the percentage deviation from the corresponding deterministic steady-state value.

Table 2: Model's parameter values

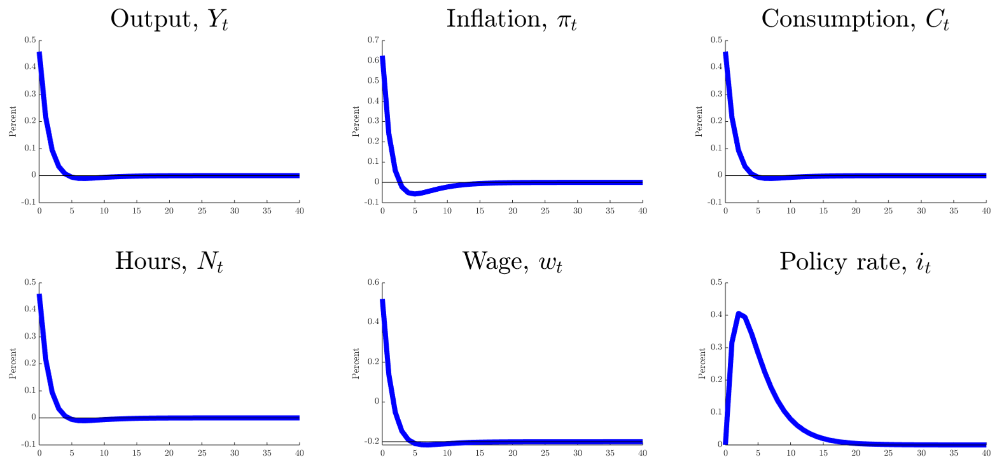
Param.	Interpretation	Value	Source
β	Deterministic discount factor	0.99	$R_{ss} = 3\%$
α	DRS parameter	0.2	Gilchrist et al. (2017)
σ_c	CRRA in consumption	2	Weber (1975)
σ_n	Frisch elasticity	2	Smets and Wouters (2007)
η	Elasticity of substitution	2	Broda and Weinstein (2006)
ψ_π	Taylor coefficient on π gap	5	Fasani and Rossi (2018)
γ_p	Level of nominal rigidity	10	Ramey (2016)
π_{ss}	Steady state value of π	1	N.A.
i_{ss}	Steady state value of i	0.0101	$\beta = 0.99$
ψ	Utility from leisure	1.3209	$N_{ss} = 0.33$
ρ_i	Monetary policy inertia	0.9	Smets and Wouters (2005)
ρ_A	Persistent of ϵ^A	0.75	Rabanal and Rubio-Ramírez (2005)
ρ_b	Persistent of ϵ^b	0.75	Rabanal and Rubio-Ramírez (2005)
σ^A	Standard deviation of ϵ^A	0.05	Rabanal and Rubio-Ramírez (2005)
σ^b	Standard deviation of ϵ^b	0.05	Rabanal and Rubio-Ramírez (2005)

Notes. β is the deterministic discount factor; α is the degree of decreasing return to scale in the production function; σ_c is the constant relative risk aversion coefficient in the consumption bundle; σ_n is the Fricsh elasticity of labor; η governs the elasticity of substitution between differentiated goods; ψ_π how the central bank responds to the inflation gap; γ_p is the strength of price adjustment cost; π_{ss} is the steady state value of inflation and it is 1; i_{ss} is the steady state value of policy rate and it is $1/\beta$; ψ governs the relative importance of consumption and leisure in generating utility; ρ_i governs the degree of nominal interest rate persistent in the Taylor rule; ρ_A and σ^A are the persistence and the variance of total factor productivity shocks ϵ_t^A; ρ_b and σ^b are the persistence and the variance of preference shocks ϵ_t^b.

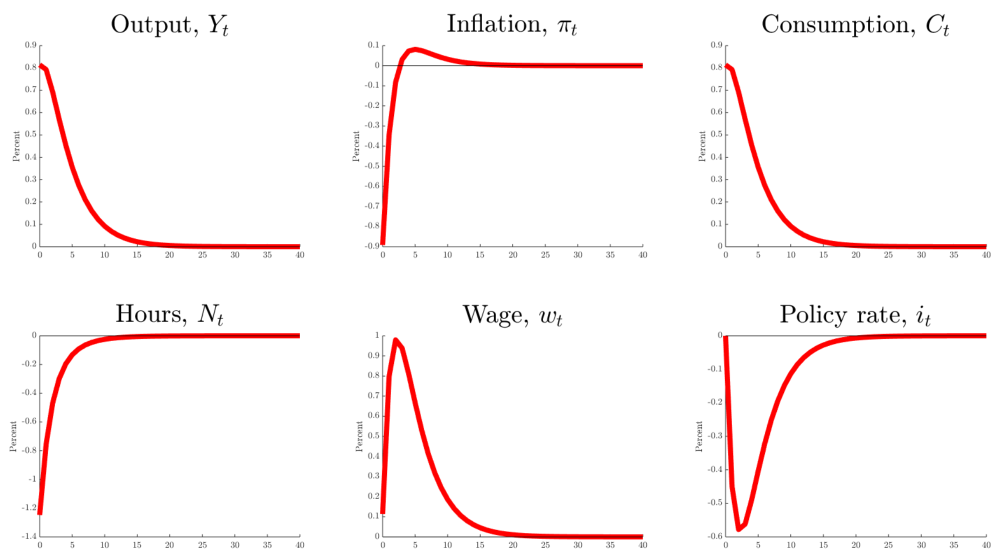
In the calibration part, we calibrate the deterministic discount factor β to match 3% annual yield on bonds. For other parameters, we list the resources of the calibration except for ψ, the utility weighting between consumption and leisure. To calibrate this parameter, we set the steady-state value of labor to $1/3$, which reflects 8-hours of working takes one-third of a day. Thus we can solve for the value of ψ using other steady-state values and parameters. Table 2 summarize the parameter values.

Figure 1: Model-implied impulse responses

(a) Demand shocks



(b) Supply shocks



Notes. Model-implied responses to a technology shock and the preference shock. Model's parameter are presented in Table 1.

The following figure shows the model implied impulse response. Figure 1(a) displays responses to a demand shock (inter-temporal shock) triggers 0.45% increase in output and consumption. Real wages also rise due to a large drop in λ_t, which is the wedge between marginal productivity of labor and the marginal cost of labor. Because of the increases in real wages, leisure becomes more expensive compared to commodity, so working hours increases and thus creating the co-movement between different components of the output. Inflation also sees an increase in response. Following the increases in the price level, the interest rate increases to stem the inflationary forces (Smets and Wouters, 2003). Figure 1(b) displays responses to a supply shock (total factor productivity shock). Output increases to the supply shock because the production is now more efficient. Working hours negatively respond to the supply shock, and inflation also negatively responds to the supply shock. The fall in hours is consistent with estimated impulse responses of identified productivity shocks in the United States (Gali, 1999). The impulse responses are consistent with the aggregate demand-supply analysis framework, demand shock imposes a positive impact on both output and inflation and supply shock imposes a positive impact on output and a negative impact on inflation. These different sets of responses can be used by our methodology to identify demand and supply shocks.

3.3 Model simulation

In the previous section, we obtain the model-implied impulse responses of output, inflation, consumption, working hours, wage rate, and policy rate to demand and supply shocks. This section introduces the model simulation process, which we use to get the simulated data from the model. Using the simulated data, we can test our identification methodology. By comparing the impulse response to the identified shocks and model-implied shocks, we know how our strategy works.

The simulation has two parts. In the first part, we use the baseline model introduced in the previous section to simulate a "long economy" worked as the population. We set the length of the population to $100,000$, this is just a large number. We start from the steady-state, randomly drawing demand and supply shocks from two independent (across time and shocks) standard normal distributions and feeding them to the model. Given the transition matrix and policy matrix we get from the first-order approximation of the model, we can compute the percentage deviation for each variable in the model in one period. We then repeat the process until the number of observations in the population is satisfied. The second part uses the baseline model to simulate the $2,000$ "short economies" working as the random samples drawn from the population. Each sample has 240 observations. We set the length of each sample to 240 to match the empirical application in the Section 4, where we use real U.S. data to identify demand and supply shocks. In this part of the simulation study, for each sample, we re-estimate a reduced-form VAR model with the optimal lag order suggested by AIC. Then use the strategy to identify the demand and supply shocks and then calculate the impulse responses. The second part of the simulation helps to construct the confidence interval of the estimated impulse response functions.

For the identification procedure in this section, we use a slightly different setup. Consider a bi-variate SVAR model has the following reduced-from

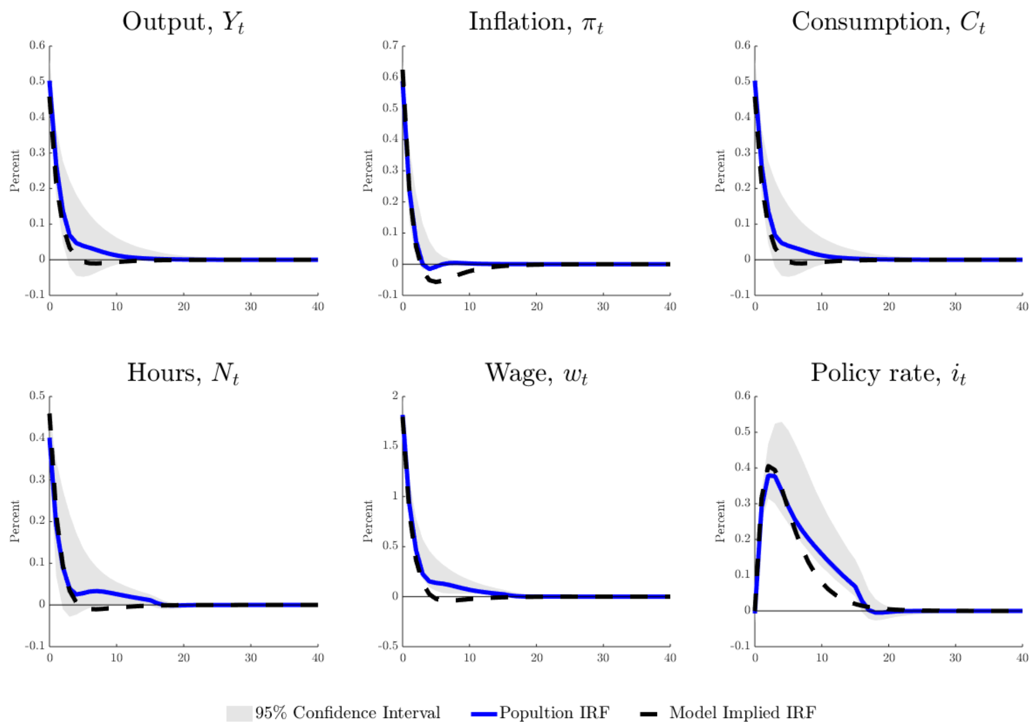
$$X_t = B(L)X_{t-1} + u_t$$

where $X_t = [Y_t, \pi_t]'$. In addition, we know demand shock imposes a positive impact on output and inflation, and supply shock imposes a positive impact on output and a negative impact on inflation. Thus, we can use the methodology described in the Section 2 to identify the demand and supply shock. Decompose the impact matrix $A = CD$ where C is the Cholesky decomposition of $\Sigma = (u_t u_t')$ and $D = [d_D, d_S]$ is an orthogonal matrix where:

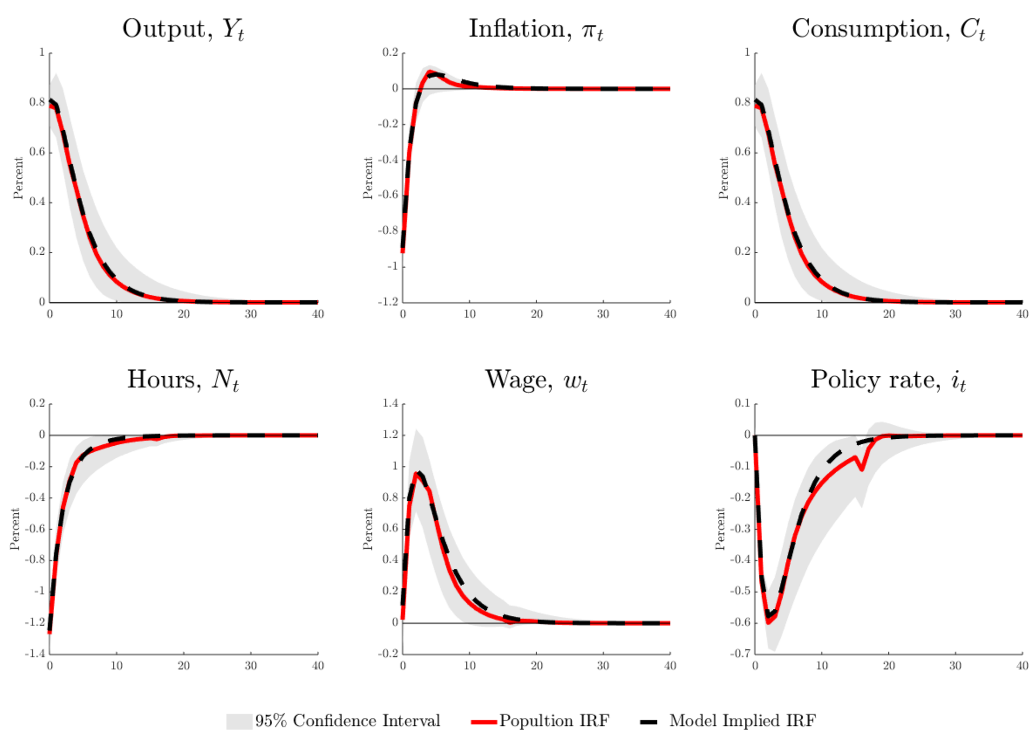
1. Column vector d_D is the solution of the following problem,
$$d_D = \arg\max_{d_1} \{(1-\delta^*)e_1 C d_1 + \delta^* e_2 C d_1, \text{ subject to } d_1' d_1 = 1\};$$
2. Column vector d_S is the solution of the following problem,
$$d_S = \arg\max_{d_2} \{(1-\delta^*)e_1 C d_2 - \delta^* e_2 C d_2, \text{ subject to } d_S' d_S = 1\};$$
3. $\delta^* \in (0, 1)$ is such that $d_D' d_S = 0$.

Figure 2: Model-implied and Estimated impulse responses

(a) Demand shocks



(b) Supply shocks

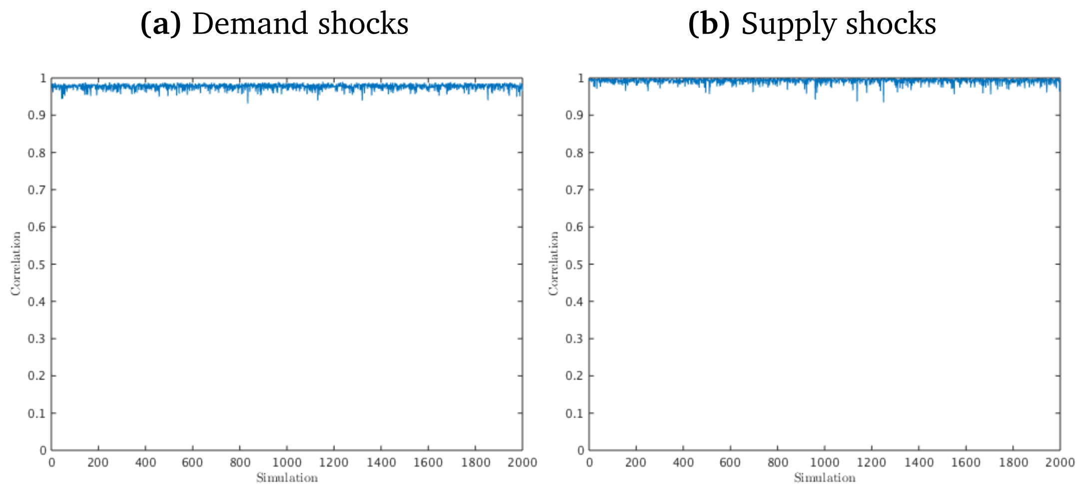


Here d_D and d_S are corresponding the demand and the supply shock. The SVAR system only contains two variables: the output and the inflation. We are not including other variables given that we only have two shocks. Putting more than two variables in the SVAR will cause a collinearity problem. This bi-variate VAR gives us the impulse responses of the output and inflation. Then, we augment the baseline bi-variate VAR with a set of auxiliary variables, S_t, including consumption, hours, real wage, and interest rate proposed by Basu et al. (2021). The auxiliary regression is regressing each of the variables in S_t to the current and lagged values of variables in the baseline bi-variate VAR X_t,

$$s_t = A(L)X_t + v_t$$

For each variable s_t in S_t, we allow a different optimal lag order suggested by AIC. Using this relation, we can recover the impulse responses of S_t from the impulse response of X_t. Figure 2 shows the model-implied and estimated (from the population) impulse responses to two shocks with the 95% confidence interval of the estimated impulse responses using $2,000$ samples. From the impulse response graph, we can see that our methodology works pretty well. The black dashed lines and colored lines are aligned with each other and are in the confidence interval. Two things are worth mentioning. First, in calculating the confidence interval for each sample, we run the methodology and get a δ^* that guarantees the impact matrix is an orthogonal matrix. We checked for every simulated economy, the δ^* is exactly the same with the theoretical one proposed by Proposition 1. Second, we can see a wiggle in the policy rate's response to the supply shock. The reason is that we only set the maximum lag order for the auxiliary regression that recovers the impulse response of the policy rate to 16. The spike happens to be at the period 16. If we set the maximum lag order to 40, there will be no more spikes, but the identification takes longer.

Figure 3: Correlation between actual shocks ϵ_t and estimated shocks $\hat{\epsilon}_t$



We also calculate the correlation coefficient of the actual shock series and estimated shock series from our econometrician methodology. The actual shock series are random draws from the normal distribution in the described data simulation process. The estimated shock series are calculated by $\hat{\epsilon}_t = \hat{A}^{-1}\hat{u}_t$, where $\hat{u}_t$ are the estimated residuals from the reduced-form VAR. For the population, the correlation of the actual demand shock ϵ_t^b and the estimated demand shock $\hat{\epsilon}_D$ is 99.07% and the correlation of the actual supply shock ϵ_t^A and the estimated supply shock $\hat{\epsilon}_S$ is 99.86%. For the samples, the confidence interval of the correlation of the actual demand shock ϵ_t^b and the estimated demand shock $\hat{\epsilon}_D$ is $[95.90\%, 98.59\%]$ and the confidence interval of the correlation of the true supply shock ϵ_t^A and the estimated supply shock $\hat{\epsilon}_S$ is $[97.50\%, 99.90\%]$.

3.4 Robustness check

In order to test the reliability, we provide a set of extensions to the baseline specification that is useful to inform on the robustness of the methodology. We extend the baseline specification in the following ways: (i) We set the $\alpha = 0$ in the baseline specification, which means the production exhibits a constant return to scale; (ii) We also add the external habit consumption to the model. Habit formation implies the inter-temporal non-separability of consumption (Duesenberry, 1949; Pollak, 1970; Abel, 1990). The functional form is quasi-difference in consumption

$$U(C_t - \phi_c C_{t-1}, 1 - N_t)$$

where we set the ϕ_c, the habit persistence parameter, to 0.9 (Alessie and Lusardi, 1997). The model with habit persistence replaces the level of consumption with consumption growth in the utility function; (iii) We use another form of nominal rigidity setting, Calvo price a la Calvo (1983). In Calvo pricing, every firm will be chosen with probability θ not allowed to change the price. We set $\theta = 0.74$ (Smets and Wouters, 2007). Aggregately, given the continuum of firms in the economy, $\theta = 0.74$ means means there will be 26% of firms in each period that can change their price.

Table 3 presents the correlation coefficient between the actual shocks and estimated shocks for each extension. We report the correlation for the population, and for simulated samples, we report the range of the correlation together with the median correlation for each specification. In every extension, the correlation coefficient for the population is above 98.5% for both demand and supply shocks. The correlations are above 95% for every extension for both demand and supply shocks for the samples. In all specifications, the estimated and actual shocks are highly correlated, indicating the identification method works well and does not depend on a special functional form of the model.

Table 3: Correlation for other specifications

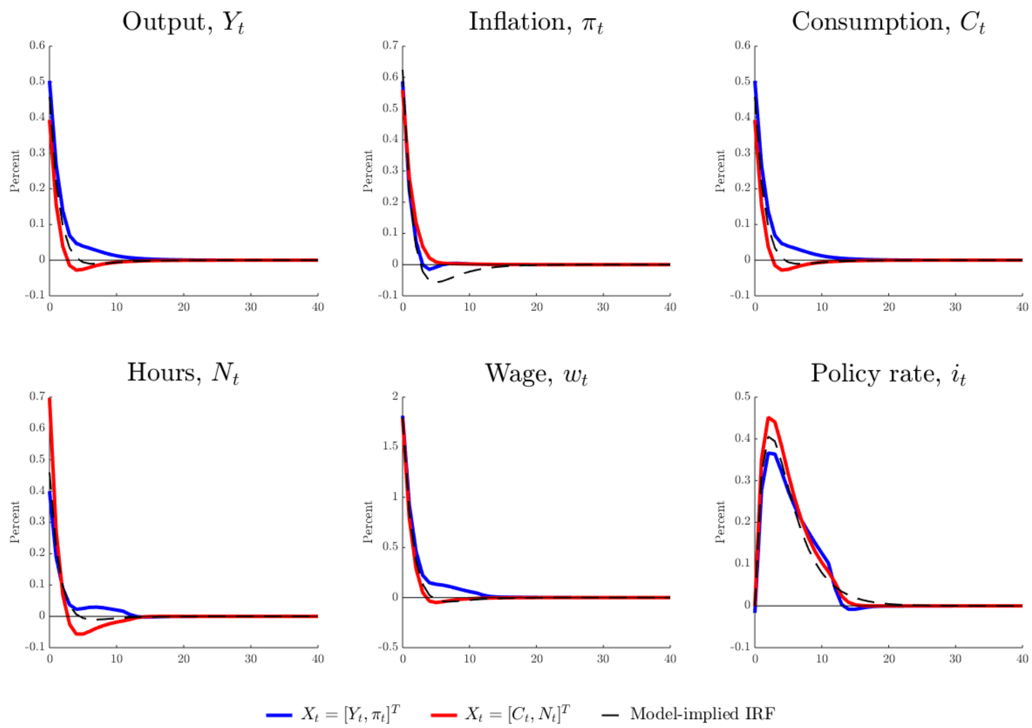
Model	Demand		Supply	
	Population	Simulated Samples	Population	Simulated Samples
(1) Baseline	99.07%	97.87% [95.90%, 98.59%]	99.86%	99.49% [97.50%, 99.90%]
(2) CRS production	99.2%	98.06% [96.30%, 98.79%]	99.93%	99.57% [97.79%, 99.94%]
(3) Habit Formation	99.15%	98.49% [96.75%, 99.14%]	98.90%	98.07% [96.05%, 99.04%]
(4) Calvo pricing	99.32%	98.50% [96.86%, 99.12%]	99.84%	99.45% [97.67%, 99.90%]

Notes. For the simulated samples, the top line is the median correlation between the estimated shocks and actual shocks. The bottom line is the range of the correlation between the estimated shocks and actual shocks. Model (1) is the baseline model, the simplest New Keynesian model described in this section. In Model (2), we set $\alpha = 0$, the production function turns to constant return to scale. Model (3) adds external habit consumption to the model. And Model (4) replaces the Rotemberg price adjustment cost with the Calvo pricing to represent nominal rigidity.

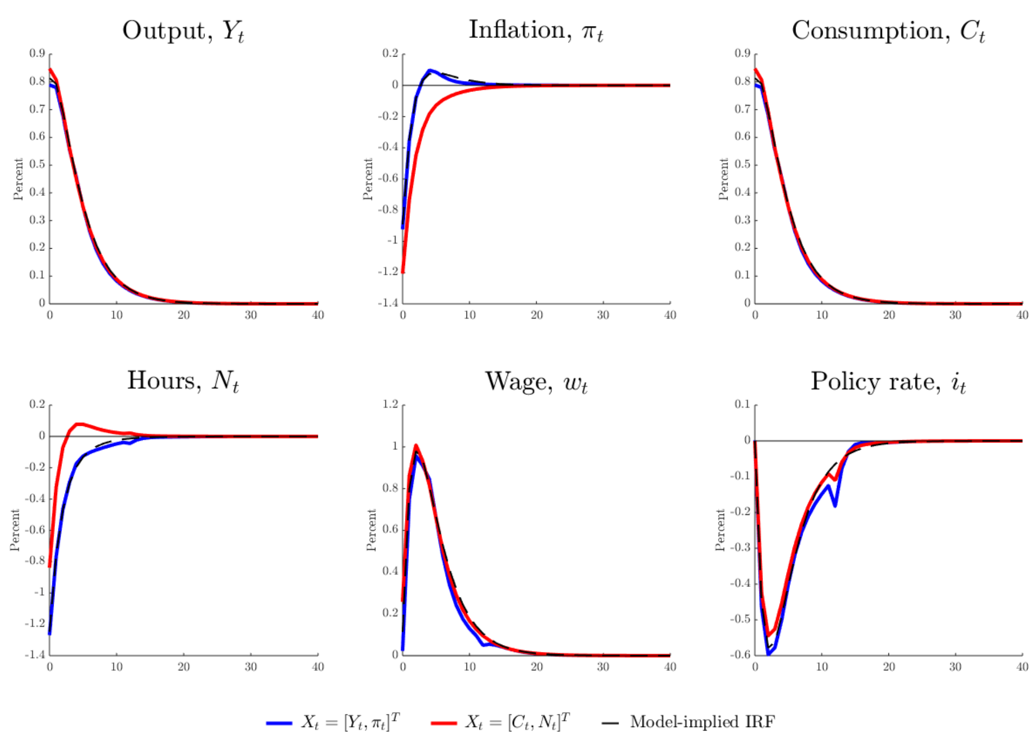
We also run another robustness test to check the reliability and stability of the methodology. Suppose we have two sets of information about the same shocks, will our method produce the same impulse responses and the same estimated shocks? From the model-implied impulse responses in the baseline specification, we can see that the consumption positively responds to the demand and supply shocks; however, the hours positively responds to the demand shock but negatively responds to the supply shock. The responses are different for these shocks. So we can use this as another set of information to identify shocks. This set of information is shown in the baseline model-implied impulse responses and it is our baseline model specific and less theory grounded. The impulse responses of consumption and hours to the demand and supply shocks may vary for other model specifications, we use this information only to test the reliability of our method. Using information regarding the output and inflation for identification is more reliable. In doing that, we set $X_t = [C_t, N_t]'$ for the SVAR and use our identification strategy to get the impulse response of the consumption and hours.

Figure 4: Comparison of different identifications

(a) Demand shocks



(b) Supply shocks



Similarly, we use the auxiliary regression to recover the impulse responses of output, inflation, real wages, and the policy rate. We can see from Figure 4 that no matter which set of information is used for identification, the results are pretty consistent (the solid red and blue lines). Figure 4 also shows two specifications and the model-implied impulse responses are aligned with each other. Similarly, we also compute the correlation between the estimated and the actual shocks. Both of the estimated shocks are highly correlated with the actual counterparts. In the setting $X_t = [Y_t, \pi_t]'$, the correlation for the demand and supply shocks are 99.07% and 99.86% and in the setting $X_t = [C_t, N_t]'$, the correlation for the demand and supply shocks are 99.61% and 99.67%. Last, the correlation coefficient for two estimated demand shocks is 98.71%, and for two estimated supply shocks is 99.15%. We can conclude from both the impulse responses and correlations that our method is reliable and stable. It can recover the true shocks no matter which information is used.

4 Application on US data

This section simultaneously identifies demand and supply shocks on U.S. aggregate data using the identifying assumption and the econometric strategy presented in the section 2

4.1 Baseline specification and main results

In the baseline specification, we estimate a reduced-form VAR with (i) log-transformation of real GDP; (ii) CPI inflation rate calculated by the difference of log CPI; (iii) log-transformation of the real investment level defined as the sum of the real consumption of durable goods and the real investment; (iv) log-transformation of the real consumption defined as the real consumption of non-durable goods plus the real services; (v) the log-transformation of total hours as hours of all persons in the non-farm business sector; (vi) the shadow federal funds rate (FFR) by Wu and Xia (2016). In order to focus on the non-Covid-19 periods, the data range is from the second quarter of 1960 to the first quarter of 2020, and reduced-form innovations are obtained by controlling for two lag of all the variables in the system with intercepts in the system. For the estimation, we also apply the bias-correction algorithm proposed by Kilian (1998) to correct for the bias when including an intercept in the VAR model for a given sample size.

$$Y_t = c + \sum_{i=1}^{4} B_i Y_{t-i} + u_t$$

We also present the 90% and 95% confidence intervals. To do so, we use the recursive-design residual bootstrap. For each simulation, we randomly select an observation from the data set and use this random draw as the start point. We then use the estimated baseline regression and this start point to construct the next period observation until the observation numbers 240 are met,

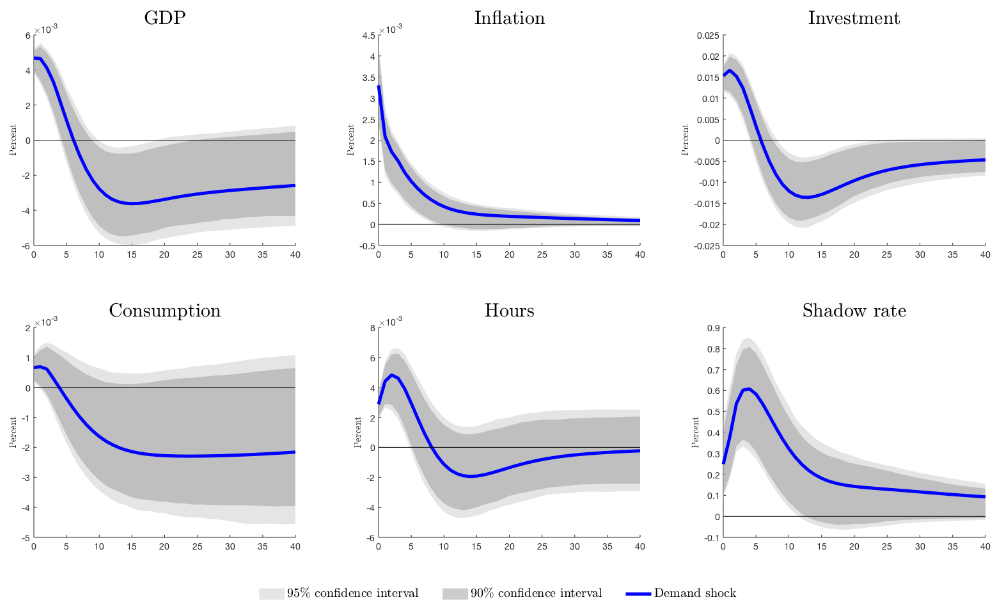
$$Y_t^s = \hat{c} + \sum_{i=1}^{4} \hat{B}_i Y_{t-i}^s + \hat{u}_t$$

where Y^s is the simulated observations, and $\hat{u}_t$ is a random draw from the residuals in the baseline regression. Once we get a bootstrap sample, we re-estimate the model and use the identification strategy introduced in Section 2 to estimate the demand and supply shock and calculate the corresponding impulse responses.

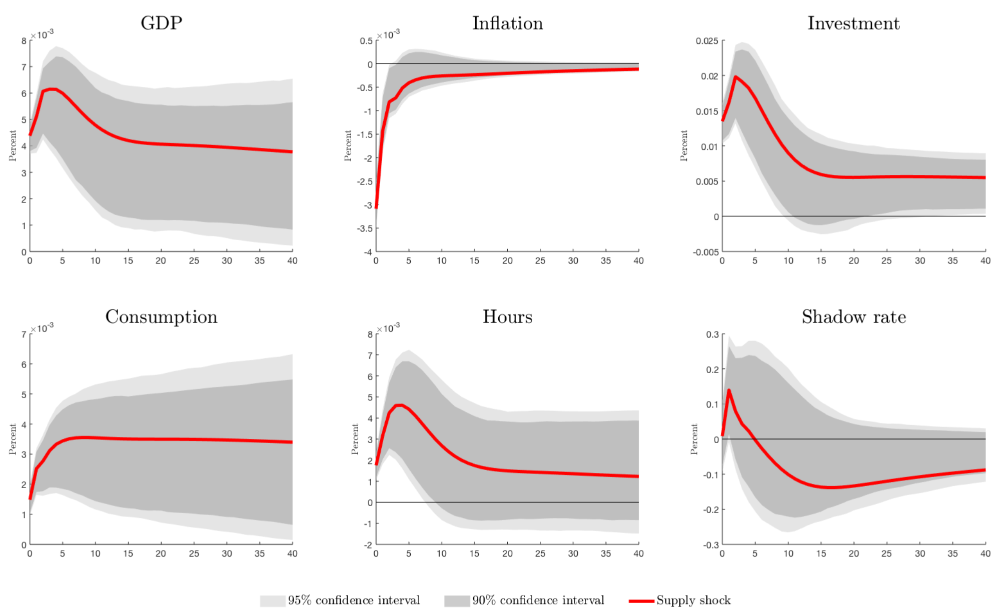
Figure 5(a) shows responses to a standard deviation demand shock. First, both GDP and inflation display a positive and significant initial impact response. However, they behave differently afterward. GDP sees a positive impact response to the demand shock, and the impact declines gradually and remains positive for about one year. After that, the impact turns negative and dies out after six years, returning to its steady state. Inflation's impact remains significantly positive for two and half years and then gradually returns to its steady-state. The inflation's impact is more persistent than GDP's. The initial inflation jump suggests that demand shocks are associated with an inflationary force. The shadow rate displays a positive response to the demand shock, and it remains significant for three years. It shows the monetary authority rise interest rate responding to the inflationary forces in the positive demand shock. Investment, consumption, and total hours exhibit an initial positive impact response and then negative and return to its steady-state value in the end. The positive impact remains less than one year and a half. The consumption seems to be the last responsive one.

Figure 5: Estimated impulse responses on U.S. aggregate data

(a) Demand shocks



(b) Supply shocks



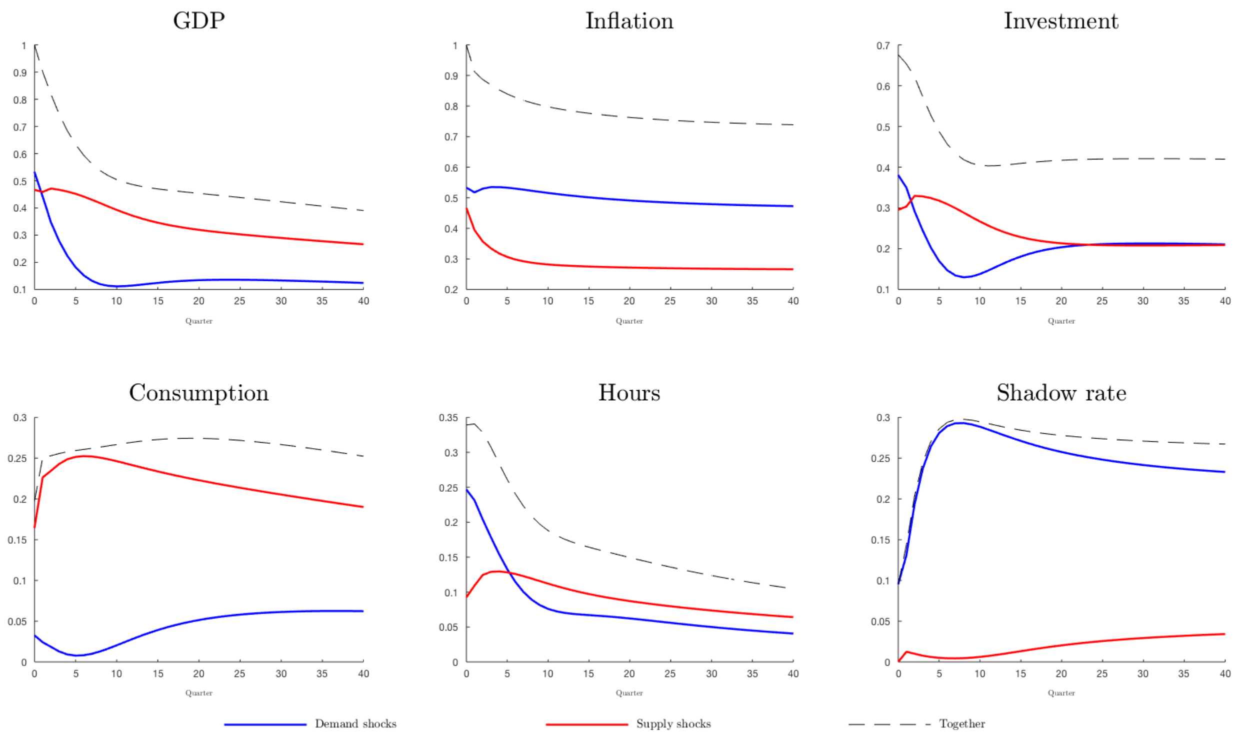
Notes. Data range: 1960:q2-2020:q1. reduced form VAR has two lag (AIC). Using bootstrap to calculate the impulse responses.

Figure 5(b) shows responses to a standard deviation supply shock. We see that GDP displays a positive and significant impact response to the supply shock. The impact response of GDP is very persistent lasts more than ten years. However, inflation displays a negative impact response to the supply shock, and it is significant only lasts about one year. After one year, the impact response dies out soon, and inflation returns to the steady-state value. We can conclude that the supply shock is associated with a deflationary force. Besides, both investment and consumption display positive impact responses to the supply shock and impacts are significant and persistent remains for over ten years. Total hours display a positive response but are less persistent. It dies out after two years. The shadow rate initially exhibits a positive response and turns negative afterward, but insignificant.

Comparing the response of demand and supply shocks, we have other findings. First, GDP is less persistent than inflation in response to the demand shock, whereas the GDP is more persistent than inflation to the supply shock. This findings is consistent with the assumption used in long run identification of demand and supply shocks by Blanchard and Quah (1988). Second, we see a cycle pattern in the system responding to the demand shock, but there is no clear cycle pattern in the system responding to the supply shock. The last empirical findings echos Brianti and Cormun (2019), the fundamental-driven fluctuations (such as technology shocks) do not generate boom-bust dynamics, whereas positive shifts in expectations do so.

Figure 6 exhibits the forecast error variance decomposition of the endogenous variables in the system explained by demand shock (solid blue lines), supply shock (solid red lines), and the two shocks together (dashed lines). In the short run, the demand shock explains more than 50% of the unexpected fluctuation in GDP, and this effect gradually declines to 10% in the long run. It also explains over half of the unexpected fluctuation of inflation for all ten years. The demand shock explains about 40% of real investment and roughly 25% of total hours in the short run and then declines to 20% for real investment and 10% for total hours. The consumption seems to be least affected by demand shocks (around 5%) in the short and the long run. Finally, the demand shock explains 30% of shadow rate over one year and a half and remains at basically the same level in the long run. In contrast, GDP is more affected by the supply shocks (over 40%), but inflation is less affected by the supply shocks (initially 45% but 25% in the long run). In the short run, the supply shocks do not have a remarkable quantitative effect on total hours (explains less than 15% of variation for hours), and it gradually declines to 10% after two years. In addition, the supply shocks explain almost 30% of real investment variation in the short and 20% the long run. The supply shock explains 20% of consumption variation over the ten-year horizon but up to 5% of shadow rate over the same horizon. This result is consistent with the finding that the response to supply shocks is more persistent for the output than the inflation. However, for the demand shock, the persistence to shocks is reversed.

Figure 6: Forecast error decomposition



Notes. Variance decomposition for the reduced-form VAR model. The lag order is 2 (AIC).

In summary, the analysis suggests three main conclusions. First, both shocks have sizable expansionary effects on macroeconomic variables such as GDP, investment, and total hours. Also, demand shock seems to be a stronger driver of boom-bust cycles. Second, inflation displays qualitatively different responses to demand and supply shocks. Also, demand shock is more critical in predicting the monetary policy. Third, GDP and inflation's response to demand and supply shock exhibit different persistence. This can be shown by the impulse response functions and forecast error variance decomposition.

5 Conclusions

This paper proposed a new identification strategy to identify the aggregate demand and supply shocks. We used a simple New Keynesian model to test the method's reliability.

Using this model to simulate data and apply the method, we find that the model-implied impulse responses and impulse responses obtained from the identification strategy are aligned with each other. Also, we calculated the correlation between real shock series and estimated shocks, the correlation could reach 96% for both aggregate demand shocks and aggregate supply shocks. We also show that this methodology is robust to different models. We add the model with some features and new the modified model to test the reliability of the model, we show that the correlation between the actual shock series and the estimated shock series is highly positively correlated.

We also apply this method to the U.S. data, the results show that both shocks have sizable expansionary effects on macroeconomic variables such as GDP, investment, and total hours. For GDP and inflation, over 50% of the forecast error can be explained by these shocks over the long run (ten years). The demand shock seems to be a stronger driver of boom-bust cycles than the supply shock. Last, GDP is less persistent than inflation in response to the demand shock, whereas the GDP is more persistent than inflation to the supply shock.

References

ABEL, A. B. (1990): "Asset prices under habit formation and catching up with the Joneses," .

ALESSIE, R. AND A. LUSARDI (1997): "Consumption, saving and habit formation," *Economics letters*, 55, 103–108.

BASU, S., G. CANDIAN, R. CHAHROUR, AND R. VALCHEV (2021): "Risky business cycles," Tech. rep., National Bureau of Economic Research.

BAUMEISTER, C. AND J. D. HAMILTON (2015): "Sign restrictions, structural vector autoregressions, and useful prior information," *Econometrica*, 83, 1963–1999.

BLANCHARD, O. J. AND D. QUAH (1988): "The dynamic effects of aggregate demand and supply disturbances," .

BRIANTI, M. (2021): *Financial Shocks, Uncertainty Shocks, and Monetary Policy Trade-Offs*, University of Alberta, Faculty of Arts, Department of Economics.

BRIANTI, M. AND V. CORMUN (2019): "What are the Sources of Boom-Bust Cycles?" .

BRODA, C. AND D. E. WEINSTEIN (2006): "Globalization and the Gains from Variety," *The Quarterly journal of economics*, 121, 541–585.

CALDARA, D., C. FUENTES-ALBERO, S. GILCHRIST, AND E. ZAKRAJŠEK (2016): "The macroeconomic impact of financial and uncertainty shocks," *European Economic Review*, 88, 185–207.

CALVO, G. A. (1983): "Staggered prices in a utility-maximizing framework," *Journal of monetary Economics*, 12, 383–398.

DIXIT, A. K. AND J. E. STIGLITZ (1977): "Monopolistic competition and optimum product diversity," *The American economic review*, 67, 297–308.

DUESENBERRY, J. S. (1949): *Income, saving, and the theory of consumer behavior*.

FASANI, S. AND L. ROSSI (2018): "Are uncertainty shocks aggregate demand shocks?" *Economics Letters*, 167, 142–146.

FAUST, J. (1998): "The robustness of identified VAR conclusions about money," in *Carnegie-Rochester conference series on public policy*, Elsevier, vol. 49, 207–244.

FRY, R. AND A. PAGAN (2011): "Sign restrictions in structural vector autoregressions: A critical review," *Journal of Economic Literature*, 49, 938–60.

FURLANETTO, F., F. RAVAZZOLO, AND S. SARFERAZ (2019): "Identification of financial factors in economic fluctuations," *The Economic Journal*, 129, 311–337.

GALI, J. (1992): "How well does the IS-LM model fit postwar US data?" *The Quarterly Journal of Economics*, 107, 709–738.

——— (1999): "Technology, employment, and the business cycle: do technology shocks explain aggregate fluctuations?" *American economic review*, 89, 249–271.

GILCHRIST, S., R. SCHOENLE, J. SIM, AND E. ZAKRAJŠEK (2017): "Inflation dynamics during the financial crisis," *American Economic Review*, 107, 785–823.

INOUE, A. AND L. KILIAN (2022): "Joint Bayesian inference about impulse responses in VAR models," *Journal of Econometrics*, 231, 457–476.

KILIAN, L. (1998): "Small-sample confidence intervals for impulse response functions," *Review of economics and statistics*, 80, 218–230.

LIPPI, M. AND L. REICHLIN (1993): "The dynamic effects of aggregate demand and supply disturbances: Comment," *The American Economic Review*, 83, 644–652.

POLLAK, R. A. (1970): "Habit formation and dynamic demand functions," *Journal of political Economy*, 78, 745–763.

RABANAL, P. AND J. F. RUBIO-RAMÍREZ (2005): "Comparing New Keynesian models of the business cycle: A Bayesian approach," *Journal of Monetary Economics*, 52, 1151–1166.

RAMEY, V. A. (2016): "Macroeconomic shocks and their propagation," *Handbook of macroeconomics*, 2, 71–162.

ROTEMBERG, J. J. (1982): "Sticky prices in the United States," *Journal of political economy*, 90, 1187–1211.

RUBIO-RAMIREZ, J. F., D. F. WAGGONER, AND T. ZHA (2010): "Structural vector autoregressions: Theory of identification and algorithms for inference," *The Review of Economic Studies*, 77, 665–696.

SMETS, F. AND R. WOUTERS (2003): "An estimated dynamic stochastic general equilibrium model of the euro area," *Journal of the European economic association*, 1, 1123–1175.

——— (2005): "Comparing shocks and frictions in US and euro area business cycles: a Bayesian DSGE approach," *Journal of Applied Econometrics*, 20, 161–183.

——— (2007): "Shocks and frictions in US business cycles: A Bayesian DSGE approach," *American economic review*, 97, 586–606.

UHLIG, H. (1998): "The robustness of identified VAR conclusions about money: A Comment," in *Carnegie-Rochester Conference Series on Public Policy*, North-Holland, vol. 49, 245–263.

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

WEBER, W. E. (1975): "Interest rates, inflation, and consumer expenditures," *The American Economic Review*, 65, 843–858.

WU, J. C. AND F. D. XIA (2016): "Measuring the macroeconomic impact of monetary policy at the zero lower bound," *Journal of Money, Credit and Banking*, 48, 253–291.

Appendix

A Proof of Proposition 1

Proof. Let's consider the bivariate case[1] with the reduced-form innovation has the following covariance matrix:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

where σ_1^2 is the variance of the reduced-form innovation in the first equation (i.e., the output equation), σ_2^2 is the variance of the reduced-form innovation in the second equation (i.e., the inflation equation), and σ_{12} is the covariance between these two reduced-form innovations.

Then the Cholesky decomposition of Σ is:

$$CC^{'} = \begin{pmatrix} c_{11} & 0 \\ c_{12} & c_{22} \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} \\ 0 & c_{22} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

The solution system is given by:

$$\begin{cases} c_{11}^2 = \sigma_1^2 \\ c_{11}c_{21} = \sigma_{12} \\ c_{21}^2 + c_{22}^2 = \sigma_2^2 \end{cases} \Rightarrow \begin{cases} c_{11} = \sigma_1 \\ c_{21} = \dfrac{\sigma_{12}}{\sigma_1} = \rho\sigma_2 \\ c_{22} = \sigma_2\sqrt{1-\rho^2} \end{cases}$$

where ρ is the correlation coefficient between two reduced-form innoviations. We focus on the positive solutions[2]. Define the orthogonal matrix D as:

$$D = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} = [d_D, d_S]$$

[1] In the Section 3, we apply the method on a bivariate VAR model $X_t = [Y_t, \pi_t]^{'}$ to identify the demand and supply shock.

[2] We have three other sets of solution: $\tilde{c}_{11} = \sigma_1$, $\tilde{c}_{12} = \sigma_{12}/\sigma_1$, $\tilde{c}_{22} = -\sqrt{\sigma_2^2 - \left(\frac{\sigma_{12}}{\sigma_1}\right)^2}$; $\tilde{c}_{11} = -\sigma_1$, $\tilde{c}_{12} = -\sigma_{12}/\sigma_1$, $\tilde{c}_{22} = \sqrt{\sigma_2^2 - \left(\frac{\sigma_{12}}{\sigma_1}\right)^2}$; and $\tilde{c}_{11} = -\sigma_1$, $\tilde{c}_{12} = -\sigma_{12}/\sigma_1$, $\tilde{c}_{22} = -\sqrt{\sigma_2^2 - \left(\frac{\sigma_{12}}{\sigma_1}\right)^2}$. But these solutions lead to the same identification

then impact matrix $A = CD$ is:

$$A = [Cd_D, Cd_S] = \begin{pmatrix} \sigma_1 d_{11} & \sigma_1 d_{12} \\ d_{11}\rho\sigma_2 + d_{21}\sigma_2\sqrt{1-\rho^2} & d_{12}\rho\sigma_2 + d_{22}\sigma_2\sqrt{1-\rho^2} \end{pmatrix}$$

Notice that the Problem (1) is equivalent to:

$$\begin{aligned} d_D &= \arg\max_{d} \{e_1 Cd + \theta e_2 Cd, \ \text{ subject to } \ dd^{'} = 1\} \\ &= \arg\max_{d_{11}, d_{21}} \left\{\sigma_1 d_{11} + \theta\left(d_{11}\rho\sigma_2 + d_{21}\sigma_2\sqrt{1-\rho^2}\right), \ \text{ subject to } \ dd^{'} = 1\right\} \end{aligned}$$

where $\theta = \delta^{-1} - 1 \in (0, +\infty)$, since we restrict $\delta \in (0, 1)$. Let λ be the Lagrangian multiplier of the equality constraint in Problem (1) and the optimal conditions are:

$$\begin{aligned} \sigma_1 + \theta\rho\sigma_2 - 2\lambda d_{11} &= 0 \\ \theta\sigma_2\sqrt{1-\rho^2} - 2\lambda d_{21} &= 0 \\ d_{11}^2 + d_{21}^2 &= 1 \end{aligned}$$

Define $\Gamma = \sqrt{\theta^2\sigma_2^2 + 2\theta\rho\sigma_1\sigma_2 + \sigma_1^2}$, then we solve for d_D as

$$d_D = \begin{pmatrix} \frac{1}{\Gamma}(\sigma_1 + \theta\rho\sigma_2) \\ \frac{1}{\Gamma}\theta\sigma_2\sqrt{1-\rho^2} \end{pmatrix}$$

We know $\rho \in [-1, 1]$, then $\theta^2\sigma_2^2 + 2\theta\rho\sigma_1\sigma_2 + \sigma_1^2 \geq \theta^2\sigma_2^2 - 2\theta\sigma_1\sigma_2 + \sigma_1^2 = (\theta\sigma_2 - \sigma_1)^2 \geq 0$, and $\Gamma \in \mathbb{R}$. So d_D always exists and is only well defined if $\Gamma \neq 0$.

The Problem (2) is equivalent to:

$$\begin{aligned} d_S &= \arg\max_{d} \{e_1 Cd - \theta e_2 Cd, \ \text{ subject to } \ dd^{'} = 1\} \\ &= \arg\max_{d_{12}, d_{22}} \left\{\sigma_1 d_{12} - \theta\left(d_{12}\rho\sigma_2 + d_{22}\sigma_2\sqrt{1-\rho^2}\right), \ \text{ subject to } \ dd^{'} = 1\right\} \end{aligned}$$

where $\theta = \delta^{-1} - 1 \in (0, +\infty)$, since we restrict $\delta \in (0, 1)$. Let μ be the Lagrangian multiplier of the equality constraint in Problem (2) and the optimal conditions are:

$$\begin{aligned} \sigma_1 - \theta\rho\sigma_2 - 2\mu d_{12} &= 0 \\ -\theta\sigma_2\sqrt{1-\rho^2} - 2\mu d_{22} &= 0 \\ d_{12}^2 + d_{22}^2 &= 1 \end{aligned}$$

Define $\Delta = \sqrt{\theta^2\sigma_2^2 - 2\theta\rho\sigma_1\sigma_2 + \sigma_1^2}$, then we solve for d_S as

$$d_S = \begin{pmatrix} \frac{1}{\Delta}(\sigma_1 - \theta\rho\sigma_2) \\ -\frac{1}{\Delta}\theta\sigma_2\sqrt{1-\rho^2} \end{pmatrix}$$

We know $\rho \in [-1, 1]$, then $\theta^2\sigma_2^2 - 2\theta\rho\sigma_1\sigma_2 + \sigma_1^2 \geq \theta^2\sigma_2^2 - 2\theta\sigma_1\sigma_2 + \sigma_1^2 = (\theta\sigma_2 - \sigma_1)^2 \geq 0$, and $\Delta \in \mathbb{R}$. So d_S always exists and is only well defined if $\Delta \neq 0$.

Last, $d_D^{'} d_S = 0$ implies:

$$(\sigma_1 - \theta\rho\sigma_2)(\sigma_1 + \theta\rho\sigma_2) = \theta^2\sigma_2^2(1-\rho^2)$$

This is a quadratic equation in θ and it has at most two roots. We keep the positive root since we need $\theta > 0$. Thus $\theta^* = \sigma_1/\sigma_2$ is the unique solution on $(0, +\infty)$. Now we restrict $\rho \neq 1$ and $\rho \neq -1$ to guarantee that Γ and Δ are nonzero and thus d_D and d_S are both well defined. The unique θ^* translates to the unique δ^*:

$$\delta^* = \frac{\sigma_2}{\sigma_1 + \sigma_2}$$

■

B Extending the Proposition 1

Still consider the bi-variate VAR case $X_t = [Y_t, \pi_t]^{'}$. But now, we allow δ in Problem (1) and Problem (2) to be different: δ_1 and δ_2. Define $\theta_1 = \delta_1^{-1} - 1$, the Lagrangian function of Problem (1) is:

$$L = e_1 C d_D + \theta_1 e_2 C d_D - \lambda(d_D d_D^{'} - 1)$$

FOC implies:

$$d_D = \frac{C^{'}}{2\lambda}(e_1^{'} + \theta_1 e_2^{'})$$

Define $\theta_2 = \delta_2/(1-\delta_2)$, the Lagrangian function of problem (2) is:

$$L = \theta_2 e_1 C d_S - e_2 C d_S - \mu(d_S d_S^{'} - 1)$$

FOC implies

$$d_S = \frac{C^{'}}{2\mu}(\theta_2 e_1^{'} - e_2^{'})$$

Last, $d_S^{'} d_D = 0$ implies:

$$d_S^{'} d_D = (\theta_2 e_1 - e_2) C C^{'}(e_1^{'} + \theta_1 e_2^{'}) = (\theta_2 \ \ -1)\Sigma \begin{pmatrix} 1 \\ \theta_1 \end{pmatrix} = 0$$

simplifies to:

$$\theta_2 = \frac{\sigma_{12} + \sigma_2^2 \theta_1}{\sigma_1^2 + \sigma_{12}\theta_1}$$

Based on the above equation, the θ_2 is increasing in θ_1 given $|\rho| \neq 1$. Taking (partial) derivative of θ_2 with respect to θ_1:

$$\begin{aligned} \frac{\partial \theta_2}{\partial \theta_1} &= \frac{\sigma_2^2(\sigma_1^2 + \sigma_{12}\theta_1) - \sigma_{12}(\sigma_{12} + \sigma_2^2\theta_1)}{(\sigma_1^2 + \sigma_{12}\theta_1)^2} \\ &= \frac{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}{(\sigma_1^2 + \sigma_{12}\theta_1)^2} \geq 0 \end{aligned}$$

Given $\rho^2 < 1$, and $\sigma_{12}^2 < \sigma_1^2\sigma_2^2$, thus θ_2 is strictly increasing in θ_1. But we need to notice that θ_2 is not continuous in $\theta_1 \in [0, +\infty)$, there exists a discontinuity point. Given the fact:

$$\begin{aligned} \lim_{\theta_1 \to +\infty} \theta_2 &= \lim_{\theta_1 \to +\infty} \frac{\sigma_{12} + \sigma_2^2\theta_1}{\sigma_1^2 + \sigma_{12}\theta_1} = \frac{\sigma_2^2}{\sigma_{12}} \\ \lim_{\theta_1 \to -\frac{\sigma_1^2}{\sigma_{12}}} \theta_2 &= \lim_{\theta_1 \to -\frac{\sigma_1^2}{\sigma_{12}}} \frac{\sigma_{12} + \sigma_2^2\theta_1}{\sigma_1^2 + \sigma_{12}\theta_1} = \infty \end{aligned}$$

then $\left(-\frac{\sigma_1^2}{\sigma_{12}}, \frac{\sigma_2^2}{\sigma_{12}}\right)$ is the discontinuity point and $\theta_1 = -\frac{\sigma_1^2}{\sigma_{12}}$ and $\theta_2 = \frac{\sigma_2^2}{\sigma_{12}}$ are the vertical and horizontal asymptote respectively. Depending on where is the discontinuity point, we have different sets of solutions.

First, we look at the $\sigma_{12} > 0$ case, and discontinuity point is in the second quadrant. Since, we need to constraint θ_1 to be positive. The corner solution $\left(0, \dfrac{\sigma_{12}}{\sigma_1^2}\right)$, on the positive part of the θ_2 axis, is the Cholesky identification. To see why, when $\theta_1 = 0$, $\delta_1 = 1$ given $\theta_1 = \delta_1^{-1} - 1$. From the Problem (1), we are choosing the first shock (demand shock) to maximize the impact on output (but in the Problem (2), we are not setting $\delta_2 = 0$ or $\delta_2 = 1$ since $\theta_2 \neq 0$). That is equivalent to the Cholesky identification where we put output on the top and then inflation. For the another case $\sigma_{12} < 0$, discontinuity point is in the fourth quadrant. Given we need to constraint θ_2 to be positive. The solution $\left(-\dfrac{\sigma_{12}}{\sigma_2^2}, 0\right)$, on the positive part of the θ_1 axis, is the Cholesky identification. To see why, when $\theta_2 = 0$, $\delta_2 = 0$ given $\theta_2 = \delta_2/(\delta_2 - 1)$. From the problem (2), we are choosing the second shock (supply shock) to maximize the impact on inflation (but in the problem (1), we are not setting $\delta_1 = 0$ or $\delta_1 = 1$ since $\theta_1 \neq 0$). That is equivalent to the Cholesky identification where we put inflation on the top and then output.

Second, a set of feasible solutions always exists, given $|\rho| \neq 1$. For the $\sigma_{12} > 0$ case, the θ_2-intercept is $\theta_2 = \sigma_{12}/\sigma_1^2 > 0$. Given $\partial\theta_2/\partial\theta_1 \geq 0$, there will always be a part of the function in the first quadrant. For the $\sigma_{12} < 0$ case, the θ_1-intercept is $\theta_1 = -\sigma_{12}/\sigma_2^2 > 0$. Given $\partial\theta_2/\partial\theta_1 \geq 0$, there will always be a part of the function in the first quadrant.

Third, if $\sigma_{12} = 0$, all the solutions are possible. Given $\sigma_{12} = 0$, then $\theta_2 = \dfrac{\sigma_2^2}{\sigma_1^2}\theta_1$. For any $\theta_1 > 0$, we have $\theta_2 > 0$, and $\theta_2(\theta_1) : [0, \infty) \to [0, +\infty)$. Intuitively, $\sigma_{12} = 0$ implies u_t^Y and u_t^π are uncorrelated. ϵ_t^i are both a linear combination of u_t^Y and u_t^π. When we pin down the first linear combination (that is fixing θ_1), the second linear combination will be uniquely pinned down in order to make ϵ_t^D and ϵ_t^S uncorrelated. Mathematically, when $\sigma_{12} = 0$, D is invertible, and rank(ϵ_t)=rank(u_t) $= 2$.

Last, no sets of feasible solutions exist given $|\rho| = 1$. If $|\rho| = 1$, then $\exists\, a, b$ such that Pr($u_t^Y = a + bu_t^\pi$) $= 1$. It is almost surely that there is a linear relation between u_t^Y and u_t^π, which implies rank(u_t) $= 1$. Given $A = CD$, A must be invertible and $\epsilon_t = A^{-1}u_t$. We know rank(ϵ_t) $\leq \min\{$rank(A^{-1}), rank$(u_t)\} = 1$, i.e., ϵ_t^D and ϵ_t^S are not orthogonal to each other. This is the most extreme value for $|\rho| = 1$ such that we have the smallest set of solutions (i.e. no solution). Graphically, when $|\rho| \to 1$, the θ_2-intercept (θ_1-intercept) will approach to the horizontal (vertical) asymptote, and size of solution will decrease.