Lecture 3: Bayesian Vector Autoregressions: Monetary Policy Evaluation

James M. Nason

Centre for Applied Macroeconomic Analysis
and
Virginia Center for Economic Policy

February 11, 2024

Introduction: Monetary Policy and SVAR Identification

- Researchers have more to identify than just monetary policy shocks.
    1. Is the policy instrument a monetary aggregate, private security, or interest rate?
    2. Is a central bank's policy rule restricted by its monetary operating mechanism?
    3. Do the money/interbank markets react to these choices by the central bank?

- Should a central bank conduct monetary policy using a monetary aggregate, a private security, or an interest rate to alter real allocations?
    1. Given monetary non-neutralities exist, a central bank alters real allocations by changing the relative price of its liability to a (nearly) riskless security traded in the money/interbank markets.
    2. Along which margin does the central bank accomplish this task?

- Are the monetary policy operating mechanism and monetary policy rule linked?

- Does the monetary policy operating mechanism restrict the monetary policy rule?

- Do the money/interbank markets, monetary policy operating system, and monetary policy rule form a simultaneous system?
    1. Are identifying shocks to the supply and demand of inside and outside money necessary to identify monetary policy shocks?
    2. $\Longrightarrow$ The reaction of the money/interbank markets to monetary policy depends on the monetary policy instrument, the monetary policy operating mechanism, and the monetary policy reaction function.

Identifying Monetary Policy from a Monetary Operating System, I

- Monetary policy VARs have to identify forecast innovations
  1. to a short (or policy) rate as the policy shock or to a central bank liability.
  2. Begs the question whether the monetary policy operating mechanism restricts these policy shocks.

- Strongin (1995), Bernanke and Mihov (1998), and Christiano, Eichenbaum, and Evans (1998) argue the Federal Reserve generates monetary policy shocks by altering the level and composition of its liabilities.

- A central bank's balance sheets consist of assets and liabilities.
  1. Assets include financial securities with a potential to range from equity, corporate debt, securitized loans (*i.e.*, mortgage backed securities), to sovereign debt plus miscellaneous items.
  2. Liabilities = cash (*i.e.*, currency) + reserves private banks deposit with the central bank + miscellaneous items.
  3. A reminder: monetary base (MB) = currency + reserves + ...

- The liabilities in play are the level and composition of reserves $\Longrightarrow$ gives the Fed several potential margins on which to manage monetary policy.

Identifying Monetary Policy from a Monetary Operating System, II

- Suppose $\mathcal{Y}_t = [y_t \;\; P_t \;\; CP_t \;\; TR_t \;\; NBR_t \;\; R_{ff,t}]'$ in the **AB**-model

$$\mathbf{A}_0 \mathcal{Y}_t = \mathbf{a} + \sum_{j=1}^{p} \mathbf{A}_j \mathcal{Y}_{t-j} + \mathbf{Q}\eta_t, \;\; \eta_t \sim \mathcal{N}(\mathbf{0}_{n\times 1}, \mathbf{I}_{n\times n}),$$

  where y_t, P_t, CP_t, TR_t, NBR_t, and $R_{ff,t}$ denote output, the aggregate price level, a commodity price index, total reserves, non-borrowed reserves, and the fed funds rate,

$$\mathbf{A}_0 = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0}_{3\times 3} \\ \mathbf{A}_{21,0} & \mathbf{I}_3 \end{bmatrix}, \; \mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{0}_{3\times 3} \\ \mathbf{0}_{3\times 3} & \mathbf{Q}_{22} \end{bmatrix}, \text{ and } \mathbf{E}\left\{\eta_{2,t}\eta_{2,t}'\right\} = \mathbf{Q}_{22}\mathbf{Q}_{22}'.$$

- The first three variables, $\mathcal{Y}_{1,t} = [y_t \;\; P_t \;\; CP_t]'$, are block exogenous w/r/t the monetary policy block of $\mathcal{Y}_{MP,t} = [TR_t \;\; NBR_t \;\; R_{ff,t}]'$ $\Longrightarrow$ can estimate these three regressions using OLS.

- The three OLS regressions of $\mathcal{Y}_{1,t}$ yield three instruments, $\hat{\eta}_{y,t}$, $\hat{\eta}_{P,t}$, and $\hat{\eta}_{CP,t}$ $\Longrightarrow$ could estimate the three regressions of the policy block, $\mathcal{Y}_{MP,t}$, one at a time by IV $\Longrightarrow$ recovers $\mathbf{Q}_{22}\eta_{2,t}$ and $\mathbf{Q}_{22}\mathbf{Q}_{22}'$, which are unrestricted.

Identifying Monetary Policy from a Monetary Operating System, III

- The three regressions $y_{MP,t}$ form a system in which $\mathbf{Q}_{22}$ is not identified $\Longrightarrow$ three restrictions on $\mathbf{Q}_{22}$ satisfy the necessary condition $\Longrightarrow$ as if identifying a $\mathbf{C}$-model.

- The elements of $y_{MP,t}$ are liabilities of the Fed and the overnight rate of a market in which banks trade reserves.
  1. The Fed's liabilities are TR_t and $BR_t \Longrightarrow$ demand by Federal Reserve member banks for total reserves and borrowed reserves,
  2. which is generated by need to meet reserve requirements (trivial) and need to settle financial claims with other Federal Reserve member banks.
  3. $R_{ff,t}$ is the intertemporal price of uncollateralized trades of reserves among these banks that clears the market for these unsecured interbank loans.

- The Fed's reaction function or policy rule sets the "net supply" of reserves.
  1. Net supply of reserves = non-borrowed reserves, NBR_t.
  2. The maintained assumption is the Fed supply of reserves to member banks is perfectly elastic $\Longrightarrow$ the Fed accommodates total, borrowed, or non-borrowed reserve shocks to banks.

Identifying Monetary Policy from a Monetary Operating System, IV

- From Strongin (1995), Christiano, Eichenbaum, and Evans (1998), and Bernanke and Mihov (1998), specify the demand for TR_t and BR_t and the Fed policy rule for NBR_t as

$$\begin{aligned} TR_t &= f(y_{1,t}) - \alpha R_{ff,t} + \sigma_d \eta_{d,t}, \\ BR_t &= f(y_{1,t}) + \beta R_{ff,t} - \gamma NBR_t + \sigma_b \eta_{b,t}, \\ NBR_t &= f(y_{1,t}) + \phi_d \sigma_d \eta_{d,t} + \phi_b \sigma_b \eta_{b,t} + \sigma_s \eta_{s,t}. \end{aligned}$$

  where $\eta_{d,t}$ is the shock to the demand for reserves, $\eta_{b,t}$ is the shock to borrowed reserves, and $\eta_{s,t}$ is the shock to the supply of reserves.

- The system of the demand for reserves restricts
  1. the demand for TR_t to be falling in $R_{ff,t}$,
  2. although a higher $R_{ff,t}$ raises BR_t, it moves opposite to NBR_t,
  3. while NBR_t responds to the shocks, $\eta_{d,t}$, $\eta_{b,t}$, and $\eta_{s,t}$, to reserves.

- Goal is to estimate the eight parameters α, β, γ, ϕ_d, ϕ_b, σ_d, σ_b, and σ_s.

Identifying Monetary Policy from a Monetary Operating System, V

- Since $TR_t = BR_t + NBR_t$, the implications for $\mathbf{A}_{21,0}$ are

$$\mathbf{Q}_{22} = \begin{bmatrix} \frac{\beta + \alpha(1-\gamma)\phi_d}{\alpha+\beta}\sigma_d & \alpha\frac{1-\gamma}{\alpha+\beta}\sigma_s & \alpha\frac{1+(1-\gamma)\phi_b}{\alpha+\beta}\sigma_b \\ \phi_d\sigma_d & \sigma_s & \phi_b\sigma_b \\ \frac{1-(1-\gamma)\phi_d}{\alpha+\beta}\sigma_d & -\frac{1-\gamma}{\alpha+\beta}\sigma_s & -\frac{1+(1-\gamma)\phi_b}{\alpha+\beta}\sigma_b \end{bmatrix}.$$

- Since $\mathbf{E}\left\{\eta_{2,t}\eta'_{2,t}\right\} = \mathbf{Q}_{22}\mathbf{Q}'_{22}$, has $0.5n(n+1) = 6$ independent moments,
  1. at least two identifying assumptions are necessary to just-identify the eight free parameters in $\mathbf{Q}_{22}$.
  2. Three or more identifying assumptions are needed to over-identify the model.

Just-Identifying Monetary Policy from a Monetary Operating System

- At least three sets of restrictions providing necessary conditions to just-identify $\mathbf{Q}_{22}$.
  1. Set $\phi_d = (1-\gamma)^{-1} = -\phi_b \Longrightarrow$ monetary policy shock = forecast innovation of $R_{ff,t} = [(1-\gamma)/(\alpha+\beta)]\sigma_s\eta_{s,t} \Longrightarrow$ the Fed targets $R_{ff,t}$; see Bernanke and Blinder (1992, "The federal funds rate and the channels of monetary transmission," *American Economic Review* 82, 901–921).
  2. $\phi_d = \phi_b = 0 \Longrightarrow$ monetary policy shock = forecast innovation of $NBR_t = \sigma_s\eta_{s,t}$ $\Longrightarrow$ MP targets NBR_t; see Christiano, Eichenbaum, and Evans (1998).
  3. $\alpha = \phi_b = 0 \Longrightarrow$ MP shock = forecast innovation of $TR_t = \sigma_d\eta_{d,t} \Longrightarrow$ Fed has a perfectly elastic TR supply schedule, which accommodates all demand for TR and $R_{ffr,t}$ has no role; see Strongin (1995) and Bernanke and Mihov (1998).
  4. These three sets of restrictions just-identify the monetary policy block of the SVAR of $\mathcal{Y}_t \Longrightarrow$ no over-identifying restrictions to test the competing models of the Fed monetary policy operating system and policy rule.

Over-Identifying Monetary Policy from a Monetary Operating System

- Bernanke and Mihov (1998) set $\gamma = 0$ to over-identify the $R_{ff,t}$, NBR_t, and TR_t targeting models to test against the model that targets BR_t.

- Over-identification is also possible by restricting $\phi_d = 1$, $\phi_b = \alpha/\beta$, and $\gamma = 0$.
  1. $\Longrightarrow$ Monetary policy targets BR_t when

$$\mathbf{Q}_{22} = \begin{bmatrix} \sigma_d & \dfrac{\alpha}{\alpha+\beta}\sigma_s & \dfrac{\alpha}{\beta}\sigma_b \\ \sigma_d & \sigma_s & \dfrac{\alpha}{\beta}\sigma_b \\ 0 & -\dfrac{1}{\alpha+\beta}\sigma_s & -\dfrac{1}{\beta}\sigma_b \end{bmatrix},$$

  2. $\Longrightarrow$ forecast innovation of $BR_t = -[\beta/(\alpha+\beta)]\sigma_s\eta_{s,t} \Longrightarrow$ sterilize the response of BR_t to $\eta_{b,t} \Longrightarrow$ MP offsets BR shocks.
  3. See Cosimano & Sheehan (1994, "The Federal Reserve Operating Procedure," 1984–1990: An Empirical Analysis," *Journal of Macroeconomics* 16, 573–588) and Bernanke & Mihov (1998).

- Identification of the impact matrices $\mathbf{A}_0$ and $\mathbf{Q}$ is needed for Bayesian estimation of the **AB**-model $\Longrightarrow$ the necessary condition is $2n^2 - 0.5n(n+1) = 51$ free parameters to set for just-identified.

Identifying Monetary Policy from a Monetary Aggregate, I

- A different approach to identifying monetary policy shocks is developed by Sims (1992, "Interpreting the macroeconomic time series facts: The effects of monetary policy," *European Economic Review* 36, 975–1000).

- Sims' critique is that tying identification of monetary policy shocks to the innovation in $R_{ff,t}$ rests on specifying supply and demand functions for the Fed's monetary aggregate.

- For example, Gordon and Leeper (1994) tie the monetary policy shock to the innovation in $R_{ff,t}$ and "the financial and goods markets do not respond to current money market disturbances."
  1. Their evaluation of monetary policy compares the monetary transmission when the Fed's monetary aggregate is TR_t compared with $M2_t$.
  2. The question is whether Fed monetary policy works first on the U.S. banking system and then the rest of the economy or on the economy's aggregate money demand function, which include banks, non-financial firms, and households.
  3. Either way, Gordon and Leeper (GL) argue "price and quantity are determined by the simultaneous interaction of supply and demand in money markets."

Identifying Monetary Policy from a Monetary Aggregate, II

- A quick review: identification of monetary policy shocks with forecast innovations in $R_{ff,t}$, TR_t, NBR_t, or BR_t requires
  1. the Fed's supply of TR_t is perfectly elastic w/r/t $R_{ff,t}$ $\Longrightarrow$ an innovation to TR_t is the policy shock,
  2. the Fed's supply of TR_t is perfectly elastic $\Longrightarrow$ an innovation to $R_{ff,t}$ is the policy shock, or
  3. banks' demand for NBR_t is perfectly elastic given TR_t is fixed $\Longrightarrow$ an innovation to NBR_t is the policy shock.

- This approach to identification focuses only on the liability on the Fed's balance sheet while ignoring responses on the balance sheets of participants in the money/interbank markets.
  1. Implicit is the assumption that *shocks to monetary policy are the dominate source of fluctuations in money/interbank markets*.
  2. If inside money/interbank market shocks are (at least) as important, estimates of monetary policy shocks are biased $\Longrightarrow$ omitted variables problem.

Identifying Monetary Policy from a Monetary Aggregate, III

- GL estimate a SBVAR on monthly data identified with a structural ordering, where $\mathcal{Y}_t = [X_t \;\; Z_t \;\; P_t, \;\; IP_t \;\; UR_t \;\; R_{10,t} \;\; CP_t]'$, $X_t = TR_t$ or M2_t and $Z_t = R_{ff,t}$ or $R_{1MTB,t}$.
    1. Starting in the market for reserves, innovations in TR_t and $R_{ff,t}$ are responses to a policy shock $\Longrightarrow$ BR_t and NBR_t are perfect substitutes,
    2. followed by the money/interbank markets (*i.e.*, M2) in which an M2_t supply shock generates innovations in M2_t and 1-month T-bill rate, $R_{1MTB,t}$,
    3. date t monetary and money/interbank market shocks have no impact on financial variables, $R_{10,t}$ (10-year Treasury bond yield), and CP_t, and macro variables, P_t, IP_t (industrial production), and UR_t (unemployment rate), and
    4. these variables are ordered recursively UR_t, IP_t, P_t, $R_{10,t}$, and CP_t.

- Identification of the reserves and money/interbank markets rests on structural shocks that are orthogonal at all leads and lags, money demand reasoning, and information restrictions on when Fed policy makers obtain data.
    1. Shocks to the demand for TR_t or M2_t are derived from opportunity costs and price and income effects.
    2. The opportunity cost is $R_{ff,t}$ when the monetary aggregate = TR_t and for M2_t the intertemporal price is $R_{1MTB,t}$.
    3. Shocks to the supply of TR_t or M2_t are identified by assuming that within a month the Fed observes the monetary aggregate and financial variables only (the FOMC meets every six weeks on average).

IDENTIFYING MONETARY POLICY FROM A MONETARY AGGREGATE, IV

- These identifying assumptions yield the supply-demand system, which integrates the money/interbank markets and the Fed's monetary operating system and policy rule

$$\begin{aligned} \varepsilon_{X,t} &= \alpha_1 \varepsilon_{Z,t} + \alpha_2 \varepsilon_{P,t} + \alpha_3 \varepsilon_{IP,t} + \eta_{d,t}, \\ \varepsilon_{Z,t} &= \alpha_4 \varepsilon_{X,t} + \alpha_5 \varepsilon_{R_{10},t} + \alpha_6 \varepsilon_{CP,t} + \eta_{s,t}. \end{aligned}$$

- The block exogeneity and recursive ordering assumptions and the identified money/interbank markets-Fed monetary operating system-policy rule supply demand system restrict the impact matrix

$$\mathbf{A}_0 = \begin{bmatrix} \mathbf{A}_{11,0} & -\alpha_1 & -\alpha_2 & -\alpha_3 & 0 & 0 & 0 \\ -\alpha_4 & \mathbf{A}_{22,0} & 0 & 0 & 0 & -\alpha_5 & -\alpha_6 \\ 0 & 0 & \mathbf{A}_{33,0} & \mathbf{A}_{34,0} & \mathbf{A}_{35,0} & 0 & 0 \\ 0 & 0 & \mathbf{A}_{43,0} & \mathbf{A}_{44,0} & 0 & 0 & 0 \\ 0 & 0 & \mathbf{A}_{53,0} & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{A}_{63,0} & \mathbf{A}_{64,0} & \mathbf{A}_{65,0} & \mathbf{A}_{66,0} & 0 \\ 0 & 0 & \mathbf{A}_{73,0} & \mathbf{A}_{74,0} & \mathbf{A}_{75,0} & \mathbf{A}_{76,0} & \mathbf{A}_{77,0} \end{bmatrix},$$

where $\mathbf{A}_0 \varepsilon_t = \eta_t = \left[\eta_{X,d,t} \;\; \eta_{Z,s,t} \;\; \eta_{L,t} \;\; \eta_{Y,s} \;\; \eta_{Y,d} \;\; \eta_{FL,t} \;\; \eta_{FS,t}\right]'$.

- There are 28 zeros $\Longrightarrow$ over-identified, non-recursive SBVAR $\Longrightarrow$ the necessary condition for identification is $0.5n(n-1) = 21$.

Gordon and Leeper: Summary

- GL's aim is to identify "an empirical model that is motivated by a traditional view of monetary policy and private sector behavior and accounts for the empirical regularities predicted by theory."
  1. $\Longrightarrow$ Separate the behavior of agents in the money/interbank markets from the monetary operating mechanism and policy rule.
  2. Motivation is SVARs identified only by monetary operating mechanism and policy rule give "odd" results.

- The odd results are the product of excluding broader monetary aggregates and/or giving little weight to the interaction of money/interbank markets from the monetary operating mechanism and policy rule, according to GL.
  1. $\Longrightarrow$ An omitted variables problem in which monetary policy drives inside money (*i.e.*, liabilities of financial firms) movements that are confused with an exogenous monetary policy shock.
  2. $\Longrightarrow$ Liquidity puzzle: a monetary policy shock produces a positive comovement between short nominal rates and monetary aggregates from impact into the short run (*i.e.*, less than two years).
  3. Advocates of the monetary operating mechanism and policy rule approach to monetary policy VARs "solve" the liquidity puzzle by including a narrow monetary aggregate (*i.e.*, TR_t or NBR_t) in $\mathcal{Y}_t$; see Christiano, Eichenbaum, and Evans (1998).
  4. $\Longrightarrow$ Price puzzle: a monetary policy shock that increases short rates also causes higher inflation (or the price level); Sims (1992) is the first to include CP_t in $\mathcal{Y}_t$ to resolve the price puzzle.

Leeper and Roush (JMCB, 2003): How Does Money Matter?

- Can monetary policy be studied without money?
  1. New Keynesian DSGE models equate monetary policy with interest rate rules (*i.e.*, Taylor rules) $\Longrightarrow$ implicit is monetary authority supplies money inelastically $\Longrightarrow$ no money demand.
  2. Central banks often describe their decisions in terms of interest rate rules $\Longrightarrow$ monetary aggregates are rarely mentioned.
  3. But, LR report responses of output and prices to an identified monetary policy shock that are sensitive to the role of money in monetary policy SVARs.

- LR study whether the
  1. monetary transmission mechanism is misspecified without money or
  2. is there information in money useful to identify a monetary policy separate from endogenous interest rate responses to changes in inflation.

- The research question is "Do inferences about monetary policy impacts depend on assumptions about how money enters the empirical model?"

Money, Monetary Policy, and Recursive and Non-Recursive Identifications, I

- LR estimate **K**-model monetary policy SVARs identified on recursive and non-recursive orderings with and without a monetary aggregate.

- Given $\mathcal{Y}_t = [Y_t \;\; C_t \;\; UR_t \;\; P_t \;\; CP_t \;\; R_{ff,t}]'$ lacks a monetary aggregate, the non-recursive identification is

$$\mathbf{A}_0 = \begin{bmatrix} \text{PM} & \text{PM} & \text{PM} & \text{PM} & \text{IN} & \text{MP} \\ \mathbf{A}_{11,0} & \mathbf{A}_{12,0} & \mathbf{A}_{13,0} & \mathbf{A}_{14,0} & \mathbf{A}_{15,0} & \mathbf{A}_{16,0} \\ 0 & \mathbf{A}_{22,0} & \mathbf{A}_{23,0} & \mathbf{A}_{24,0} & \mathbf{A}_{25,0} & 0 \\ 0 & 0 & \mathbf{A}_{33,0} & \mathbf{A}_{34,0} & \mathbf{A}_{35,0} & 0 \\ 0 & 0 & 0 & \mathbf{A}_{44,0} & \mathbf{A}_{45,0} & \mathbf{A}_{46,0} \\ 0 & 0 & 0 & 0 & \mathbf{A}_{55,0} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{A}_{65,0} & \mathbf{A}_{66,0} \end{bmatrix},$$

  where $\mathbf{A}_0\varepsilon_t = \eta_t$, PM, IN, and MP denote the product market, information, and monetary policy sectors, and the SVAR is over-identified $\Longrightarrow$ 17 zeros and $0.5n(n-1) = 15$.

- The non-recursive SVAR is built on three impact identification assumptions.
  1. Shocks to the information variable, CP_t, affect the rest of the variables in $\mathcal{Y}_t$.
  2. PM shocks only effect the variables of the PM.
  3. Y_t, P_t, and $R_{ff,t}$ are the only elements of $\mathcal{Y}_t$ responding to MP shocks at impact.

Money, Monetary Policy, and Recursive and Non-Recursive Identifications, II

- LR add money to their SVARs in several ways.
- The idea is to include in η_t demand shocks to $\mathrm{M2}_t$ that move
  1. it, C_t, P_t, $R_{ff,t}$, and the opportunity cost of holding $\mathrm{M2}_t$, $R_{ff,t} - R_{\mathrm{M2},t}$, at impact, where $R_{\mathrm{M2},t}$ is the return on $\mathrm{M2}_t$.
  2. $\Longrightarrow$ A linear restriction equates coefficients on $R_{ff,t}$ and $R_{\mathrm{M2},t}$ in the MD block.
  3. $\mathrm{M2}_t$ and $R_{ff,t}$ respond to the MP shock at impact.
- The non-recursive over-identified (32 zeros and $0.5n(n-1) = 28$) SVAR has

$$\mathbf{A}_0 = \begin{bmatrix} \mathrm{PM} & \mathrm{PM} & \mathrm{PM} & \mathrm{PM} & \mathrm{IN} & \mathrm{IN} & \mathrm{MD} & \mathrm{MP} \\ \mathbf{A}_{11,0} & \mathbf{A}_{12,0} & \mathbf{A}_{13,0} & \mathbf{A}_{14,0} & \mathbf{A}_{15,0} & \mathbf{A}_{16,0} & 0 & 0 \\ 0 & \mathbf{A}_{22,0} & \mathbf{A}_{23,0} & \mathbf{A}_{24,0} & \mathbf{A}_{25,0} & \mathbf{A}_{26,0} & \mathbf{A}_{27,0} & 0 \\ 0 & 0 & \mathbf{A}_{33,0} & \mathbf{A}_{34,0} & \mathbf{A}_{35,0} & \mathbf{A}_{36,0} & 0 & 0 \\ 0 & 0 & 0 & \mathbf{A}_{44,0} & \mathbf{A}_{45,0} & \mathbf{A}_{46,0} & \mathbf{A}_{47,0} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{A}_{55,0} & \mathbf{A}_{56,0} & \mathbf{A}_{57,0} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{A}_{66,0} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{A}_{75,0} & \mathbf{A}_{76,0} & \mathbf{A}_{77,0} & \mathbf{A}_{78,0} \\ 0 & 0 & 0 & 0 & \mathbf{A}_{85,0} & \mathbf{A}_{86,0} & \mathbf{A}_{87,0} & \mathbf{A}_{88,0} \end{bmatrix},$$

where $\mathcal{Y}_t = [Y_t \;\; C_t \;\; UR_t \;\; P_t \;\; R_{\mathrm{M2},t} \;\; CP_t \;\; \mathrm{M}_{2,t} \;\; R_{ff,t}]'$ and $\mathbf{A}_{57,0} = \mathbf{A}_{87,0}$.

-

Money, Monetary Policy, and Recursive and Non-Recursive Identifications, III

- LR report recursive identifications produce the liquidity puzzle.
    1. The magnitude of the liquidity effect predicts whether monetary policy shocks drive movements in real activity.
    2. Estimated liquidity effect's size depends on ordering of $M2_t$ and $R_{ff,t}$.
    3. Given $M2_t$ is ordered before $R_{ff,t}$, monetary policy shocks generate larger non-neutralities than when this order is reversed.
    4. These orderings map the reduced form correlation of $M2_t$ and $R_{ff,t}$ into a different monetary policy shocks.
    5. A monetary policy shock identified by the innovation to $M2_t$ has large real effects because it Granger causes output and inflation.

- Non-recursive identification schemes generate larger real responses to monetary policy shocks $\Longrightarrow$ identifying money supply and demand shocks place less restrictions on the map from reduced form dynamics to structural responses.

Introduction: The Fed Funds Futures Market

- The Chicago Merchantile Exchange (CME) offers futures contracts on $R_{ff,t}$, which are traded on the Chicago Board of Trade (CBOT).

- These are American style options $\Longrightarrow$ the buyer can exercise the option at any moment before the expiration date of the contract.

- Fed funds futures contracts are available 30-days, 6-months, and 12-months ahead.

- 30-day contracts can be "rolled over" to price $R_{ff,t+j}$ at 30 day intervals, $j = 30, 60, 90, \ldots \Longrightarrow$ price fed funds future forward contracts curve.

- The fed funds future contract is a "bet" on $R_{ff,t+h}$, date $t+h$ effective fed funds rate.
  1. Let $f_{d,h}$ = the fed funds future contract set at day d, which pays off h months ahead
  2. Net payoff on fed funds future contract = $f_{d,h} - R_{ff,t+h}$.

- The prices of these contracts are also used to construct the probability of a future interest rate action by the FOMC on the dates it will meet.

Faust, Swanson, and Wright (JME, 2004)

- Disagreement over identification of monetary policy SVARs.
- Identifications seem a priori reasonable, but
  1. incomplete descriptions of DSGE model predictions and
  2. create empirical puzzles $\Longrightarrow$ monetary policy SVARs yield different estimates by tweaking the identification.
- Faust, Swanson, and Wright (FWR) impose the path
  1. of fed fund futures on the response of $R_{ff,t}$ to an own shock.
  2. $\Longrightarrow$ The monetary policy shock is conditioned on data sampled at a higher frequency than is $\mathcal{Y}_t$.

Identifying Monetary Policy SVARs with Fed Funds Futures

- Start from the unrestricted VAR, $\mathcal{Y}_t = \mathbf{B}(\mathbf{L})\mathcal{Y}_{t-1} + \varepsilon_t$, $\varepsilon_t \sim \mathcal{N}(\mathbf{0}_{n\times 1}, \mathbf{\Omega})$.
  1. Assume $\mathbf{C}(\mathbf{L}) = [\mathbf{I}_n - \mathbf{B}(\mathbf{L})]^{-1}$ and $\varepsilon_t = \mathbf{D}\eta_t \Longrightarrow \mathcal{Y}_t = \mathbf{C}(\mathbf{L})\mathbf{D}\eta_t$.
  2. Let $R_{ff,t}$ be the first element of $\mathcal{Y}_t \Longrightarrow \mathbf{D}_{R_{ff}}$ is the first column of $\mathbf{D}$.
  3. $\mathbf{C}(\mathbf{L})\mathbf{D}_{R_{ff}} = \sum_{\ell=0}^{\infty} \mathbf{C}_\ell \mathbf{D}_{R_{ff}} \mathbf{L}^\ell \Longrightarrow$ these MA(∞) are the IRFs w/r/t the monetary policy shock, which at $\ell = 0$ is a 25 basis point increase in $R_{ff,t}$.
  4. Often the FOMC moves the target fed funds rate by this amount $\Longrightarrow$ IRFs equivalent to a one standard deviation shock, but confidence bands will differ.
  5. FSW identify “the impulse response of the funds rate to the policy shock” with “the response measured in the futures market data.”
  6. Label this IRF $\mathbf{G}_{R_{ff},h}$ at horizon $h = 0, 1, 2, \ldots, H \Longrightarrow \mathbf{G}_{R_{ff},h} = \mathbf{C}_h\mathbf{D}_{R_{ff}}$.
  7. Stack the $H+1$ equations, $h = 0, 1, 2, \ldots, H$ to obtain the system $\mathcal{R}\mathbf{D}_{R_{ff}} = \mathbf{G}_{R_{ff}}$ $\Longrightarrow$ the unique solution is $\mathbf{D}_{R_{ff}} = \mathcal{R}^{-1}\mathbf{G}_{R_{ff}}$, given $\mathcal{R}$ is of rank $H+1$.
  8. Otherwise, $\mathbf{D}_{R_{ff}} = \mathcal{R}^{-1}\mathbf{G}_{R_{ff}}$ is only partially identified $\Longrightarrow$ violate the necessary condition to identify SVARs because the lack of full rank in $\mathcal{R}$ is equivalent to near linear dependence of monetary policy IRFs $\Longrightarrow \mathbf{G}_{R_{ff},h} \approx \mathbf{G}_{R_{ff},h+1}$.

- Constructing confidence bands of $\mathbf{G}_{R_{ff}}$ have to account for
  1. its sampling uncertainty because its elements are estimated and
  2. its partial identification when $\mathcal{R}$ is rank deficient $\Longrightarrow$ need to employ robust methods to compute confidence bands.

Estimating the response of $R_{ff,t}$ to $f_{d,h}$ Shocks, I

- No cash is required upfront to buy $f_{d,h}$.
  1. Along with no arbitrage $\Longrightarrow \mathbf{E}_t\left\{m_{t+h}(f_{d,h} - R_{ff,t+h})\right\} = 0$, where m_{t+h} is the stochastic discount factor (SDF).
  2. Since $Cov(x_1, x_2) = \mathbf{E}\{x_1 x_2\} - \mathbf{E}\{x_1\}\mathbf{E}\{x_2\}$

$$\Longrightarrow \qquad f_{d,h} = \mathbf{E}_t R_{ff,t+h} + \frac{Cov_d(m_{t+h}, R_{ff,t+h})}{\mathbf{E}_t m_{t+h}},$$

     where $Cov_d(\cdot, \cdot)$ is the day d conditional covariance of the SDF and fed funds rate h-months ahead.
  3. The payoff on a day d, h-month ahead fed funds future contract = expected effective fed funds rate at date $t+h$ plus a risk premium.

- FSW invoke three assumptions to move from the pricing function of $f_{d,h}$ to the response of $R_{ff,t+h}$ w/r/t a shock to $f_{d,h}$.
  1. Day to day changes in the fed funds future risk premium are tiny $\Longrightarrow f_{d,h} - f_{d-1,h} = (\mathbf{E}_d - \mathbf{E}_{d-1})R_{ff,t+h} \Longrightarrow \Delta_d f_{d,h} = \Delta_d \mathbf{E}_t R_{ff,t+h}$.
  2. Changes in $f_{d,h}$ are exogenous monetary policy shocks only on days when the FOMC releases a policy statement.
  3. Assumes data updates on these days have no effect on the FOMC and the Fed has no private information revealed by these statements $\Longrightarrow$ Identifying assumption is changes in $f_{d,h}$ are unanticipated.

Estimating the response of $R_{ff,t}$ to $f_{d,h}$ Shocks, II

- If the identified SVAR recovers the "true" monetary policy shock, $\mathbf{E}_t R_{ff,t+h} = \sum_{\ell=0}^{\infty} \mathbf{C}_{\ell+h}(1,\cdot)\mathbf{D}\eta_{t-\ell}$.

- Change $\Delta_d \mathbf{E}_t R_{ff,t+h}$ is driven only by changes in the day d expectations of the structural shocks, $\eta_t \Longrightarrow \eta_{t-\ell}$ is known on day d during month t.
  1. $\Delta_d \mathbf{E}_t R_{ff,t+h} = \mathbf{C}_h(1,\cdot)\mathbf{D}\Delta_d \mathbf{E}_t \eta_t \Longrightarrow \Delta_d \mathbf{E}_t R_{ff,t+h} = \mathbf{C}_h(1,\cdot)\mathbf{D}_{R_{ff}} \Delta_d \mathbf{E}_t \eta_{R,ff,t} + \nu_{h,t}$, where $\nu_{h,t}$ is non-FOMC news about $R_{ff,t+h} \Longrightarrow$ FSW assume $\nu_{h,t} = 0$.
  2. $\Longrightarrow \Delta_d f_{d,h} = \mathbf{G}_{R_{ff},h} \Delta_d \mathbf{E}_t \eta_{R,ff,t}$, where $\mathbf{G}_{R_{ff},h} = \mathbf{C}_h \mathbf{D}_{R_{ff}}$, $h = 0, 1, 2, \ldots, H$.
  3. Since $\Delta_d \mathbf{E}_t \eta_{R,ff,t}$ is independent of h, $\mathbf{G}_{R_{ff},0} \Delta_d \mathbf{E}_t \eta_{R,ff,t} = 0.25 \times$ the scale of the shock change $= \Delta_d \mathbf{E}_t \eta_{R,ff,t}$.
  4. $\Longrightarrow \Delta_d f_{d,h} = \left(\mathbf{G}_{R_{ff},h} / \mathbf{G}_{R_{ff},0}\right) \Delta_d f_{d,0} \Longrightarrow$ the change in the fed funds futures payoff is relative to the 25bps FOMC policy change on day d.
  5. The unobserved fed funds rate shock $\eta_{R,ff,t}$ is identified with the known change in the fed funds futures payoff, which is proportional to the hypothesized policy experiment of a 25bps change in $R_{ff,t}$ for $h = 1, 2, 3, 4, 5$.

- FSW estimate $\mathbf{G}_{R_{ff},h}$, $h \geq 1$, by regressing FOMC-statement day $\Delta_d f_{d,h}$ on change in the target fed funds rate.

Estimating the response of $R_{ff,t}$ to $f_{d,h}$ Shocks, III

- The VAR information set is $\mathcal{Y}_t = [IP_t \;\; P_t \;\; CP_t \;\; R_{ff,t} \;\; NBR_t \;\; TR_t]'$.
- FSW estimate an unrestricted VAR and recover $\widehat{\mathbf{D}} = \widehat{\boldsymbol{\Omega}}^{0.5} \Longrightarrow$ the structural innovation to $R_{ff,t}$, $\eta_{R_{ff},t}$, is identified as the monetary policy shock.
- The estimated structural IRFs replicate the extant literature.
  1. A 25bps increase in $R_{ff,t} \Longrightarrow IP_t$ falls with a trough at a year and a half and
  2. there is a price puzzle in the short run, which is made worse by dropping CP_t.
- When $f_{d,h}$ identifies monetary policy SVARs, reject recursive identifications
  1. in which $\eta_{R_{ff},t}$ is the monetary policy shock,
  2. the IRF of IP_t is similar to that identified by $\eta_{R_{ff},t}$,
  3. contributes only a small fraction to the FEVD of IP_t,
  4. and P_t responds to this monetary policy shock at impact.
  5. $\Longrightarrow f_{d,h}$ identifies a monetary policy shock that is causally structural prior to P_t.

THE LUCAS CRITIQUE

- Analyzing changes in monetary policy as once and for all is a long standing tradition among macroeconomists and central bankers.
  1. This is the approach of taken by Lucas (1976, "Econometric policy evaluation: a critique," in Brunner, K. and A.H. Meltzer (eds.), *Carnegie-Rochester Conference Series on Public Policy* 1, 104–130).
  2. However, monetary policy is a repeated process, which involves switching or moving between alternative policy regimes $\Longrightarrow$ the process of changing policy regimes is stochastic and not known with certainty.
  3. $\Longrightarrow$ Monetary policy afflicted by regime switching introduces a source of nonlinear dynamics into the economy.
  4. See Sims (1987, "A rational expectations framework for short-run policy analysis," in Barnett, W.A., and K.J. Singleton (eds.), NEW APPROACHES TO MONETARY ECONOMICS, Cambridge University Press, Cambridge, UK, pp. 293–308) and Sargent (1999, THE CONQUEST OF AMERICAN INFLATION, Princeton University Press, Princeton, NJ).

- Leeper and Zha (JME, 2003) build on this insight.
  1. Monetary policy regimes are governed by a Markov chain process $\Longrightarrow$ the true DGP of the economy is nonlinear.
  2. Within a policy regime the dynamics of the economy are approximately linear.
  3. A monetary policy regime is characterized by the central bank's linear policy reaction function or rule.

Leeper and Zha (JME, 2003)

- Suppose a central bank asks an economist for advice about monetary policy.
  1. Only linear econometric tools are available to the economist.
  2. The economist only studies exogenous monetary policy shocks within a regime.
  3. Conditioning on a specific monetary policy regime, the economist evaluates potential policy experiments/intervention $\Longrightarrow$ changes in policy are linear.

- LZ decompose effects of a policy experiment under the current regime into
  1. *direct effects* $\Longrightarrow$ measured by tools that are standard for SVAR policy analysis.
  2. However, private agents update their expectations in response to monetary policy interventions $\Longrightarrow$ *expectations formation effect*.
  3. The *expectations formation effect* is a probability statement about the extent to which private agents revise their beliefs about the stability of the current regime in response to a monetary policy intervention.

- Private agents do not engage in substantial updating of their expectations about a regime switch
  1. when a monetary policy experiment is consistent with the current regime $\Longrightarrow$ a modest policy change.
  2. Immodest policy occurs if private agents revise their beliefs a central bank has moved to a different policy regime defined by an alternative policy rule.

Modest and Immodest Monetary Policy Experiments

- The LZ notions of modest and immodest monetary policy experiments are built on the IRFs w/r/t monetary policy shocks.

- The policy intervention measure = $\sum_{\ell=0}^{H} \mathbf{C}_\ell(MP, \cdot)\mathbf{D}_{MP}\eta_{MP,T+1+H-\ell}$, where
  1. $\mathbf{C}_\ell(MP, \cdot)$ and $\mathbf{D}_{MP}$ denote MA(∞) of the monetary policy variable (*i.e.*, $R_{ff,t}$)
  2. and the column of responses to the identified monetary policy shock, $\eta_{MP,t}$.

- Modest policy interventions are not "too large" $\Longrightarrow$ a sequence of $\{\eta_{MP,T+1+H-\ell}\}_{\ell=0}^{H}$
  1. could be large as long as M does not signal a persistent policy intervention.
  2. Monetary policy shocks can be volatile but not persistent.

- Small $\eta_{MP,t}$s and persistent policy interventions generate a large $\sum_{\ell=0}^{H} \mathbf{C}_\ell(MP, \cdot)\mathbf{D}_{MP}\eta_{MP,T+1+H-\ell} \Longrightarrow$ immodest policy experiments.
  1. Ex: A central bank forced to generate a persistent sequence of monetary policy surprises to achieve its goals
  2. causes private agents to revise their expectations about the stability of the current monetary policy regime.

- Private agents revise their expectations and alter their decision rules when they believe the central bank will change its policy regime $\Longrightarrow$ nonlinear dynamics.

Introduction to Nonlinear VARs, I

- Immodest policy experiments suggest policy analysis grounded in linear econometric models is limited $\Longrightarrow$ this idea predates Leeper and Zha (JME, 2003).
  1. See Cogley and Sargent (2002, "Evolving Post-World War II U.S. inflation dynamics," in Bernanke, B.S., and K. Rogoff NBER Macroeconomics Annual 2001, Volume 16, MIT Press, Cambridge, MA, pp. 331–388).
  2. Cogley and Sargent study the joint dynamics of $\mathcal{Y}_t = [\pi_t \;\; UR_t \;\; R_{3MTB,t}]'$ with a recursively identified SVAR $\Longrightarrow$ a **C**-model.

- Their time-varying parameter (TVP-)SVAR is

$$\mathcal{Y}_t = \sum_{\ell=1}^{p} \mathbf{B}_{\ell,t}\mathcal{Y}_{t-\ell} + \mathbf{D}\eta_t, \;\; \mathbf{D}\eta_t = \varepsilon_t, \;\; \varepsilon_t \sim \mathcal{N}\left(\mathbf{0}_{3\times 1}, \mathbf{\Omega}_\varepsilon\right),$$

  where the TVPs evolve as random walks, $\mathbf{b}_t = \mathbf{b}_{t-1} + \vartheta_t$, $\vartheta_t \sim \mathcal{N}\left(\mathbf{0}_{9p\times 1}, \mathbf{\Omega}_\vartheta\right)$, and $\mathbf{E}\{\varepsilon_{t+j}\vartheta_{t+s}\} = \mathbf{V}$ for $j = s = 0$, and zero otherwise.

- The goal is to test natural rate theories and the accelerationist Phillips curve; see King and Watson (1997, "Testing long-run neutrality," Federal Reserve Bank of Richmond *Economic Quarterly* 83/3(Summer), 69–101).

Introduction to Nonlinear VARs, II

- Cogley and Sargent (2002) assert their estimates show
  1. the $\mathbf{E}_t \pi_{t+j}$ and its time-varying persistence have positive comovement
  2. and the persistence of π_t has fallen since 1990 $\Longrightarrow$ natural rate theories are consistent with the estimated time-varying persistence in π_t.
  3. The Phillips curve is vertical in the LR, which signals there is no LR π_t-UR_t trade off.

- Econometricians and policy makers face the problem, according to Cogley and Sargent, that fixed coefficient models will reject natural rate theory as π_t persistence falls.

- $\Longrightarrow$ A reduced form π_t-UR_t trade off reappears, but attempts to exploit the trade off only raises π_t and restarts its persistence $\Longrightarrow$ immodest policy.

Introduction to Nonlinear VARs, III

- The Cogley and Sargent (2002) TVP-SVAR is criticized by
  1. Sims (2002, "Comment on 'Evolving Post-World War II U.S. inflation dynamics'," in Bernanke, B.S., and K. Rogoff NBER Macroeconomics Annual 2001, Volume 16, MIT Press, Cambridge, MA, pp. 373–379) and
  2. Stock (2002, "Comment on 'Evolving Post-World War II U.S. inflation dynamics'," in Bernanke, B.S., and K. Rogoff NBER Macroeconomics Annual 2001, Volume 16, MIT Press, Cambridge, MA, pp. 379–387).

- Sims and Stock argue the Cogley and Sargent TVP-SVAR specification lacks time-varying or stochastic volatility (SV) in η_t $\Longrightarrow$ the TVP-SVAR estimates conflate SV with $\mathbf{b}_t$, which is an omitted variable problem fixed by including $\mathbf{\Omega}_{\varepsilon,t}$.
  1. For example, time-varying mean reversion can appear observationally equivalent to stochastic volatility in forecast innovations, but these have different economic, econometric, and policy implications.
  2. $\Longrightarrow$ If π_t and UR_t suffer from SV, but a decline in SV is misconstrued as a drop in persistence, monetary policy experiments that aim to trade lower UR_t for higher π_t will be viewed as immodest by private agents
  3. $\Longrightarrow$ revive a dormant π_t process and raise the costs of lowering UR_t.

Example: AR(1) with TVP and Stochastic Volatility

- Consider the TVP-AR(1) with stochastic volatility (SV)

$$y_t \; = \; (1 - \rho_t)\overline{y}_t \; + \; \rho_t y_{t-1} \; + \; \sigma_{e,t} e_t,$$

  where $\overline{y}_t$ is the time-varying population average of y_t, the time-varying AR1 coefficient $\rho_t \in (-1, 1)$, and $\sigma_{e,t}$ is time-varying heteroskedasticity of the Gaussian forecast innovation $e_t \sim \mathcal{N}(0, 1)$.

- Persistence and volatility of y_t has three sources: $\overline{y}_t$, ρ_t, and $\sigma_{e,t}$.

- Identification of $\overline{y}_t$, ρ_t, and $\sigma_{e,t}$ rely on assumptions about the stochastic processes driving these latent factors and the correlation of the innovations.

Recovering Hidden Economic Structure: TVP or Markov Switching

- Cogley and Sargent (2005) estimate a TVP-SVAR that includes SV in $\varepsilon_t \Longrightarrow \varepsilon_t = \mathbf{\Omega}_{\varepsilon,t}^{0.5}\eta_t$,
  1. where $\mathbf{\Omega}_{\varepsilon,t} = \mathbf{D}_t\mathbf{\Gamma}_t^{-1}\mathbf{D}_t'$, $\mathbf{D}_t$ is a lower triangular matrix with ones on its diagonal,
  2. and $\mathbf{\Gamma}_t$ = diagonal matrix of precision (inverse of volatility) scaling parameters.
  3. Estimates show SV dominates TVP slope coefficients because of difficulties in "detecting" the latter $\Longrightarrow$ tests for TVP slope coefficients have low power.
  4. $\Longrightarrow$ "Agnosticism about" TVP slope coefficients is a safer position according to Cogley and Sargent.
  5. Cogley and Sargent (2002, 2005) view $\mathbf{b}_t$ (and the decomposition of $\mathbf{\Omega}_{\varepsilon,t}$) as state variables hidden from the econometrician.

Recovering Hidden Economic Structure: TVP or Markov Switching, II

- Another tradition in econometrics maintains that the underlying state or regime of the economy is latent and to be estimated.
    1. See Hamilton (1989, "A new approach to the economic analysis of nonstationary time series and the business cycle," *Econometrica* 57, 357–384) and Kim and Nelson (1999, State-space models with regime switching: Classical and Gibbs-Sampling approaches with applications, Cambridge, MA: MIT Press).
    2. The economy moves or switches between states or regimes differing by growth rates (*i.e.*, state dependent conditional means), lagged dynamic responses, volatilities, and/or the specification of policy rules $\Longrightarrow$ Sims and Zha (2006).

- Sims, Waggoner, and Zha (2008, SWZ) argue regime switching identifies continuous changes or drift and abrupt or discontinuous shifts in BVAR parameters while a TVP-SV-VAR only models continuous parameter drift.

Sims and Zha (2006): Monetary Policy Evaluation with MS-BVARs, I

- Sims and Zha (2006) identify regime switching with a Markov switching (MS) process.
  1. MS places testable restrictions on the BVAR of Sims and Zha (1998).
  2. $\Longrightarrow$ Identify a SVAR in which the latent state vector is driven by MS.

- MS is imposed on the impact coefficients, $\boldsymbol{a}_0(\mathcal{S}_t)$, slope coefficients, $\boldsymbol{a}_{\mathbb{A}}(\mathcal{S}_t)$, and the diagonal elements of the vector $\boldsymbol{\Gamma}(\mathcal{S}_t)$, which scales the precision (inverse volatility) of η_t, where $\mathcal{S}_t$ is the vector of states or regimes.
  1. $\mathcal{S}_t = [\mathcal{S}_{1,t}\ \ \mathcal{S}_{2,t}\ \ldots\ \mathcal{S}_{h,t}]' \Longrightarrow h$ regimes, where
  2. $\Pr(\mathcal{S}_t = i \,|\, \mathcal{S}_{t-1} = k) = \omega_{i,k},\ i, k = 1, \ldots, h.$

- The MS-BVAR is a $\mathbf{K}$-model, $\mathbf{A}_0(\mathcal{S}_t)\mathcal{Y}_t = \mathbf{a}(\mathcal{S}_t) + \sum_{j=1}^{p} \mathbf{A}_j(\mathcal{S}_t)\mathcal{Y}_{t-j} + \boldsymbol{\Gamma}^{-1}(\mathcal{S}_t)\eta_t$.

- The SZ priors, Λ, are applied to the MS-BVAR plus additional priors on the probabilities of the latent regimes.

- The vector of probabilities of the latent regimes, ω, $\boldsymbol{a}_0(\mathcal{S}_t)$, $\boldsymbol{a}_{\mathbb{A}}(\mathcal{S}_t)$, and $\boldsymbol{\Gamma}(\mathcal{S}_t)$, are estimated using a Metropolis within Gibbs simulator.

- MCMC adapts Sims and Zha (1998) by scaling the prior variances of $\boldsymbol{a}_0(\mathcal{S}_t)$ and $\boldsymbol{a}_{\mathbb{A}}(\mathcal{S}_t)$ by h.

Sims and Zha (2006): Monetary Policy Evaluation with MS-BVARs, II

- Sims and Zha estimate MS-BVARs where $\mathcal{Y}_t = [PC_t \ \ M2_t \ \ R_{ff,t} \ \ y_t \ \ P_t \ \ UR_t]'$
  1. with fixed slope coefficients and intercepts and only MS in the SV,
  2. only MS in slope coefficients of monetary policy variables change and SV,
  3. only MS in slope coefficients of non-monetary policy variables change and SV,
  4. and MS in slope coefficients, intercepts, and SV.

- The (fixed coefficient) impact matrix displays a non-recursive identification

$$\mathbf{A}_0 = \begin{bmatrix} \text{IN} & \text{MP} & \text{MD} & \text{PM} & \text{PM} & \text{PM} \\ \mathbf{A}_{11,0} & 0 & 0 & 0 & 0 & 0 \\ \mathbf{A}_{21,0} & \mathbf{A}_{22,0} & \mathbf{A}_{23,0} & 0 & 0 & 0 \\ \mathbf{A}_{31,0} & \mathbf{A}_{32,0} & \mathbf{A}_{33,0} & 0 & 0 & 0 \\ \mathbf{A}_{41,0} & 0 & \mathbf{A}_{43,0} & \mathbf{A}_{44,0} & \mathbf{A}_{45,0} & \mathbf{A}_{46,0} \\ \mathbf{A}_{51,0} & 0 & \mathbf{A}_{53,0} & 0 & \mathbf{A}_{55,0} & \mathbf{A}_{56,0} \\ \mathbf{A}_{61,0} & 0 & 0 & 0 & 0 & \mathbf{A}_{66,0} \end{bmatrix}.$$

- The identification consists
  1. of a simultaneous money supply-money demand system, which generates $M2_t$ and $R_{ff,t}$,
  2. the product market sector is block exogenous w/r/t shocks to the money supply-money demand system and CP_t,
  3. and shocks to CP_t drive the other five variables in $\mathcal{Y}_t$.

Sims and Zha (2006): Monetary Policy Evaluation with MS-BVARs, III

- The data rank the BVAR with MS in the volatility of η_t and the slope coefficients of the monetary policy variable regressions are driven by a four regime MS process.
  1. A regime running from about 1980 to 1984 $\Longrightarrow$ Volcker disinflation.
  2. A Greenspan regime starts in the late 1980s and runs to the end of the sample in 2003 and for much of the 1960s.
  3. The Burns regime begin in the late 1960s and last through the 1970s $\Longrightarrow$ rising inflation to the great inflation.
  4. Counterfactual $\Longrightarrow$ feed the 1970s Burns regime shocks into the MS-BVAR with the Volcker or Greenspan policy rule (fixed coefficients $\Longrightarrow$ posterior means) $\Longrightarrow$ produce lower inflation and less output growth than observed in-sample.

- Repeat the counterfactual experiments, but draw from the posterior distributions.
  1. $\Longrightarrow$ Explore the risk of the Burns, Volcker, and Greenspan regimes.
  2. These counterfactual show (median) inflation in the Volcker and Greenspan regimes compared with the Burns regime.
  3. Volcker or Greenspan regime have greater risk of deflation and negative output growth while the Burns regime does not in the 1970s.
  4. $\Longrightarrow$ Inflation does not fall in the 1980s under the Burns regime, but its median output growth is higher than in-sample (not by much though and is less than that produced by the Volcker and Greenspan regimes).
  5. $\Longrightarrow$ Burns regime would not have ended the inflation of the 1970s.

Sims and Zha (2006): Monetary Policy Evaluation with MS-BVARs, IV

- Data prefer the BVAR with only MS in the volatility of $\eta_t \Longrightarrow$ there are nine MS volatility regimes.
- This is evidence, according to SZ, of a "stable monetary policy reactions to a changing array of major disturbances generated the historical pattern."
  1. Monetary policy reacts to fiscal (Vietnam War), oil price, and labor market shocks, which dominate aggregate fluctuations.
  2. $\Longrightarrow$ SZ narrative combines good luck-bad luck hypothesis with an unchanging Fed policy rule that can produce "sub-optimal" responses to bad shocks.
  3. $\Longrightarrow$ Burns regime is not a sunspot equilibrium $\Longrightarrow$ the Fed reacted strongly to inflation shocks or the Taylor principle is not violated.
  4. The Fed's policy rule valued real activity $\Longrightarrow$ the reaction to bad shocks implies the Fed was unwilling to trade or sacrifice less real activity for lower inflation.

Sims and Zha (2006): Monetary Policy Evaluation with MS-BVARs, V

- If little prior weight is placed on the Fed never altering its policy rule and operating mechanism, the evidence is
    1. monetary aggregates were important during the Burns regime of the 1970s $\Longrightarrow$ monetarism,
    2. the Volcker regime targeted reserves, while the Greenspan regime returned to the interest rate target of the Martin era (pre-1970).
    3. $\Longrightarrow$ Changes in Fed policy rules and operating mechanisms "were of uncertain timing, not permanent, and not easily understood."
    4. $\Longrightarrow$ Monetary policy evaluation needs to confront evidence the Fed changed the variable(s) on which its policy operated, which requires econometric tools that treat these regime shifts as transitory, stochastic, and opaque.

A Structural MS-BVAR: Specification

- Sims and Zha (2006) and Sims, Waggoner, and Zha (2008) write the MS-BVAR as

$$\mathcal{Y}_t'\mathbf{A}_0(\mathcal{S}_t) = \mathbf{a}(\mathcal{S}_t) + \sum_{j=1}^{p}\mathcal{Y}_{t-j}'\mathbf{A}_j(\mathcal{S}_t) + \eta_t'\mathbf{\Gamma}^{-1}(\mathcal{S}_t).$$

1. The density of η_t is $\mathcal{P}\left(\eta_t \,\middle|\, \mathcal{Y}_{t-1}, \boldsymbol{\mathcal{S}}_t, \omega, \mathbf{\Theta}\right) = \mathcal{N}\left(\eta_t \,\middle|\, \mathbf{0}_{n\times 1}, \mathbf{I}_n\right)$ and
2. the density of $\mathcal{Y}_t$ is $\mathcal{P}\left(\mathcal{Y}_t \,\middle|\, \boldsymbol{\mathcal{Z}}_{t-1}, \boldsymbol{\mathcal{S}}_t, \omega, \mathbf{\Theta}\right) = \mathcal{N}\left(\mathcal{Y}_t \,\middle|\, \mu_{\mathcal{Y}}(\mathcal{S}_t), \mathbf{\Omega}_{\mathcal{Y}}(\mathcal{S}_t)\right)$,
3. where $\boldsymbol{\mathcal{Z}}_t = \left[\mathcal{Y}_1' \;\; \mathcal{Y}_2' \;\ldots\; \mathcal{Y}_t'\right]'$, $\boldsymbol{\mathcal{S}}_t = \left[\mathcal{S}_0' \;\; \mathcal{S}_1' \;\ldots,\; \mathcal{S}_t'\right]'$,

$$\mu_{\mathcal{Y}}(\cdot) = \left[\mathbb{A}(\cdot)\;\mathbf{a}(\cdot)\right]\mathbf{A_0}^{-1}(\cdot)\left[\boldsymbol{\mathcal{Z}}_t \;\; 1\right]', \qquad \mathbf{\Omega}_{\mathcal{Y}}(\cdot) = \left[\mathbf{A}_0(\cdot)\mathbf{\Gamma}(\cdot)^2\mathbf{A}_0'(\cdot)\right]^{-1},$$

and

$$\mathbf{\Theta} = \left[\mathbf{A}_0(1)\;\mathbf{A}_0(2)\ldots\mathbf{A}_0(h)\;\mathbb{A}(1)\;\mathbb{A}(2)\ldots\mathbb{A}(h)\;\mathbf{a}(1)\;\mathbf{a}(2)\ldots\mathbf{a}(h)\;\mathbf{\Gamma}(1)\;\mathbf{\Gamma}(2)\ldots\mathbf{\Gamma}(h)\right]'.$$

A Structural MS-BVAR: Over-Parameterization

- A MS-BVAR becomes over-parameterized as n, p, and h rise.
  1. Let $n = p = 6$, and suppose all the slope coefficients are permitted to shift in all the regimes of a MS-BVAR.
  2. The number of coefficients per regime equals $n^2 p = 216$, which would strain the information content of sample sizes, T, often found in macro given $h \geq 2$.
- However, many regimes often last far less than T observations.
- Discipline time variation imposed on $\mathbf{A}_0(\mathcal{S}_t)$, $\mathbb{A}(\mathcal{S}_t)$, $\mathbf{\Gamma}(\mathcal{S}_t)$, and ω by MS to protect against singularities in the likelihood.

A Structural MS-BVAR: Transition Matrix Restrictions

- SWZ restrict the (first-order) Markov transition matrices $\Longrightarrow$ the laws of motion of the Markov chains in which the regime probabilities reside.
  1. The transition matrix $\mathfrak{T}$ permits switching only between adjacent regimes, and this switching is symmetric, which gives

$$\mathfrak{T} = \begin{bmatrix} \varrho_1 & 0.5(1-\varrho_2) & 0 & 0 & \ldots & 0 & 0 \\ 1-\varrho_1 & \varrho_2 & 0.5(1-\varrho_3) & 0 & \ldots & 0 & 0 \\ 0 & 0.5(1-\varrho_2) & \varrho_3 & 1-\varrho_4 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & \varrho_{h-1} & 1-\varrho_h \\ 0 & 0 & 0 & 0 & \ldots & 0.5(1-\varrho_{h-1}) & \varrho_h \end{bmatrix}.$$

- Estimate the MS-BVAR to obtain the transition probabilities $\varrho_1,\ \varrho_2, \ldots,\ \varrho_h$.
- The vector of Markov-chain probabilities ω are mapped into the transition matrix $\mathfrak{T}$ as $\mathfrak{T}_{\cdot,j} = \mathfrak{M}_{\cdot,j}\omega_{\cdot,j} \Longrightarrow \mathfrak{M}$ maps ω into the probability of remaining within a regime,
  1. where $\mathfrak{M}$ is a matrix of zeros and ones whose dimension is a function of the number of Markov chains and the regimes within each chain.
  2. Priors placed on duration (in periods) of remaining within a regime $\Longrightarrow \dfrac{1}{1-\varrho_i}$.

A Structural MS-BVAR: The Likelihood and Posterior

- The log likelihood of the MS-BVAR model is

$$\ln \mathcal{P}\Big(\mathcal{Z}_T \,\Big|\, \omega, \mathbf{\Theta}\Big) = \sum_{t=1}^{T} \ln\left[\sum_{\mathcal{S}_t \in H} \mathcal{P}\Big(\mathcal{Y}_t \,\Big|\, \mathcal{Z}_{t-1}, \mathcal{S}_t, \omega, \mathbf{\Theta}\Big) \mathcal{P}\Big(\mathcal{S}_t \,\Big|\, \mathcal{Z}_{t-1}, \omega, \mathbf{\Theta}\Big)\right],$$

  where sampling the density $\mathcal{P}\Big(\mathcal{S}_t \,\Big|\, \mathcal{Z}_{t-1}, \omega, \mathbf{\Theta}\Big)$ gives $\Pr(\mathcal{S}_t = i \,|\, \mathcal{S}_{t-1} = k) = \omega_{i,k}$.

- SWZ propose Gibbs sampling methods to construct $\ln \mathcal{P}\Big(\mathcal{Z}_T \,\Big|\, \omega, \mathbf{\Theta}\Big)$, the conditional densities of $\mathbf{\Theta}$, $\mathcal{P}\Big(\mathbf{\Theta} \,\Big|\, \mathcal{Z}_{t-1}, \mathcal{S}_t, \omega\Big)$, and ω, $\mathcal{P}\Big(\omega \,\Big|\, \mathcal{Z}_{t-1}, \mathcal{S}_t, \mathbf{\Theta}\Big)$.
  1. Sampling $\mathbf{\Theta}$ and ω involves a backward recursion, which generates $\mathcal{S}_T$.
  2. $\Longrightarrow$ Integrate $\mathcal{S}_T$ out of $\mathcal{P}\Big(\mathcal{Z}_T \,\Big|\, \omega, \mathbf{\Theta}\Big)$; see appendix A of SWZ.

- Evaluate MS-BVARs on the joint posterior distribution of ω and $\mathbf{\Theta}$, conditional on $\mathcal{Z}_T$ and priors, using Bayes' rule $\Longrightarrow$

$$\mathcal{P}\Big(\omega, \mathbf{\Theta} \,\Big|\, \mathcal{Z}_T\Big) \propto \mathcal{P}\Big(\mathcal{Z}_T \,\Big|\, \omega, \mathbf{\Theta}\Big) \mathcal{P}\Big(\omega, \mathbf{\Theta}\Big),$$

  where $\mathcal{P}\Big(\omega, \mathbf{\Theta}\Big) \Longrightarrow$ priors of ω and Θ to compute posterior odds for MS-BVAR(p)s differing by the number of regimes h embedded in $\mathbf{\Theta}$.

A Structural MS-BVAR: Priors

- Start from the SZ priors $\Longrightarrow \Lambda = \begin{bmatrix} \lambda_0 & \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 & \lambda_5 \end{bmatrix}$.
- Sims and Zha (2006) and SWZ impose prior restrictions to limit the dimension of the time variation of the slope coefficients, the $\mathbf{A}_j$s, and the intercepts, $\mathbf{a}$.
  1. The restrictions $[\mathbb{A}(\mathcal{S}_t)\ \mathbf{a}(\mathcal{S}_t)]' = \boldsymbol{\mathcal{F}}(\mathcal{S}_t) + \overline{\boldsymbol{\mathcal{F}}}\mathbf{A}_0(\mathcal{S}_t)$, where $\overline{\boldsymbol{\mathcal{F}}} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0}_{n\times 1} \end{bmatrix}'$ and the matrix of restrictions $\boldsymbol{\mathcal{F}}(\mathcal{S}_t)$ conforms with $[\mathbb{A}(\mathcal{S}_t)\ \mathbf{a}(\mathcal{S}_t)]'$ and $\overline{\boldsymbol{\mathcal{F}}}\mathbf{A}_0(\mathcal{S}_t)$.
  2. A mean zero prior distribution on $\boldsymbol{\mathcal{F}}(\mathcal{S}_t)$ matches the SZ random walk prior.
  3. A tighter random walk prior reduces $\mathbf{\Gamma}^{-1}(\cdot)$, which contains the factor loadings that scale the SV of $\eta_t \Longrightarrow$ increases persistence in $\mathbb{A}$.
  4. SWZ see the random walk prior as independent of beliefs about $\mathbf{\Omega}_y(\mathcal{S}_t) \Longrightarrow$ a normal prior is placed on $\mathbf{A}_0$, while the squared diagonal elements of $\mathbf{\Gamma}(\cdot)$ are drawn from the gamma distribution $\Longrightarrow$ elements of $\mathbf{\Gamma}$ have independent priors.
  5. $\Longrightarrow \mathbf{A}_0$, $\mathbf{a}$, $\mathbf{A}_1, \ldots, \mathbf{A}_p$, $\mathbf{\Gamma}$, and ω are estimated simultaneously when computing $\ln \mathcal{P}\left(\boldsymbol{\mathcal{Z}}_T \,\middle|\, \omega, \mathbf{\Theta}\right) \Longrightarrow \mathbf{\Gamma}$ is not a direct function of $\mathbf{A}_0$, $\mathbf{a}$, and $\mathbf{A}_1, \ldots, \mathbf{A}_p$.
- SWZ place a Dirichlet prior on the transition probabilities $\omega \Longrightarrow$ belief $\omega_{i,k}$ is the probability of regime i conditional on regime k given this transition is observed $\alpha_{i,k}-1$ times, which reflects uncertainty about which regimes are most likely.
- Prior on the probability, ϱ_i, of the average duration of remaining in regime i at date t given regime i at date $t-1 \Longrightarrow$ average duration in periods is $\dfrac{1}{1-\varrho_i} \Longrightarrow$ if prior is to stay in the same regime 10 quarters, ϱ_i is centered on 0.9.

A Structural MS-BVAR: Estimation, I

- Code to estimate the SWZ MS-BVAR is available in Adjemian, Bastani, Juillard, Maih, Mihoubi, Prerndia, Ratto, and Villemot (2014, "Dynare: Reference Manual Version 4.4.3," Dynare Working Papers number 1, CEPREMAP), which is available at `http://www.dynare.org/documentation-and-support/manual`.
- A description of commands to implement the code is in section 4.18 of the Dynare: Reference Manual Version 4.4.3 (pp. 87–97).
- Dynare wiki, `http://www.dynare.org/DynareWiki/TableOfContents`, has more details about using the SWZ MS-BVAR code and information updates to the code.

A Structural MS-BVAR: Estimation, II

- The steps to estimate a sequence of MS-BVARs and ask which is or are most favored by the data are
    1. set the priors of Λ, and the Dirichlet duration priors,
    2. construct the posterior mode of a MS-BVAR(p) model using optimization methods robustifed for the possibility of multiple peaks in the likelihood and a potentially flat posterior,
    3. equate the posterior mode of the MS-BVAR(p) model with initial conditions for $\boldsymbol{\Theta}$ to begin a MCMC simulator of x_1 steps,
    4. construct the posterior of an MS-BVAR(p) by generating x_2 draws from the proposals created by the MCMC simulator,
    5. choose among the competing MS-BVAR(p) models by calculating $\mathcal{P}\left(\omega, \boldsymbol{\Theta} \middle| \mathcal{Z}_T\right)$ using log marginal data densities (MDD), which are computed using the posterior distributions of the previous step, and
    6. rerun the MS-BVAR(p) model(s) most favored by the data to obtain the regime probabilities, $\varrho_1, \ldots, \varrho_h$ and regime-dependent residuals
- The MCMC simulator is a Metropolis within Gibbs algorithm $\Longrightarrow$ similar to estimating a fixed coefficient $\mathbf{K}$-model except the Gibbs step conditions on ω when drawing $\mathbf{A}_0(\mathcal{S}_t)$, $\mathbb{A}(\mathcal{S}_t)$, $\mathbf{a}(\mathcal{S}_t)$, and $\boldsymbol{\Gamma}(\mathcal{S}_t)$ and on these parameters when drawing updates of ω.

A Structural MS-BVAR: Estimation, III

- MS intercepts, $\mathbf{a}(\mathcal{S}_t)$, are not estimated in the Dynare code $\Longrightarrow$ the SZ prior drives these to zero to satisfy the Minnesota random walk prior.
- Dynare's MS-BVAR code employs an optimizer adapted from the `csminwel` software of Chris Sims $\Longrightarrow$ iterate back and forth between
  1. block that solves for $\mathbf{\Theta}$ given ω and
  2. a block that solves ω conditional $\mathbf{\Theta}$
  3. until a convergence criterion is satisfied.
- The Dynare code has no provisions for restarting a MS-BVAR at the posterior mode.
  1. There are large computational costs to generating a complete set of results for a MS-BVAR $\Longrightarrow$ only compute MDDs in a first pass at estimation.
  2. $\Longrightarrow$ Reestimate the MS-BVAR(s) preferred by the data, verify the favored MS-BVAR(s) retains most favored status,
  3. and, if it true, produce a complete set of results for that (these) model(s).

(Re-)Introduction to Recursive TVP-VARs: Primiceri (2006)

- Primiceri (2005) estimates a recursive TVP-SV-BVAR
    1. "*to provide a flexible framework for the estimation and interpretation of time variation in the systematic and nonsystematic part of monetary policy and their effect on the rest of the economy.*"
    2. ". . . *any reasonable attempt to model changes in policy, structure and their interaction must include time variation of the variance covariance matrix of the innovations,*"
    3. . . . "*time variation of the simultaneous relations among the variables of the model and heteroscedasticity of the innovations.*"

- The TVP-SV-SVAR is estimated on quarterly π_t, UR_t, and $R_{3MTB,t}$ from 1953 to 2001.
    1. The volatility of $\eta_{\pi,t}$, $\eta_{UR,t}$, and $\eta_{R_{3MTB,t}}$ falls post-1984.
    2. IRFs of π_t and UR_t w/r/t $\eta_{R_{3MTB,t}}$ are similar across time.
    3. Primiceri argues this is evidence against nonlinearities (*i.e.*, structural change) in the monetary transmission mechanism in the U.S.
    4. Impose Greenspan policy rule on the 1970s generates synthetic data matching the sample data.
    5. Systematic monetary policy was not responsible for the inflation of the 1970s.

- Smaller information set and recursive identification explains at least part of the disparities in results across Primiceri (2005) and Sims and Zha (2006).

Primiceri's TVP-SV-SVAR: The Model

- Primiceri's TVP-SV-SBVAR is

$$\mathcal{Y}_t = \mathbf{c}_t + \sum_{\ell=1}^{p} \mathbf{B}_{t,\ell}\mathcal{Y}_{t-\ell} + \mathbf{D}_t\mathbf{\Gamma}_t\eta_t, \quad \eta_t \sim \mathcal{N}\left(\mathbf{0}_{3\times1}, \mathbf{I}_3\right),$$

which is a **C**-model, where $\mathcal{Y}_t = [\pi_t \;\; UR_t \;\; R_{3MTB,t}]'$, $\eta_t = [\eta_{\pi,t} \;\; \eta_{UR,t} \;\; \eta_{R_{3MTB,t}}]'$,

$$\mathbf{c}_t = \begin{bmatrix} c_{1,t} \\ c_{2,t} \\ c_{3,t} \end{bmatrix}, \quad \mathbf{B}_{\ell,t} = \begin{bmatrix} B_{11,\ell,t} & B_{12,\ell,t} & B_{13,\ell,t} \\ B_{21,\ell,t} & B_{22,\ell,t} & B_{23,\ell,t} \\ B_{31,\ell,t} & B_{32,\ell,t} & B_{33,\ell,t} \end{bmatrix},$$

$$\mathbf{D}_t = \begin{bmatrix} 1 & 0 & 0 \\ D_{21,t} & 1 & 0 \\ D_{31,t} & D_{32,t} & 1 \end{bmatrix}, \quad \mathbf{\Gamma}_t = \begin{bmatrix} \gamma_{1,t} & 0 & 0 \\ 0 & \gamma_{2,t} & 0 \\ 0 & 0 & \gamma_{3,t} \end{bmatrix},$$

which imposes the LDL decomposition on $\mathbf{\Omega}_{\varepsilon,t} \Longrightarrow$ only recursive identifications.

- Define $\mathcal{X}_t' = \mathbf{I}_3 \otimes [\mathcal{Y}_{t-1}' \; \ldots \; \mathcal{Y}_{t-p}' \;\; 1]$ and $\mathbb{B}_t = \text{vec}\Big([\mathbf{B}_{1,t} \; \ldots \; \mathbf{B}_{p,t} \;\; \mathbf{c}_t]\Big) \Longrightarrow$ stack the regressions to create the "static" system of regressions

$$\mathcal{Y}_t = \mathcal{X}_t'\mathbb{B}_t + \varepsilon_t, \quad \varepsilon_t \sim \left(\mathbf{0}_{3\times1}, \mathbf{\Omega}_{\varepsilon,t}\right),$$

where $\mathbf{\Omega}_{\varepsilon,t} = \mathbf{D}_t\mathbf{\Gamma}_t\mathbf{\Gamma}_t'\mathbf{D}_t'$.

Primiceri's TVP-SV-SVAR: Hidden State Processes

- Assume the intercepts, slope coefficients, impact coefficients, and log of the volatility scaling on the structural shocks evolve as (driftless) random walks with mean zero Guassian innovations

$$\begin{aligned} \mathbb{B}_{t+1} &= \mathbb{B}_t + \vartheta_{t+1}, \\ D_{t+1} &= D_t + \zeta_{t+1}, \\ \ln \gamma_{t+1} &= \ln \gamma_t + \xi_{t+1}, \end{aligned}$$

where $D_t = [D_{21,t} \;\; D_{31,t} \;\; D_{32,t}]'$, $\gamma_t = [\gamma_{1,t} \;\; \gamma_{2,t} \;\; \gamma_{3,t}]'$, and

$$\text{Var}\begin{pmatrix} \eta_t \\ \vartheta_t \\ \zeta_t \\ \xi_t \end{pmatrix} = \begin{bmatrix} \mathbf{I} & 0 & 0 & 0 \\ 0 & \mathbf{\Omega}_\vartheta & 0 & 0 \\ 0 & 0 & \mathbf{\Omega}_\zeta & 0 \\ 0 & 0 & 0 & \mathbf{\Omega}_\xi \end{bmatrix},$$

- Cogley and Sargent (2005) assume η_t is correlated with ϑ_t, ζ_t, and ξ_t.
- In finance, SV is defined as $\ln \gamma_{t+1}^2 = \ln \gamma_t^2 + \Omega_\xi^{0.5} \xi_{t+1}$, $\xi_{t+1} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.

Primiceri's TVP-SV-SBVAR: Priors

- The priors are
  1. $\mathbb{B}_t \sim \mathcal{N}\left(\overline{\mathbb{B}}, \overline{\mathbf{\Omega}}_{\mathbb{B}}\right)$, $D_t^{-1} \sim \mathcal{N}\left(\overline{D}^{-1}, \overline{\mathbf{\Omega}}_{D^{-1}}\right)$, $\ln \gamma_t \sim \mathcal{N}\left(\ln \overline{\gamma}, \overline{\mathbf{\Omega}}_{\gamma}\right)$,
  2. $(\vartheta_t \vartheta_t')^{-1} \sim \mathcal{W}^{-1}\left(\overline{\mathbf{\Omega}}_{\vartheta}, g_{\vartheta}\right)$, $(\zeta_t \zeta_t')^{-1} \sim \mathcal{W}^{-1}\left(\overline{\mathbf{\Omega}}_{\zeta}, g_{\zeta}\right)$, and $(\xi_t \xi_t')^{-1} \sim \mathcal{W}^{-1}\left(\overline{\mathbf{\Omega}}_{\xi}, g_{\xi}\right)$.

- Initial conditions and priors are calibrated to the first 10 years of sample data $\Longrightarrow$ the training sample
  1. $\mathbb{B}_0 \sim \mathcal{N}\left(\widehat{\mathbb{B}}_{OLS}, \mathrm{x}_1 \mathbf{\Omega}_{\widehat{\mathbb{B}}_{OLS}}\right)$, $D_0 \sim \mathcal{N}\left(\widehat{D}_{OLS}, \mathrm{x}_2 \mathbf{\Omega}_{\widehat{\mathbf{D}}_{OLS}}\right)$, $\ln \gamma_0 \sim \mathcal{N}\left(\ln \widehat{\gamma}_{OLS}, \mathbf{I}_3\right)$,
  2. $(\vartheta_t \vartheta_t')^{-1} \sim \mathcal{W}^{-1}\left(\mathrm{x}_3^2\, g_{\vartheta}\, \mathbf{\Omega}_{\widehat{\mathbb{B}}}, g_{\vartheta}\right)$, $(\xi_t \xi_t')^{-1} \sim \mathcal{W}^{-1}\left(\mathrm{x}_6^2\, g_{\xi}\, \mathbf{I}_3, g_{\xi}\right)$, and
  3. separate the priors of the non-policy and policy block shock innovation covariance matrices as

     $$\left(\zeta_{\pi,UR,t}\zeta_{\pi,UR,t}'\right)^{-1} \sim \mathcal{W}^{-1}\left(\mathrm{x}_4^2\, g_{\zeta,\pi,UR}\, \mathbf{\Omega}_{\widehat{D}_{OLS},\pi,UR}, g_{\zeta,\pi,UR}\right) \text{ and}$$
     $$\left(\zeta_{R3MTB,t}\zeta_{R3MTB,t}'\right)^{-1} \sim \mathcal{W}^{-1}\left(\mathrm{x}_5^2\, g_{\zeta,R3MTB}\, \mathbf{\Omega}_{\widehat{D}_{OLS},R3MTB}, g_{\zeta,R3MTB}\right),$$
  4. where x_i^2 are tuning parameters, $i = 1, \ldots, 6$.

- The priors of $\mathbb{B}_t$ and $\mathbf{D}_t$ are normal, which makes the Gibbs sampling straightforward because the identification is recursive $\Longrightarrow$ estimate regression by regression.

An Aside on the Gibbs Sampler

- Simplest Gibbs samplers are a special case of the MH-MCMC algorithm.
- Gibbs sampling solves the problem
  1. of drawing from a multivariate distribution when its form is unknown.
  2. $\Longrightarrow$ Analytic computation of marginal distributions from a multivariate distribution is not possible.
  3. Joint distribution can consist of observed and latent random variables.
- Consider drawing samples of the random variables x and y having an unknown bivariate distribution $p(x, y)$.
- Suppose the conditional distributions, $p(x \mid y)$ and $p(y \mid x)$, are known.
  1. Elements of the covariance matrix of errors in x and
  2. y has the intercepts and lag coefficients of a VAR.
- Knowledge of the conditional distributions makes Gibbs sampling possible.

A Generic Gibbs Sampler

- Given initial conditions $Z^0 = (x^0, y^0)$, draw
  1. $x^1 \sim p(x \,|\, y^0)$ and $y^1 \sim p(y \,|\, x^1) \Longrightarrow Z_1 = (x^1, y^1)$,
  2. $x^2 \sim p(x \,|\, y^1)$ and $y^2 \sim p(y \,|\, x^2) \Longrightarrow Z_2 = (x^2, y^2)$,
  3. $x^3 \sim p(x \,|\, y^2)$ and $y^3 \sim p(y \,|\, x^3) \Longrightarrow Z_3 = (x^3, y^3), \ldots,$
  4. $x^j \sim p(x \,|\, y^{j-1})$ and $y^j \sim p(y \,|\, x^j) \Longrightarrow Z_j = (x^j, y^j), \ldots,$
  5. $x^{\mathcal{J}} \sim p(x \,|\, y^{\mathcal{J}-1})$ and $y^{\mathcal{J}} \sim p(y \,|\, x^{\mathcal{J}}) \Longrightarrow Z_{\mathcal{J}} = (x^{\mathcal{J}}, y^{\mathcal{J}})$.
- Yields an implicit Markov chain switching between $x \,|\, y$ and $y \,|\, x$ $\Longrightarrow$ transition probabilities moving from y to x and the converse.
- Gibbs sampling involves Monte Carlo methods because drawing Z_j from conditional probability distributions for $j = 1, 2, \ldots, \mathcal{J}$.
- This generalizes to $Z = (x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n)$.
- Key to Gibbs sampling is ordering the random variables in Z correctly.

Intuition for Gibbs Sampling

- Remember that $p(x \mid y) = \dfrac{p(x, y)}{p(y)} \Longrightarrow p(x \mid y) \propto p(x, y)$.
- Normalizing constant of $p(x \mid y)$ is the marginal distribution of y, $p(y)$ $\Longrightarrow$ since x is independent of $p(y)$, it is unaffected by changes in x.
- Repeated sampling from $p(x \mid y)$ and $p(y \mid x)$ approximates $p(x, y)$.
- Given x and y are continuously distributed, $p(y)$ restored after sampling.
- Let $p(\cdot \mid w) \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Omega_{x,y}\right)$, where $w = x, y$, and $\Omega_{x,y} = \begin{bmatrix} 1 & \varrho \\ \varrho & 1 \end{bmatrix}$.
- Approximate $p(x, y)$ using recipe for conditional Gaussian distribution.
  1. $x \mid y \sim \mathcal{N}\left(\varrho y, 1 - \varrho^2\right) \Longrightarrow x$ is conditionally normal on y,
  2. and similar for y, $y \mid x \sim \mathcal{N}\left(\varrho x, 1 - \varrho^2\right)$.

Primiceri's TVP-SV-SBVAR: MCMC Algorithm

- The Gibbs sampler described in Primiceri (2005) has errors.

- Primiceri (2005) uses a Gibbs sampler that
  1. conditions on the data and existing draw of the indicator when drawing $\ln \gamma_t$,
  2. draws the indicator conditional on $\mathcal{Y}^t$, the updated draw of $\ln \gamma_t$, and the previous draws of $\mathbb{B}_{t-1}$, D_{t-1}, $\mathbf{\Omega}_\vartheta$, $\mathbf{\Omega}_\zeta$, and $\mathbf{\Omega}_\xi$,
  3. draw $\mathbb{B}_t$, D_t, and (new) estimates of $\mathbf{\Omega}_\vartheta$, $\mathbf{\Omega}_\zeta$, and $\mathbf{\Omega}_\xi$ conditional on $\mathcal{Y}^t$ and the updated draw of $\ln \gamma_t$ and not the indicator.

- The problem is this Gibbs sampling algorithm draws from different likelihoods at each step in the algorithm . . . more about the problem below.

Pros and Cons of TVP-SV-VARs and MS-BVARs

- TVP-SV-BVARs and MS-BVARs each have pluses and minuses.
- TVP-SV-SBVARs are easy to estimate and compute time-dependent IRFs and FEVDs.
- TVP-SV-VARs have several problems.
  1. Cogley and Sargent (2002, 2005) and Primiceri (2005) estimate recursive SBVARs $\Longrightarrow$ limited identification schemes.
  2. SBVARs lack a steady state $\Longrightarrow$ random walks source of time-variation in $\mathbb{B}_t$, D_t, and γ_t $\Longrightarrow$ initial conditions are found in a training sample not the likelihood.
  3. Stationarity of a SBVARs is imposed using inequality $\Longrightarrow$ priors on $(\mathbb{B}_t, D_t)$ should inform stationarity of a SBVAR constraints.
  4. Cogley and Sargent invoke a "virtual prior" that tosses draws of any $\left(\mathbb{B}_t, D_t\right) \in \left\{\mathbb{B}_t, D_t\right\}_{t=1}^{T}$ that have roots inside the unit circle.
  5. This rejection procedure often requires a large number of draws for the MCMC to converge $\Longrightarrow$ posterior MSEs display wide coverage intervals.
  6. Probability a draw $\left(\mathbb{B}_t, D_t\right) \in \left\{\mathbb{B}_t, D_t\right\}_{t=1}^{T}$ has roots inside the unit circle $\longrightarrow 1$ $\Longrightarrow$ reject every draw; see Koop and Potter (2011).
- A MS-BVARs recovers estimates of the Markov transition probabilities, ϱ_i, $i = 1, \ldots, h$, which is unique, but economic identification of ϱ_i is arbitrary.
  1. Computing IRFs and FEVDs is difficult except in special cases and
  2. estimating MS-BVARs are computationally and time intensive.

Canova and Pérez Forero (2015): Introduction

- Canova and Pérez Forero (CPF) develop methods to estimate TVP-SV-BSVARs in which $\mathbf{A}_{0,t}$ is non-recursive.
- Restrictions on $\mathbf{A}_{0,t}$ can be linear or nonlinear.
  1. Linear restrictions are imposed on $\mathbf{A}_{0,t}$ using format of Amisano and Giannini (1997) $\Longrightarrow$ Kalman filter and smoother generate TVPs for MH step in a Gibbs sampler algorithm to estimate SBVARs.
  2. CPF rely on the extended Kalman filter to approximate nonlinear restrictions, which involve long-run or sign restrictions $\Longrightarrow$ the extended Kalman filter replaces the Kalman filter in the MH in Gibbs sampler simulators.
- Study the Gordon and Leeper (1994) and Leeper and Roush (2003) critiques of recursive identifications of SVARs $\Longrightarrow$ monetary policy makers and financial market participants interact in money markets.
- Also, can explore the impact of Leeper and Zha (2003) immodest monetary policy actions on the real economy.

Canova and Pérez Forero (2015): A Fixed Coefficient Example

- Consider the fixed coefficient-static structural vector regression (SVR) $\Longrightarrow \mathbf{A}_0 y_t = \eta_t,\ \eta_t \sim \mathcal{N}(\mathbf{0}_{n\times 1}, \mathbf{I}_n)$.

- The likelihood of this SVR is

$$\mathcal{H}\left(\mathbf{Y} \middle| \boldsymbol{a}_0\right) = (2\pi)^{-0.5NT} \det\left|\mathbf{A}_0\right|^{T} \exp\left\{-\frac{1}{2}\sum_{t=1}^{T} y_t'\mathbf{A}_0'\mathbf{A}_0 y_t\right\}.$$

- $\mathcal{H}\left(\mathbf{Y} \middle| \boldsymbol{a}_0\right)$ is nonlinear in $\boldsymbol{a}_0 \Longrightarrow$ a source of the Bayesian estimation problem.

- Recursive identifications draw from $\boldsymbol{\Omega}^{-1} = \mathbf{A}_0'\mathbf{A}_0 \Longrightarrow$ triangular restrictions.

- Otherwise, need to recast non-recursive SVR as a linear system.

Canova and Pérez Forero (2015): A Static Fixed Coefficient SVR

- Trick is to vectorize the fixed coefficient-static SBVR: $\text{vec}(\mathbf{A}_0 y_t) = \eta_t$.
  1. Amisano and Giannini (1997) explicit form linear restrictions of the **K**-SVAR are $\text{vec}(\mathbf{A}_0) = \mathbf{S}_{\mathbf{A}_0}\boldsymbol{a}_0 + \mathbf{s}_{\mathbf{A}_0}$, where $\mathbf{S}_{\mathbf{A}_0}$ and $\mathbf{s}_{\mathbf{A}_0}$ are matrices of zeros and ones.
  2. $\Longrightarrow \text{vec}\left(\mathbf{A}_0 y_t\right) = \left(y_t' \otimes \mathbf{I}_{n^2}\right)\left[\mathbf{S}_{\mathbf{A}_0}\boldsymbol{a}_0 + \mathbf{s}_{\mathbf{A}_0}\right]$.
  3. $\Longrightarrow \left(y_t' \otimes \mathbf{I}_{n^2}\right)\mathbf{s}_{\mathbf{A}_0} = -\left(y_t' \otimes \mathbf{I}_{n^2}\right)\mathbf{S}_{\mathbf{A}_0}\boldsymbol{a}_0 + \eta_t$ or $\tilde{y}_t = \boldsymbol{z}_t\boldsymbol{a}_0 + \eta_t \Longrightarrow$ only linear restrictions imposed on system of regressions.
  4. The SVR's likelihood becomes

$$\mathcal{H}\left(\mathbf{Y}\middle|\boldsymbol{a}_0\right) = (2\pi)^{-0.5NT}\det\left|\frac{\partial\text{vec}\left(\mathbf{A}_0 y_t\right)}{\partial y_t'}\right|^{T}\exp\left\{-\frac{1}{2}\sum_{t=1}^{T}\left[\tilde{y}_t - \boldsymbol{z}_t\boldsymbol{a}_0\right]'\left[\tilde{y}_t - \boldsymbol{z}_t\boldsymbol{a}_0\right]\right\},$$

where $\dfrac{\partial\text{vec}\left(\mathbf{A}_0 y_t\right)}{\partial y_t'} = \dfrac{\left(y_t' \bigotimes \mathbf{I}_{n^2}\right)\mathbf{S}_{\mathbf{A}_0}\boldsymbol{a}_0}{\partial y_t'} + \dfrac{\left(y_t' \bigotimes \mathbf{I}_{n^2}\right)\mathbf{s}_{\mathbf{A}_0}}{\partial y_t'}$,

$\text{vec}\left(\dfrac{\left(y_t' \bigotimes \mathbf{I}_{n^2}\right)\mathbf{S}_{\mathbf{A}_0}\boldsymbol{a}_0}{\partial y_t'}\right) = \mathbf{S}_{\mathbf{A}_0}\boldsymbol{a}_0$ and $\text{vec}\left(\dfrac{\left(y_t' \bigotimes \mathbf{I}_{n^2}\right)\mathbf{s}_{\mathbf{A}_0}}{\partial y_t'}\right) = \mathbf{s}_{\mathbf{A}_0}$.

  5. Given the system of regressions $\tilde{y}_t = \boldsymbol{z}_t\boldsymbol{a}_0 + \eta_t$, let $\hat{\boldsymbol{a}}_0 = \left[\sum_{t=1}^{T}\boldsymbol{z}_t'\boldsymbol{z}_t\right]^{-1}\left[\sum_{t=1}^{T}\boldsymbol{z}_t'\tilde{y}_t\right]$ and $\hat{\boldsymbol{\Omega}}_{\boldsymbol{a}_0} = \left[\sum_{t=1}^{T}\boldsymbol{z}_t'\left(\sum_{t=1}^{T}\hat{\eta}_t'\hat{\eta}_t\right)^{-1}\boldsymbol{z}_t\right]^{-1}$, where $\hat{\eta}_t = \tilde{y}_t - \boldsymbol{z}_t\hat{\boldsymbol{a}}_0$.

Canova & Pérez Forero (2015): A Sampler for a Static Fixed Coefficient SBVR

- Multi-step sampler is initialized at the "OLS" estimates, $\boldsymbol{a}_{0,0} = \hat{\boldsymbol{a}}_0$ and equates $\boldsymbol{\Omega}_{\boldsymbol{a}_0,j}$ with the (inverse) precision of $\boldsymbol{a}_{j,0}$, $j = 1, \ldots, \mathcal{J}$.
- Running the multi-step sampler at iteration j involves
  1. sample a potential $\boldsymbol{a}_0^{\blacklozenge} \sim \mathcal{P}\left(\boldsymbol{a}_{j,0} \middle| \boldsymbol{a}_{j-1,0}\right) = t\left(\boldsymbol{a}_{j-1,0}, \mathrm{r}\boldsymbol{\Omega}_{\boldsymbol{a}_0,j-1}, \nu\right)$, where $0 < r,\ \ 4 \leq \nu$, and $t(\cdot, \cdot, \cdot)$ is the t-distribution,
  2. construct $\kappa = \dfrac{\mathcal{P}\left(\boldsymbol{a}_0^{\blacklozenge} \middle| \tilde{\mathbf{Y}}\right) \mathcal{P}\left(\boldsymbol{a}_{j,0} \middle| \boldsymbol{a}_{j-1,0}\right)}{\mathcal{P}\left(\boldsymbol{a}_{j-1,0} \middle| \tilde{\mathbf{Y}}\right) \mathcal{P}\left(\boldsymbol{a}_{j-1,0} \middle| \boldsymbol{a}_{j,0}\right)}$, where the posterior of $\boldsymbol{a}_0$ is $\mathcal{P}\left(\boldsymbol{a}_0 \middle| \tilde{\mathbf{Y}}\right) = \mathcal{H}\left(\tilde{\mathbf{Y}} \middle| \boldsymbol{a}_0\right) \mathcal{P}\left(\boldsymbol{a}_0 \middle| \cdot\right) \mathcal{I}_{\boldsymbol{a}}$ and $\mathcal{I}_{\boldsymbol{a}}$ is an indicator function (*i.e.*, either zero or one) flagging whether $\mathbf{A}_0$ is full rank, and
  3. the update is $\boldsymbol{a}_{j,0} = \boldsymbol{a}_0^{\blacklozenge}$ if $\kappa > u \sim \mathcal{U}(0, 1) \Longrightarrow$ otherwise $\boldsymbol{a}_{j,0} = \boldsymbol{a}_{j-1,0}$.
- Algorithm need tails of the t-distribution to capture the range of the posterior of $\boldsymbol{a}_0$,
  1. non-zero elements of $\boldsymbol{a}_{j,0}$ are sampled jointly $\Longrightarrow \boldsymbol{\Omega}_{\boldsymbol{a}_0,j}$ is not diagonal, and
  2. the MH step is needed because $\hat{\boldsymbol{a}}_0$ and $\hat{\boldsymbol{\Omega}}_{\boldsymbol{a}_0}$ are OLS constructs $\Longrightarrow$ ignore information in $\dfrac{\partial \mathrm{vec}\left(\mathbf{A}_0 \mathcal{Y}_t\right)}{\partial \mathcal{Y}_t'}$ about $\mathcal{H}\left(\tilde{\mathbf{Y}} \middle| \boldsymbol{a}_0\right) \Longrightarrow$ maybe a problem when the Jacobian matters for the shape of the likelihood and thus the posterior.

Canova and Pérez Forero (2015): A Static TVP-SBVR Example

- Time variation in $\mathbf{A}_0$ complicates estimation of the static SBVR $\Longrightarrow$ it is nonlinear in $\mathbf{A}_{0,t}$ and the posterior distribution of $\mathbf{A}_{0,t}$ conditions on $\mathcal{Z}_t \Longrightarrow \mathcal{Z}_t$ is endogenous.
- The static TVP-SBVR is $\mathbf{A}_{0,t}\mathcal{Y}_t = \eta_t$, where $\text{vec}(\mathbf{A}_{0,t}) = \boldsymbol{a}_{0,t}$ and $\boldsymbol{a}_{0,t+1} = \boldsymbol{a}_{0,t} + \psi_{t+1}$, $\psi_{t+1} \sim \mathcal{N}\left(\mathbf{0}_{n^2}, \boldsymbol{\Omega}_\psi\right)$, and given an initial value for $\mathbf{A}_0$, which is labeled $\mathbf{A}_{0,0}$.
  1. Assume $\boldsymbol{a}_{0,t} \in \boldsymbol{\mathcal{A}} \subset \mathbb{R}^{k_1} \Longrightarrow \boldsymbol{\mathcal{A}}$ is large "enough" to hold draws that maybe eliminated.
  2. Given $f_t = f(\boldsymbol{a}_{0,t})$, which evolves as $f_{t+1} = f_t + \mu_{t+1}$, $\mu_{t+1} \sim \mathcal{N}\left(\mathbf{0}_{n^2}, \boldsymbol{\Omega}_\mu\right)$, where $\boldsymbol{\Omega}_\mu$ is a full rank and pd matrix, the static TVP-SBVR is $\tilde{\mathcal{Y}}_t = \mathcal{Z}_t f_t + \eta_t$.
  3. The static TVP-SBVR's likelihood is

$$\mathcal{H}\left(\tilde{\mathbf{Y}} \middle| f_{0,t}, \boldsymbol{\Omega}_\mu\right) = (2\pi)^{-0.5NT} \det\left|\mathcal{F}\left(\boldsymbol{a}_{0,t}\right)\right|^T \exp\left\{-\frac{1}{2}\sum_{t=1}^{T}\left[\tilde{\mathcal{Y}}_t - \mathcal{Z}_t f_t\right]'\left[\tilde{\mathcal{Y}}_t - \mathcal{Z}_t f_t\right]\right\},$$

  where $\mathcal{F}\left(\boldsymbol{a}_{0,t}\right) = \mathbf{S}_{\mathbf{A}_0} f_t + \mathbf{s}_{\mathbf{A}_0}$.

- Need to generate the joint distribution of $f_{1:T} = \left\{f_t\right\}_{t=1}^T$ and $\boldsymbol{\Omega}_\mu$.
  1. $\Longrightarrow$ Gibbs sampling but need $\mathcal{P}\left(f_{1:T} \middle| \tilde{\mathbf{Y}}, \boldsymbol{\Omega}_\mu\right)$ and $\mathcal{P}\left(\boldsymbol{\Omega}_\mu \middle| \tilde{\mathbf{Y}}, f_{1:T}\right)$,
  2. where $\mathcal{P}\left(\boldsymbol{\Omega}_\mu \middle| \tilde{\mathbf{Y}}, f_{1:T}\right) \sim \mathcal{IW}$ by choosing the correct assumptions.

Canova and Pérez Forero (2015): A Static TVP-SBVR Simulator, I

- If we could sample $\mathcal{P}\left(f_{1:T}\,\middle|\,\mathbf{Y},\,\mathbf{\Omega}_{\mu}\right)$, estimating the static TVP-SBVR becomes a straightforward Gibbs sampling problem.

- The nonlinear static TVP-SBVR, $\tilde{y}_t = \mathcal{Z}_t f_t + \eta_t$, is the problem, but given f_t the static TVP-SBVR is conditionally linear $\Longrightarrow$ a Rao-Blackwellization procedure to create a more efficient simulator.

- The random walk assumption suggests the Kalman filter (KF) and smoother (KS) provide "prior draws" for f_t and its MSE, given initial conditions $f_{0|0}$ and $\mathbf{\Sigma}_{f,0|0}$.

- The static TVP-SBVR are observation equations and $f_t = f_{t-1} + \mu_t$ are state equations.
  1. KF updates of f_t are $f_{t|t} = f_{t|t-1} + \mathcal{K}_t\left[\tilde{y}_t - \mathcal{Z}_t f_{t|t-1}\right]$, $t = 1, \ldots, T$,
  2. MSE updates of f_t are $\mathbf{\Sigma}_{f,t|t} = \mathbf{\Sigma}_{f,t|t-1} + \mathbf{\Sigma}_{f,t|t-1}\mathcal{Z}_t'\mathbf{\Sigma}_{\mathcal{Z},t|t-1}^{-1}\mathcal{Z}_t\mathbf{\Sigma}_{f,t|t-1}'$,
  3. where $f_{t|t-1} = f_{t-1|t-1}$, the Kalman gain is $\mathcal{K}_t = \mathbf{\Sigma}_{f,t|t-1}\mathcal{Z}_t'\mathbf{\Sigma}_{\mathcal{Z},t|t-1}^{-1}$, $\mathbf{\Sigma}_{f,t|t-1} = \mathbf{\Sigma}_{f,t-1|t-1} + \mathbf{\Omega}_{\mu}$, and the MSE of $\mathcal{Z}_t$ is $\mathbf{\Sigma}_{\mathcal{Z},t|t-1} = \mathcal{Z}_t'\mathbf{\Sigma}_{f,t|t-1}\mathcal{Z}_t + \mathbf{I}_{n^2}$.

- Starting from smoothed $\overleftrightarrow{f}_{T|T} = f_{T|T}$ and $\overleftrightarrow{\mathbf{\Sigma}}_{f,T|T} = \mathbf{\Sigma}_{f,T|T}$, the KS is run to produce
  1. smoothed updates, $\overleftrightarrow{f}_{t|t+1} = f_{t|t} + \mathbf{\Sigma}_{f,t|t}\mathcal{Z}_t'\mathbf{\Sigma}_{f,t+1|t}^{-1}\left(\overleftrightarrow{f}_{t+1|t+2} - \mathcal{Z}_t' f_{t|t}\right)$, and
  2. MSEs, $\overleftrightarrow{\mathbf{\Sigma}}_{f,t|t+1} = \mathbf{\Sigma}_{f,t|t} - \mathbf{\Sigma}_{f,t|t}\mathcal{Z}_t'\mathbf{\Sigma}_{f,t+1|t}^{-1}\mathcal{Z}_t\mathbf{\Sigma}_{f,t|t-1}'$, $t = T-1, \ldots, 1$.

Canova and Pérez Forero (2015): A Static TVP-SBVR Simulator, II

- Set initial conditions $f_{0|0}$ and $\mathbf{\Sigma}_{f,0|0}$, draw $\mathbf{\Omega}_{\mu,0} \sim \mathcal{IW}(\underline{\mathbf{\Omega}}_{\mu}, \nu)$, given $\underline{\mathbf{\Omega}}_{\mu}$ and $\tilde{\mathcal{Y}}_t$.

- Run the multi-step sampler for $j = 1, \ldots, \mathcal{J}$ iterations to construct $\left\{ \overleftrightarrow{f}_{j,t|t+1} \right\}_{t=1}^{T}$ and $\left\{ \overleftrightarrow{\mathbf{\Sigma}}_{f,j,t|t+1} \right\}_{t=1}^{T}$ using the KS updating equations.
  1. Draw $f_t^{\blacklozenge} \sim \mathcal{P}\left(f_{j,t} \,\middle|\, f_{j-1,t}\right) = t\left(\overleftrightarrow{f}_{j,t|t+1}, \mathrm{r}\overleftrightarrow{\mathbf{\Sigma}}_{f,j-1,t|t+1}, \nu\right)$, where $0 < \mathrm{r}$ and $4 \leq \nu$.
  2. Define $\boldsymbol{f}^{\blacklozenge,T} \equiv \left\{f_t^{\blacklozenge}\right\}_{t=1}^{T}$ and $\mathcal{P}\left(\boldsymbol{f}_j^{\blacklozenge,T} \,\middle|\, \boldsymbol{f}_{j-1}^{\blacklozenge,T}\right) \equiv \prod_{t=1}^{T} \mathcal{P}\left(f_{j,t} \,\middle|\, f_{j-1,t}\right)$.
  3. Next, construct $\kappa = \dfrac{\mathcal{P}\left(\boldsymbol{f}^{\blacklozenge,T} \,\middle|\, \tilde{\mathbf{Y}}\right) \mathcal{P}\left(\boldsymbol{f}_{j-1}^{\blacklozenge,T} \,\middle|\, \boldsymbol{f}_{j}^{\blacklozenge,T}\right)}{\mathcal{P}\left(\boldsymbol{f}_{j-1}^{T} \,\middle|\, \tilde{\mathbf{Y}}\right) \mathcal{P}\left(\boldsymbol{f}^{\blacklozenge,\boldsymbol{T}}{}_{j} \,\middle|\, \boldsymbol{f}^{\blacklozenge,\boldsymbol{T}}{}_{j-1}\right)}$, where the posterior of $\boldsymbol{f}^T$ is $\mathcal{P}\left(\boldsymbol{f}^T \,\middle|\, \mathbf{Y}\right) = \mathcal{H}\left(\tilde{\mathbf{Y}} \,\middle|\, \boldsymbol{f}^T, \mathbf{\Omega}_{\mu}\right) \mathcal{P}\left(\boldsymbol{f}^T\right) \mathcal{I}_f$, $\mathcal{P}\left(\boldsymbol{f}^T\right)$ is the prior of f_t, and $\mathcal{I}_f$ is an indicator function (*i.e.*, either zero or one) flagging whether $\mathbf{A}_0$ is full rank, and
  4. the update is $\boldsymbol{f}_j^T = \boldsymbol{f}^{\blacklozenge,T}$ if $\kappa > u \sim \mathcal{U}(0, 1) \Longrightarrow$ otherwise $\boldsymbol{f}_j^T = \boldsymbol{f}_{j-1}^T$.
  5. The final step is to draw $\mathbf{\Omega}_{\mu,j}^{-1} \sim \mathcal{P}\left(\mathbf{\Omega}_{\mu,j-1}^{-1} \,\middle|\, \boldsymbol{f}_j^T, \tilde{\mathbf{Y}}\right) = \mathcal{IW}(\overline{\mathbf{\Omega}}_{\mu}, \overline{\nu})$, given $\boldsymbol{f}_j^T$ and $\tilde{\mathbf{Y}}$, where $\overline{\nu} = T + \underline{\nu}$ and $\overline{\mathbf{\Omega}}_{\mu}^{-1} = \left[\underline{\mathbf{\Omega}}_{\mu} + \sum_{t=1}^{T} \left(\boldsymbol{f}_{j,t} - \boldsymbol{f}_{j,t-1}\right)\left(\boldsymbol{f}_{j,t} - \boldsymbol{f}_{j,t-1}\right)'\right]^{-1}$.

Canova and Pérez Forero (2015): The TVP-SV-SBVAR

- The **K**-model version of the standard TVP-SV-SBVAR is

$$\mathbf{A}_{0,t}\mathcal{Y}_t \;=\; \mathbf{A}_{0,t}\mathbf{c}_t \;+\; \mathbf{A}_{0,t}\sum_{\ell=1}^{p}\mathbf{B}_{t,\ell}\mathcal{Y}_{t-\ell} \;+\; \mathbf{\Gamma}_t\eta_t, \quad \eta_t \sim \mathcal{N}\left(\mathbf{0}_{n\times 1}, \mathbf{I}_n\right),$$

  where $\mathbf{\Gamma}_t = \text{diag}(\gamma_{1,t} \, \dots \, \gamma_{p,t})$.

- The reduced form intercepts and slope coefficients, impact coefficients, and volatility scaling of the structural errors evolve as (driftless) random walks with mean zero Gaussian innovations

$$\begin{aligned} \mathbb{B}_{t+1} &= \mathbb{B}_t + \vartheta_{t+1}, \\ \boldsymbol{a}_{0,t+1} &= \boldsymbol{a}_{0,t} + \psi_{t+1}, \\ \ln\gamma_{t+1} &= \ln\gamma_t + \xi_{t+1}, \end{aligned}$$

  where

$$\boldsymbol{\mathcal{V}} \equiv \text{Var}\begin{pmatrix} \eta_t \\ \vartheta_t \\ \psi_t \\ \xi_t \end{pmatrix} = \begin{bmatrix} \mathbf{I} & 0 & 0 & 0 \\ 0 & \mathbf{\Omega}_\vartheta & 0 & 0 \\ 0 & 0 & \mathbf{\Omega}_\psi & 0 \\ 0 & 0 & 0 & \mathbf{\Omega}_\xi \end{bmatrix}.$$

Canova and Pérez Forero (2015): Casting the TVP-SV-SBVAR in State Space Form

- Remember $\mathcal{X}'_t = \mathbf{I}_n \otimes \left[y'_{t-1} \ \ldots \ y'_{t-p} \ \ 1 \right]$ and $\mathbb{B}_t = \text{vec}\left(\left[\mathbf{B}_{1,t} \ \ldots \ \mathbf{B}_{p,t} \ \ \mathbf{c}_t\right]\right) \Longrightarrow$ the standard TVP-SV-SVAR can be written in concentrated form

$$\mathbf{A}_{0,t}\left(y_t - \mathcal{X}'_t \widehat{\mathbb{B}}_t\right) = \boldsymbol{\Gamma}_t \eta_t,$$

where $\widehat{\mathbb{B}}_t$ are estimates of the reduced from TVP intercepts and slope coefficients $\Longrightarrow$ draw $\mathbf{A}_{0,t}$ and $\boldsymbol{\Gamma}_t$ conditional on estimates of $\mathbb{B}_t$.

- Define $\hat{y}_t \equiv y_t - \mathcal{X}'_t \widehat{\mathbb{B}}_t$, $\text{vec}\left(\mathbf{A}_{0,t}\right) = \mathbf{S}_{\mathbf{A}_0} \boldsymbol{f}_t + \mathbf{s}_{\mathbf{A}_0}$, $f_t = f\left(\boldsymbol{a}_{0,t}\right)$, $\mathbf{S}_{\mathbf{A}_0}$ is $n^2 \times \dim(f_t)$, and $\mathbf{s}_{\mathbf{A}_0}$ is column vector of length n^2.
  1. $\Longrightarrow$ Reparameterize the TVP-SV-SVAR as a "static" system of regressions $\left(\hat{y}'_t \otimes \mathbf{I}_{n^2}\right)\left[\mathbf{S}_{\mathbf{A}_0} \boldsymbol{f}_t + \mathbf{s}_{\mathbf{A}_0}\right] = \boldsymbol{\Gamma}_t \eta_t$.
  2. Let $\tilde{y}'_t \equiv \left(\hat{y}'_t \otimes \mathbf{I}_{n^2}\right) \mathbf{s}_{\mathbf{A}_0}$ and $\mathcal{Z}_t = -\left(\hat{y}'_t \otimes \mathbf{I}_{n^2}\right) \mathbf{S}_{\mathbf{A}_0} \Longrightarrow \tilde{y}'_t = \mathcal{Z}_t f_t + \boldsymbol{\Gamma}_t \eta_t$, which is the system of observation equations and the system of state equations is $f_{t+1} = f_t + \mu_{t+1}$, $\mu_{t+1} \sim \mathcal{N}\left(\mathbf{0}_{n^2}, \boldsymbol{\Omega}_\mu\right)$, where $\boldsymbol{\Omega}_\mu = \boldsymbol{\Omega}_\psi$.

Canova and Pérez Forero (2015): MCMC Sampling of the TVP-SV-SBVAR

- The sampling algorithm needs
  1. to draw $f^T \sim \mathcal{P}\left(f^T \,\middle|\, \tilde{\mathbf{Y}},\ \gamma^T,\ \boldsymbol{\mathcal{V}},\ \hat{\mathbb{B}}^T\right)$, where $f^T \equiv \left\{f_t\right\}_{t=1}^T$, $\gamma^T \equiv \{\gamma_t\}_{t=1}^T$, and $\hat{\mathbb{B}}^T \equiv \left\{\hat{\mathbb{B}}_t\right\}_{t=1}^T \Longrightarrow$ sample f^T conditional on γ^T, $\boldsymbol{\mathcal{V}}$, and $\hat{\mathbb{B}}^T$.
  2. Joint sampling of f_t, where $\mathbf{\Omega}_\mu$ is non-diagonal for the TVP-SV-SBVAR.
  3. Similar to the TVP-SBVR sampling algorithm in which $\mathbf{\Omega}_\mu$ is non-diagonal.

- Algorithms for drawing f^T and $\hat{\mathbb{B}}^T$ have existed for 25 years or more.
  1. A leading example is Carter and Kohn (1994, "On Gibbs sampling for state space models," *Biometrika* 81, 541–553).
  2. They propose to sample f^T and $\hat{\mathbb{B}}^T$ using the Kalman filter and smoother.
  3. Within Kalman filtering and smoothing routines, estimate the covariance matrices in $\boldsymbol{\mathcal{V}}$ by drawing updates from $\mathcal{IW}$ distributions.

- Unresolved is the problem of sampling from the posterior of γ^T.

Canova and Pérez Forero (2015): How to Estimate SV in the TVP-SV-SBVAR

- Credit Primiceri (2005) for adapting an estimator developed in the finance literature for modeling the SV of asset returns.

- CPF apply the same approach to the system of static regressions, $\widehat{\mathbf{A}}_{0,t}\widehat{y}_t = \mathbf{\Gamma_t}\eta_t$, which is conditionally linear on $\widehat{y}^T$ and $\widehat{\mathbf{A}}_{0,t}$ evaluate at $\boldsymbol{f}_j^T$, $j = 1, \ldots, \mathcal{J}$.

  1. Define $\widehat{\widehat{y}}_t \equiv \widehat{\mathbf{A}}_{0,t}\widehat{y}_t \Longrightarrow$ the system of static regressions $\widehat{\widehat{y}}_t = \mathbf{\Gamma_t}\eta_t$.
  2. Square both sides of its ℓ-th equation, pass the ln through, and add a small constant, c, to bound $\ln\widehat{\widehat{y}}^2_{\ell,t}$ away from $-\infty \Longrightarrow$ approximate the heteroskedastic variance as $\ln\left(\widehat{\widehat{y}}^2_{\ell,t} + c\right) \approx 2\ln\gamma_{\ell,t} + \ln\eta^2_{\ell,t}$, where $\ell = 1, \ldots, n$.
  3. Assuming Gaussian structural errors $\Longrightarrow \ln\eta^2_{\ell,t} \sim \ln\chi^2(1)$, which has a mean $= -1.2704$ and variance $= 0.5\pi^2$; see Harvey, Ruiz, and Shephard (1994, "Multivariate stochastic variance models," *Review of Economic Studies* 61, pp. 247–264).

Canova and Pérez Forero (2015): How to Estimate SV in the TVP-SV-SBVAR, cont.

- Primiceri draws $\ln \gamma_{\ell,t}$ by adapting the sampling algorithm of
  1. Kim, Shephard, and Chib (1998, "Stochastic volatility: Likelihood inference and comparison with ARCH models," *Review of Economic Studies* 65, 361–393).
  2. Before drawing $\ln \gamma_{\ell,t}$, sample a discrete indicator, $\mathcal{s}_t$, of the underlying volatility state of the SBVAR $\Longrightarrow$ information about the structural shocks, η_t.
  3. The volatility state $\mathcal{s}_t$ is drawn from a mixture normal distribution that is calibrated to approximate a $\ln \chi^2(1)$ random variable.
  4. Assume there are $\mathcal{K}$ volatility states, $\mathcal{s}_t \equiv [\mathcal{s}_{1,t} \ \mathcal{s}_{2,t} \ \dots \ \mathcal{s}_{\mathcal{K},t}]'$ and $\mathcal{S}^T \equiv \{\mathcal{s}_t\}_{t=1}^T$.
  5. The $\mathcal{K}$ volatility states match the $\mathcal{K}$ normal distributions in the mixture.

- The calibration of the mixture normal distribution is not model dependent.
  1. See table 4 of Kim, Shephard, and Chib (1998) for a calibration involving a 7-point mixture normal.
  2. A more refined 10-point mixture is in table 1 of Omori, Chib, Shephard, and Nakajima (2007, "Stochastic volatility with leverage: Fast and efficient likelihood inference," *Journal of Econometrics* 140, 425–449).

The MCMC Algorithm of Del Negro and Primiceri (2015)

- As noted previously, Primiceri (2005) employs a flawed sampler to estimate recursive TVP-SV-SBVARs (remember he draws model parameters from different likelihoods at each step in the algorithm).

- Del Negro and Primiceri (2015) place the source of the problem in the conditioning information needed to sample s_t efficiently from a mixture normal distribution that approximates the true $\ln \chi^2(1)$ distribution.

- Their fix is to draw $\mathbb{S}^T$ after $\mathbb{B}^T$ and f^T, but before $\ln \gamma^T$, and $\mathbf{\Omega}_\vartheta$, $\mathbf{\Omega}_\psi$, and $\mathbf{\Omega}_\xi$.
  1. Draw $\mathbb{B}^T$ given $\mathcal{Y}^T$ and current draws of f^T, $\mathbb{S}^T$, $\ln \gamma^T$, $\mathbf{\Omega}_\vartheta$, $\mathbf{\Omega}_\psi$, and $\mathbf{\Omega}_\xi$.
  2. Draw f^T conditional on $\mathcal{Y}^T$, the update of $\mathbb{B}^T$ and current draws of $\mathbb{S}^T$, $\ln \gamma^T$, $\mathbf{\Omega}_\vartheta$, $\mathbf{\Omega}_\psi$, and $\mathbf{\Omega}_\xi$.
  3. Draw $\mathbb{S}^T$ conditional on $\mathcal{Y}^T$, the updates of $\mathbb{B}^T$ and f^T, and the current accepted draws of $\ln \gamma^T$, $\mathbf{\Omega}_\vartheta$, $\mathbf{\Omega}_\psi$, and $\mathbf{\Omega}_\xi$.
  4. Draw $\ln \gamma^T$ conditional on $\mathcal{Y}^T$, the updates of $\mathbb{B}^T$, f^T, and $\mathbb{S}^T$, and the current accepted draws of $\ln \gamma^T$, $\mathbf{\Omega}_\vartheta$, $\mathbf{\Omega}_\psi$, and $\mathbf{\Omega}_\xi$.
  5. Draw $\mathbf{\Omega}_\vartheta$, $\mathbf{\Omega}_\psi$, and $\mathbf{\Omega}_\xi$ conditional on $\mathcal{Y}^T$ and updates of $\mathbb{B}^T$, f^T, $\mathbb{S}^T$, and $\ln \gamma^T$.

Canova and Pérez Forero (2015): A Sampling Algorithm for the TVP-SV-SBVAR

- The multi-step algorithm generating the posterior of the revised TVP-SV-SVAR is
  1. starts by initializing $\left[\mathbb{B}_0^T \; f_0^T \; \mathcal{S}_0^T \; \gamma_0^T \; \boldsymbol{\mathcal{V}}_0\right]$, followed by iterating from $j = 1$ to $\mathcal{J}$,
  2. sample $\mathbb{B}_j^T \sim \mathcal{P}\left(\mathbb{B}_j^T \middle| \tilde{\mathbf{Y}}, f_{j-1}^T, \mathcal{S}_{j-1}^T, \gamma_{j-1}^T, \boldsymbol{\mathcal{V}}_{j-1}\right) \mathcal{I}\left(\mathbb{B}_j^T\right)$, where $\mathcal{I}\left(\mathbb{B}_j^T\right)$ is an indicator function tossing draws of $\mathbb{B}^T$ with roots inside the unit circle and $\mathcal{P}\left(\mathbb{B} \middle| \tilde{\mathbf{Y}}, \cdot, \cdot, \cdot, \cdot\right) \sim \mathcal{N}\left(\mathbb{B}, \boldsymbol{\Omega}_{\mathbb{B}}\right) \Longrightarrow$ several samplers are available to draw $\mathbb{B}^T$,
  3. next draw $f_j^T \sim \mathcal{P}\left(f_j^T \middle| \tilde{\mathbf{Y}}, \mathcal{S}_{j-1}^T, \gamma_{j-1}^T, \boldsymbol{\mathcal{V}}_{j-1}, \mathbb{B}_j^T, \right)$ and use the KS updating equations to sample $f_t^{\blacklozenge} \sim \mathcal{P}\left(f_{j,t} \middle| f_{j-1,t}\right) = t\left(\overleftrightarrow{f}_{j,t|t+1}, \mathrm{r}\overleftrightarrow{\boldsymbol{\Sigma}}_{f,j-1,t|t+1}, \nu\right)$, where $0 < \mathrm{r}$ and $4 \leq \nu$,
  4. construct $\kappa = \dfrac{\mathcal{P}\left(\boldsymbol{f}^{\blacklozenge,T} \middle| \tilde{\mathbf{Y}}, \mathcal{S}^T, \gamma^T, \boldsymbol{\mathcal{V}}, \hat{\mathbb{B}}^T\right) \mathcal{P}\left(\boldsymbol{f}_{j-1}^{\blacklozenge,T} \middle| \boldsymbol{f}_j^{\blacklozenge,T}\right)}{\mathcal{P}\left(\boldsymbol{f}_{j-1}^T \middle| \tilde{\mathbf{Y}}, \mathcal{S}^T, \gamma^T, \boldsymbol{\mathcal{V}}, \hat{\mathbb{B}}^T\right) \mathcal{P}\left(\boldsymbol{f}^{\blacklozenge,\boldsymbol{T}}{}_j \middle| \boldsymbol{f}^{\blacklozenge,\boldsymbol{T}}{}_{j-1}\right)}$, where $\mathcal{P}\left(\boldsymbol{f}^T \middle| \tilde{\mathbf{Y}}, \mathcal{S}^T, \gamma^T, \boldsymbol{\mathcal{V}}, \hat{\mathbb{B}}^T\right) = \mathcal{H}\left(\tilde{\mathbf{Y}} \middle| \boldsymbol{f}^T, \mathcal{S}^T, \gamma^T, \boldsymbol{\mathcal{V}}, \hat{\mathbb{B}}^T\right) \mathcal{P}\left(\boldsymbol{f}^T\right) \mathcal{I}_f$ is the posterior of $\boldsymbol{f}^T$, $\mathcal{P}\left(\boldsymbol{f}^T\right)$ is the prior of $\boldsymbol{f}^T$, and $\mathcal{I}_f$ is an indicator function flagging whether $\mathbf{A}_0$ is full rank, and
  5. the update is $\boldsymbol{f}_j^T = \boldsymbol{f}^{\blacklozenge,T}$ if $\kappa > u_1 \sim \mathcal{U}(0, 1) \Longrightarrow$ otherwise $\boldsymbol{f}_j^T = \boldsymbol{f}_{j-1}^T$.

Canova and Pérez Forero (2015): A Sampling Algorithm for the TVP-SV-SBVAR, cont.

6. Sample s_m^T $(m = 1, \ldots, \mathcal{K})$ from the mixture normal distribution using $\Pr\left(s_{\ell,t} = m \,\middle|\, \hat{\hat{y}}_t, \ln \gamma_{\ell,t}\right) \propto q_m \times \phi\left(\dfrac{\hat{\hat{y}}_t - 2\ln \gamma_{\ell,t} - \tau_m + 1.2704}{\varsigma_m}\right)$, conditional on $\hat{\hat{\mathbf{Y}}}$, $\boldsymbol{f}_j^T$, $\mathbb{B}_j^T$, and γ_{j-1}^T, where $\Pr(s_{\ell,t} = m) = q_m$ is the unconditional probability of state m, $\phi(\cdot)$ is the pdf of the standard normal distribution, (τ_m, ς_m) = (mean, standard deviation) of state m of the mixture normal distribution, and 1.2704 is the mean of $\ln \chi^2(1)$, which centers the numerator on zero to control realizations of $2\ln \gamma_{\ell,t}$.
7. Next, if $\Pr\left(s_{\ell,t} = j-1 \,\middle|\, \hat{\hat{y}}_t, \ln \gamma_{\ell,t}\right) < u_2 \leq \Pr\left(s_{\ell,t} = j \,\middle|\, \hat{\hat{y}}_t, \ln \gamma_{\ell,t}\right)$, $s_{\ell,t} = j$, otherwise, $s_{\ell,t} = j-1$, where $u_2 \sim \mathcal{U}(0, 1)$.
8. Similar to the Kalman smoothing updating of $\boldsymbol{f}_j^T$, generate updates of γ^T using its random walk law of motion conditional on $\hat{\hat{\mathbf{Y}}}$, $\boldsymbol{f}_j^T$, $\mathbb{B}_j^T$, and $\mathbb{S}_j^T$, where sampling of $\gamma_{\ell,t}$ depends on a diagonal $\boldsymbol{\Omega}_\xi \Longrightarrow$ independence of $\gamma_{\ell,t}$, $\ell = 1, \ldots, n$.
9. Sample the blocks of $\boldsymbol{\mathcal{V}}$ as independent $\mathcal{IW}$s, where $\boldsymbol{\mathcal{V}}_j \sim \mathcal{P}\left(\boldsymbol{\mathcal{V}}_j \,\middle|\, \tilde{\mathbf{Y}}, \boldsymbol{f}_j^T, \mathbb{S}_j^T, \gamma_j^T, \mathbb{B}_j^T\right)$.
10. Repeat these 9-steps in the Metropolis in Gibbs sampling algorithm $j = 1, \ldots, \mathcal{J}$ times.

Canova and Pérez Forero (2015): More on Estimating TVP-SV-SBVARs

- There are several MatLab™ programs provided by Canova and Pérez Forero (2015) to estimate non-recursive TVP-SV-SBVARs; see `http://www.qeconomics.org/upcoming/305/QE305_code_and_data.zip`.

- These MatLab™ programs include several algorithms for sampling $\mathbb{B}_t$, given the TVP-SV-SBVAR is identified on the short-run impact matrix $\mathcal{A}_0$.
  1. A multi-move MH sampler is described in Carter and Kohn (1994).
  2. A single-move MH sampler is proposed by Koop and Potter (2011, "Time varying VARs with inequality restrictions," *Journal of Economic Dynamics & Control* 35, 1126–1138).
  3. By recasting the TVP-SV-SVAR as a hierarchical model, Chib and Greenberg (1995, "Hierarchical analysis of SUR models with extensions to correlated serial errors and time-varying parameter models," *Journal of Econometrics* 68, 339–360) develop a multi-move MH sampler, which is refined by Koop and Korobilis (2010).

- Trade speed for efficiency when choosing single- or multi-move sampling algorithms.
  1. Multi-move samplers converge faster, but to satisfy the restriction the roots of $\mathbb{B}^T$ are outside the unit circle need many draws to achieve convergence.
  2. Single move algorithms require fewer draws, but convergence is slow.

Canova and Pérez Forero (2015): Sampling the Reduced Form TVPs

- A multi-move sampler draws many random variables at once $\Longrightarrow$ draw $\mathbb{B}^T$ in step 1 of the Canova-Pérez Forero algorithm $\Longrightarrow$ check the stationarity of $\mathbb{B}^T$.
- Koop and Potter (2011) sample $\mathbb{B}_t$ instead of $\mathbb{B}^T$ in their single-move sampler.
  1. $\Longrightarrow$ Draw $\mathbb{B}_t$, check if it is stationary, and then accept or reject it.
  2. This procedure will alter the process for sampling $\boldsymbol{\mathcal{V}}$.
- The hierarchical model of Chib and Greenberg (1995) alters the law of motion of the reduced-form intercept and slope parameters of the TVP-SV-SVAR to
  1. $\mathbb{B}_t = \boldsymbol{\Phi}\Lambda_t + \varrho_t$ and $\Lambda_t = \Lambda_{t-1} + \varpi_t$, where $\mathrm{Var}(\varrho_t) = \boldsymbol{\Omega}_\varrho$ and $\mathrm{Var}(\varpi_t) = \boldsymbol{\Omega}_\varpi$.
  2. Koop and Korobilis (2010) identify $\boldsymbol{\Phi}$ by assuming its upper block, $\boldsymbol{\Phi}_1 = \mathbf{I}_{np_1}$, $\Longrightarrow$ the lower block, $\boldsymbol{\Phi}_2$, is $np_2 \times np_2$, where $np_2 = n^2p + n - np_1 > 0$.

Canova and Pérez Forero (2015): More on Sampling the Reduced Form TVPs, III

- Under the hierarchical model, Canova and Pérez Forero develop a single-move sampler for $\mathbb{B}_t$, but Λ^T is drawn using a multi-move step.
  1. Given f_t, y_t, Φ, Λ_t, and $\mathbf{\Omega}_\varrho$, $\mathbb{B}_t \sim \mathcal{N}\left(\overline{\mathbb{B}}_t, \overline{\mathbf{\Omega}}_{\mathbb{B},t}\right)$, where $\overline{\mathbf{\Omega}}_{\mathbb{B},t} = \left(\underline{\mathbf{\Omega}}_{\mathbb{B}}^{-1} + \boldsymbol{\mathcal{X}}_t \mathbf{\Omega}_{\varepsilon,t}^{-1} \boldsymbol{\mathcal{X}}_t'\right)^{-1}$, $\overline{\mathbb{B}}_t = \overline{\mathbf{\Omega}}_{\mathbb{B},t}\left(\underline{\mathbf{\Omega}}_{\mathbb{B}}^{-1}\underline{\mathbb{B}}_t + \boldsymbol{\mathcal{X}}_t \mathbf{\Omega}_{\varepsilon,t}^{-1} \mathcal{Y}_t\right)^{-1}$, $\underline{\mathbb{B}}_t = \Phi\Lambda_t$, and $\underline{\mathbf{\Omega}}_{\mathbb{B}} = \mathbf{\Omega}_\varrho$.
  2. Sample $\mathbf{\Omega}_\varrho$ from the residuals $\mathbb{B}_t - \mathbf{\Phi}\Lambda_t$, which are distributed $\mathcal{IW}$.
  3. Draw $\Lambda_t \,|\, \mathbb{B}_t \Longrightarrow$ treat $\mathbb{B}_t$ as observable and cast the hierarchical model as a state space model to generate proposals for $\mathbb{B}_t$ and Λ_t, where $\mathbf{\Omega}_\varpi \sim \mathcal{IW}$, given Λ_t.
  4. Given $\mathbb{B}_t$, Λ_t, and $\mathbf{\Omega}_\varpi$, $\mathbf{\Psi}_2 \sim \mathcal{N}\left(\overline{\mathbf{\Psi}}_2, \overline{\mathbf{\Omega}}_{\mathbf{\Psi}_2}\right)$, where $\overline{\mathbf{\Omega}}_{\mathbf{\Psi}_2} = \left(\underline{\mathbf{\Omega}}_{\mathbf{\Psi}_2}^{-1} + \mathbf{\Omega}_{\varpi,2}^{-1}\Lambda^T\Lambda^{T\prime}\right)^{-1}$, $\overline{\mathbf{\Psi}}_2 = \overline{\mathbf{\Omega}}_{\mathbf{\Psi}_2}\left(\underline{\mathbf{\Omega}}_{\mathbf{\Psi}_2}^{-1}\underline{\mathbf{\Psi}}_2 + \mathbf{\Omega}_{\varpi,2}^{-1}\Lambda^T\mathbb{B}^T\right)^{-1}$, where Λ^T is $np_2 \times T$, $\mathbf{\Omega}_{\mathbf{\Psi}_2}$ is the associated covariance matrix, $\underline{\mathbf{\Psi}}_2 = \mathbf{0}_{np_2\times 1}$ and $\underline{\mathbf{\Omega}}_{\mathbf{\Psi}_2} = \mathrm{r}_\Psi^2 \mathbf{I}_{np_2}$, and $\mathrm{r}_\Psi^2 = 0.01$.
  5. Finally, as previously described, sample f_t, y_t, and $\boldsymbol{\mathcal{V}}$, conditional on $\mathbb{B}_t$ and $\underline{\mathbf{\Omega}}_{\mathbf{\Psi}}$.

Recent Literature on TVP-SV-SVARs

- Several parts of the current literature on SVARs are worth mentioning.
- Petrova (2019) develops a class of Gibbs samplers to estimate TVPs and SV nonparametrically.
  1. No parametric assumptions imposed on TVPs and SV (*i.e.*, random walks).
  2. Trade loss of efficiency of nonparametric methods against misspecification error if, say, random walk assumption is incorrect.
- Nonparametric methods almost always rely of kernel estimators.
  1. Estimate log likelihood of a TVP-SV-VAR as a weighted average conditional on draws of the TVPs and SVs.
  2. The weights are calculated using a time-varying normal kernel estimator.
  3. $\Longrightarrow$ Kernel smooths the log likelihood over a time-varying subsample or a rolling window with changing endpoints.
  4. Choice of the kernel is a (hidden) parametric assumption that is not innocuous.
  5. Silverman (1986, Density Estimation for Statistics and Data Analysis, New York, NY: Chapman and Hall) and Pagan and Ullah (1999, Nonparametric Econometrics, Cambridge, U.K.: Cambridge University Press).
- Empirical question is impact on estimator of posterior log likelihood
  1. of misspecification error of random walk of TVPs and SV and
  2. relative inefficiency of rolling window kernel to produce.

Recent Literature on TVP-SV-SVARs

- Research proposing estimators and tests of SVARs with SV has been ongoing for more than ten years.

- Begin with Rigobon (2003, "Identification through heteroskedasticity," *Review of Economics and Statistics* 85, 777–792) and an example is Lanne and Lütkephol (2008).
  1. Rigobon shows specifying enough breaks or change points gives identifying restrictions to estimate structural shocks with time-varying heteroskedasticity.
  2. Lanne and Lütkephol apply this idea to monetary policy SVARs to study the identification of monetary policy shocks.

- Lütkephol and Netšunajev (2017, "Structural vector autoregressions with heteroskedasticity: A review of different volatility models," *Econometrics and Statistics* 1, 2–18) is a survey and current examples are
  1. Lütkephol and Netšunajev (2017, "Structural vector autoregressions with smooth transition in variances: The interaction between US monetary policy and the stock market," *Journal of Economic Dynamics and Control* 84, 43–57),
  2. Lewis (2021) for an example applying nonparametric methods to estimate time-varying heteroskedastic SVARs applied to fiscal policy shocks,
  3. Brunnermeier, et al (2021) estimate a fixed coefficient SVAR with MS-SV identified by forcing the cross-section structural variances to sum to unity $\Longrightarrow$ forces at least one variance to change at each date t,
  4. Lütkephol, Meitz, Netšunajev, and Saikkonen (2021, "Testing identification via heteroskedasticity in structural vector autoregressive models," *Econometrics Journal* 24, 1–22) for recent inference methods.