var ivsvar — Instrumental-variables structural vector autoregressive models

Description	Quick start	Menu	Syntax
Options	Remarks and examples	Stored results	Methods and formulas
References	Also see		

Description

`ivsvar` estimates the parameters of structural vector autoregressive (SVAR) models by instrumental variables. Instrumental-variables SVAR models are an alternative to the short-run SVAR models fit by `svar`, requiring fewer constraints than would be necessary in those models to identify the structural impulse–response functions (IRFs). They need fewer constraints because the shocks of interest, target shocks, are modeled using instrumental variables. The structural IRFs are then estimated only for the target shocks. Instrumental-variables SVAR models are also called proxy SVAR models. `ivsvar` provides two estimators: a generalized method of moments (GMM) estimator for a single target shock and a minimum distance estimator for multiple target shocks.

Quick start

Fit an instrumental-variables SVAR model for the variables y1, y2, and y3, with y3 instrumented by z using the GMM estimator

```
ivsvar gmm y1 y2 (y3 = z)
```

Same as above, but run the reduced-form vector autoregressive (VAR) model with lags 1 through 4 instead of the default 1 through 2

```
ivsvar gmm y1 y2 (y3 = z), lags(1/4)
```

Add exogenous variables x1 and x2

```
ivsvar gmm y1 y2 (y3 = z), lags(1/4) exog(x1 x2)
```

Instrumental-variables SVAR with the minimum distance estimator, using short-run constraints given by predefined matrix P

```
ivsvar mdist y1 (y2 y3 = z1 z2), peq(P)
```

Menu

Statistics > Multivariate time series > Instrumental-variables SVAR

Syntax

Generalized method of moments (GMM) estimator

> ivsvar gmm *depvarlist* (*varname*_{target} = *varlist*_{iv}) [*if*] [*in*] [, *options gmm_options*]

Minimum distance estimator

> ivsvar mdist *depvarlist* (*varlist*_{target} = *varlist*_{iv}) [*if*] [*in*] [, *options mdist_options*]

*varname*_{target} is a dependent variable corresponding to the target shock.

*varlist*_{iv} is a list of instruments.

*varlist*_{target} is a list of dependent variables corresponding to the target shocks.

options	Description
Model	
noconstant	suppress constant term
nozconstant	suppress constant terms for the instruments **z**
lags(*numlist*)	specify a list of lags for the VAR model; default is lags(1 2)
exog(*varlist*_{exog})	specify exogenous variables
zlags(*numlist*)	specify a list of lags for the instruments **z**; default is no lags for the instruments
Reporting	
level(*#*)	set confidence level; default is level(95)
nocnsreport	do not display constraints
display_options	control columns and column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling
coeflegend	display legend instead of statistics

gmm_options	Description
Model 2	
scale(*#*)	set scale for impact effect; default is scale(1)
showgmm	display underlying GMM output
norescale	do not rescale GMM output
SE/Robust	
vce(*vcetype*)	*vcetype* may be robust, cluster *clustvar*, or hac *hacspec*

mdist_options	Description
Model 2	
beq(*matrix*$_{beq}$)	define and apply to **B** equality constraint matrix *matrix*$_{beq}$
peq(*matrix*$_{peq}$)	define and apply to **P**$_z$ equality constraint matrix *matrix*$_{peq}$
qmatrix(*matrix*)	specify minimum distance weight matrix **Q**
showvar	display underlying VAR output
showzvar	display underlying VAR output for the instruments
Maximization	
maximize_options	control the maximization process; seldom used

You must tsset your data before using ivsvar; see [TS] **tsset**.

collect, rolling, and statsby are allowed; see [U] **11.1.10 Prefix commands**.

coeflegend does not appear in the dialog box.

See [U] **20 Estimation and postestimation commands** for more capabilities of estimation commands.

Options

> Model

noconstant suppresses the constant term in the reduced-form VAR model.

nozconstant suppresses the constant terms in preliminary regressions on the instruments **z**.

lags(*numlist*) specifies the lags to be included in the reduced-form VAR model. The default is lags(1 2). This option takes a *numlist* and not simply an integer for the maximum lag. For instance, lags(2) would include only the second lag in the model, whereas lags(1/2) would include both the first and second lags in the model. See [U] **11.1.8 numlist** and [U] **11.4.4 Time-series varlists** for further discussion of *numlist*s and lags.

exog(*varlist*$_{exog}$) specifies a list of exogenous variables to be included in the reduced-form VAR model.

zlags(*numlist*) specifies that the lags in *numlist* be partialed out when running preliminary regressions on the instruments **z**. The residuals from these regressions are then used as the instruments. See [U] **11.1.8 numlist** and [U] **11.4.4 Time-series varlists** for further discussion of *numlist*s and lags.

> Model 2

The following options are specific to the GMM estimator:

scale(#) specifies the scaling factor used in computing impact effects. For example, an impact effect of 0.25 may be computed with scale(0.25). The default is scale(1).

showgmm specifies that the output from gmm also be displayed. By default, it is fit quietly.

norescale removes rescaling entirely. The coefficients estimated by GMM are reported.

The following options are specific to the minimum distance estimator:

beq(*matrix*$_{beq}$) and peq(*matrix*$_{peq}$) specify the short-run constraints in an instrumental-variables SVAR model. Short-run constraints are required any time there is more than one target shock. The beq() option specifies constraints on the parameters of the **B** matrix; the peq() option specifies constraints on the parameters of the **P**$_z$ matrix (see *Multiple target shocks* for more details on the **B** and **P**$_z$ matrices). An instrumental-variables SVAR model requires at least $r(r-1)/2$ constraints, where r is the number of target shocks.

beq($matrix_{beq}$) specifies a matrix that defines a set of equality constraints. The **B** matrix must be of dimension $k \times r$, where k is the number of variables in the VAR model and r is the number of target shocks. The elements of this matrix must be missing or real numbers. A missing value in the (i, j) element of this matrix specifies that the (i, j) element of **B** is a free parameter. A real number in the (i, j) element of this matrix constrains the (i, j) element of **B** to this real number. If beq($matrix_{beq}$) is not specified, all elements in **B** are assumed to be free parameters.

peq($matrix_{peq}$) specifies a matrix that defines a set of equality constraints. The $\mathbf{P_z}$ matrix must be of dimension $r \times r$. It follows the same rules as described in beq($matrix_{beq}$), except that it applies to $\mathbf{P_z}$ rather than **B**. If peq($matrix_{peq}$) is not specified, all elements in $\mathbf{P_z}$ are assumed to be free parameters. For example, if there are two instruments for two target shocks, the matrix

$$\mathbf{P_z} = \begin{bmatrix} \cdot & 0 \\ 0 & \cdot \end{bmatrix}$$

specifies that the first instrument is correlated with the first target shock, the second instrument is correlated with the second target shock, and the cross-correlations are constrained to be zero. By contrast, the matrix

$$\mathbf{P_z} = \begin{bmatrix} \cdot & 0 \\ \cdot & \cdot \end{bmatrix}$$

continues to impose that the first instrument is not affected by the second target shock, but allows the second instrument to be correlated with both target shocks.

qmatrix($matrix$) specifies the weight matrix **Q** used in minimum distance estimation. By default, the inverse variance matrix of the reduced-form parameters is used.

showvar specifies that the output from the underlying VAR model also be displayed.

showzvar specifies that the output from the underlying VAR model for the instruments **z** also be displayed.

SE/Robust

vce($vcetype$) specifies the type of standard error reported, which includes types that are robust to some kinds of misspecification (robust) and that allow for intragroup correlation (cluster $clustvar$); see [R] **vce_option**.

vce(hac $hacspec$) requests a heteroskedasticity- and autocorrelation-consistent (HAC) variance–covariance matrix. The full syntax of $hacspec$ is one of the following:

vce(hac $kernel$ [#]) requests a HAC variance–covariance matrix using the specified kernel (see below) with optional # lags. The bandwidth of a kernel is equal to $\# + 1$. If # is not specified, a kernel with $N - 2$ lags is used, where N is the sample size.

vce(hac $kernel$ opt [#]) requests a HAC variance–covariance matrix using the specified kernel (see below), and the lag order is selected using Newey and West's (1994) optimal lag-selection algorithm. # is an optional tuning parameter that affects the lag order selected; see the discussion in *Methods and formulas* in [R] **ivregress**.

$kernel$ may be one of the following:

bartlett or nwest requests the Bartlett (Newey–West) kernel.

parzen or gallant requests the Parzen (Gallant 1987) kernel.

quadraticspectral or andrews requests the quadratic spectral (Andrews 1991) kernel.

level(*#*); see [R] **Estimation options**.

nocnsreport; see [R] **Estimation options**.

display_options: noci, nopvalues, cformat(*%fmt*), pformat(*%fmt*), and sformat(*%fmt*); see [R] **Estimation options**.

Maximization

maximize_options: difficult, technique(*algorithm_spec*), iterate(*#*), [no]log, trace, gradient, showstep, hessian, tolerance(*#*), ltolerance(*#*), nrtolerance(*#*), nonrtolerance, and from(*init_specs*); see [R] **Maximize**. These options are seldom used.

The following option is available with ivsvar but is not shown in the dialog box:

coeflegend; see [R] **Estimation options**.

Remarks and examples

Remarks are presented under the following headings:

Introduction
Structural VAR models with external instruments
Multiple target shocks

Introduction

This entry assumes that you are familiar with VAR model and SVAR model estimation. If you are not, please look at [TS] **var intro**, [TS] **var**, [TS] **var svar**, and the references therein. Here, we illustrate how to fit SVAR models with instrumental variables, also known as proxies. This method is an alternative to that described in [TS] **var svar**, which fits SVAR models subject to short-run and long-run restrictions. For more detailed information on SVAR models, see Amisano and Giannini (1997) and Hamilton (1994). For good introductions to VAR models, see Lütkepohl (2005), Hamilton (1994), Stock and Watson (2001), and Becketti (2020).

Instrumental-variables SVAR models were introduced by Stock and Watson (2012), with early applications by Mertens and Ravn (2013) and Gertler and Karadi (2015). Montiel Olea, Stock, and Watson (2021) provide a treatment of the GMM estimator. The minimum distance estimator for instrumental-variables SVAR models is due to Angelini and Fanelli (2019).

Structural VAR models with external instruments

A reduced-form VAR model without exogenous variables can be written as

$$\mathbf{y}_t = \mathbf{A}_1\mathbf{y}_{t-1} + \cdots + \mathbf{A}_p\mathbf{y}_{t-p} + \mathbf{u}_t$$

where $\mathbf{y}_t$ is a $k \times 1$ vector of endogenous variables, $(\mathbf{A}_1, \ldots, \mathbf{A}_p)$ are $k \times k$ matrices of parameters, and $\mathbf{u}_t$ is a $k \times 1$ vector of residuals. The residuals have $k \times k$ covariance matrix $\boldsymbol{\Sigma}$. As in [TS] **var svar**, we assume the reduced-form residuals $\mathbf{u}_t$ can be written as linear combinations of underlying independent shocks $\mathbf{e}_t$ by

$$\mathbf{u}_t = \mathbf{B}\mathbf{e}_t$$

where $\mathbf{B}$ is a $k \times k$ matrix and $\mathbf{e}_t$ is a $k \times 1$ vector of shocks. The matrix $\mathbf{B}$ is related to the reduced-form residual covariance matrix by

$$\boldsymbol{\Sigma} = \mathbf{B}\mathbf{B}' \tag{1}$$

where $E(\mathbf{e}_t\mathbf{e}_t') = \mathbf{I}$.

There are k^2 entries in $\mathbf{B}$, but the data provide only $k(k-1)/2$ pieces of information in the symmetric matrix $\boldsymbol{\Sigma}$. An SVAR model places restrictions on some entries of $\mathbf{B}$ to estimate the remaining entries. For example, a researcher might set $\mathbf{B}$ to be lower triangular, leaving the $k(k-1)/2$ diagonal and lower-diagonal elements to be estimated.

Instrumental-variables SVAR models estimate the same parameters as traditional SVAR models; the object of interest is one or more columns of $\mathbf{B}$. Instrumental-variables SVAR models impose fewer restrictions on $\mathbf{B}$ than would be required in an SVAR model. This is accomplished by incorporating additional information that supplements the covariance restriction in (1). In an instrumental-variables SVAR model, there exists a variable z_t with the following properties:

$$E(z_t e_{1,t}) = p_z$$
$$E(z_t e_{j,t}) = 0 \quad \forall j \neq 1$$

We call $e_{1,t}$ the target shock. The variable z_t is correlated with the shock $e_{1,t}$ and is uncorrelated with all other shocks. We say z_t is an instrument, and it can be used to recover the first column of $\mathbf{B}$.

One can write

$$
\begin{aligned}
E(\mathbf{u}_t z_t) &= E(\mathbf{B}\mathbf{e}_t z_t) \\
&= \mathbf{B}_1 E(z_t e_{1,t}) + \mathbf{B}_2 E(z_t e_{2,t}) + \cdots + \mathbf{B}_k E(z_t e_{k,t}) \\
&= \mathbf{B}_1 p_z
\end{aligned}
$$

so that the column of $\mathbf{B}$ associated with the target shock $e_{1,t}$ is identified, up to scale factor p_z. The instrumental-variables SVAR model does not attempt to identify the remaining columns of $\mathbf{B}$, which correspond to the nontarget shocks.

The instrumental-variables SVAR model identifies the first column of $\mathbf{B}$ up to a scale factor. The final normalization is the unit effect normalization: for a specified target variable, the effect on impact is normalized to 1.

Instrumental-variables SVAR models come with benefits and costs. The benefit is that the researcher needs to impose fewer restrictions on the impact effect matrix $\mathbf{B}$ than would be necessary in an SVAR model identified by short-run restrictions. The cost is that the instrumental-variables SVAR model estimates only the columns of $\mathbf{B}$ related to the target shocks; it can say nothing about the columns of $\mathbf{B}$ related to the nontarget shocks.

▷ Example 1: Instrumental-variables SVAR model

In this example, we use common US macroeconomic data. We have data on US industrial production growth (`ip_growth`), consumer price index inflation (`inflation`), and the interest rate (`fedfunds`). These three variables are the outcome variables of interest. We also have two instruments. The first is a proxy for monetary policy shocks (`money_inst`), the surprise component of interest rate movements used in Romer and Romer (2004) and Wieland and Yang (2020). The second is a proxy variable for oil shocks (`oil_inst`), the net oil price increase over the previous 12 months, similar to a measure constructed in Hamilton (2003). These two proxies are used to instrument for interest rate shocks and inflation shocks, respectively.

We first fit an instrumental-variables SVAR model on inflation, the interest rate, and industrial production growth, using the oil price instrument as a proxy for the inflation shock.

```
. use https://www.stata-press.com/data/r19/usmacro3
(Federal Reserve Economic Data - St. Louis Fed, 2023-09-01)

. ivsvar gmm fedfunds ip_growth (inflation = oil_inst)

Step 1:
Iteration 0:   GMM criterion = .74206787
Iteration 1:   GMM criterion = 2.677e-31
Iteration 2:   GMM criterion = 1.678e-31

Step 2:
Iteration 0:   GMM criterion = 4.547e-31
Iteration 1:   GMM criterion = 4.547e-31  (backed up)

note: model is exactly identified.

Instrumental-variables SVAR                           Number of obs = 783
VAR sample: 1954m10 thru 2019m12
GMM sample: 1954m10 thru 2019m12
 ( 1)  [e.inflation]inflation = 1
```

Effect	Coefficient	Robust std. err.	z	P>|z|	[95% conf. interval]	
e.inflation						
fedfunds	.0046142	.271441	0.02	0.986	−.5274004	.5366288
ip_growth	−.31198	.4330713	−0.72	0.471	−1.160784	.5368241
inflation	1	(constrained)				

```
Note: Underlying VAR fit with 2 lags.
Dependent variables: fedfunds ip_growth inflation
Instrumented shock: inflation
Instrument: oil_inst
```

The output displays only the impact coefficients; the reduced-form VAR lag coefficients are suppressed. The impact effects are the elements of the column of **B** that correspond to the structural inflation shock. The impact effects are shown only for the target shock, in this case, the inflation shock (`e.inflation`). The first response is the impact effect of the inflation shock on the interest rate (`fedfunds`), which is close to zero. The second response is the impact effect of the inflation shock on the growth rate of industrial production (`ip_growth`); the estimated impact effect is −0.31, though the 95% confidence interval includes zero. The third response is the effect of an inflation shock on inflation itself (`inflation`), which is scaled to 1.

In a short-run SVAR model fit by the `svar` command, all elements of the impact matrix **B** are displayed, arranged by column. In the present `ivsvar` output, only columns of **B** associated with the target shock are estimated and displayed. Hence, the output above contains only the impact effects of the inflation shock.

To trace out the dynamic effects of the inflation shock, we create and graph the structural IRFs with the `irf` suite of commands:

```
. irf set ivsvarirf.irf, replace
. irf create model1, step(36)
. irf graph sirf, impulse(inflation)
```

`sirf` requests that structural IRFs be displayed.

Graphs by irfname, impulse variable, and response variable
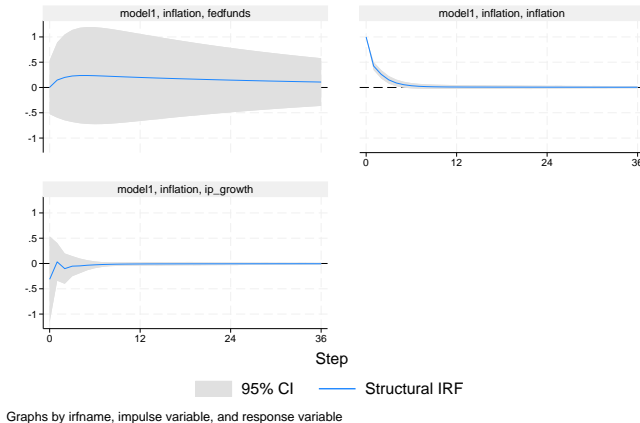
Each panel displays the dynamic response of one variable to the shock. The x axis measures time in the units of the data, in this case months. The y axis has the same units as the variables, in this case, percentage points. The interest rate (top left) shows little response to the shock at any time horizon. Inflation (top right) rises by 1% on impact by construction, falling back to its long-run average within 6 steps (6 months). Industrial production falls on impact but returns to its long-run average quickly.

◁

▷ Example 2: Including additional lags in the VAR model

We continue with the above example and now include 12 lags in the reduced-form VAR model instead of the default 2 lags. These additional lags can detect more complicated short-run dynamics in the periods after a shock. We can then graph the dynamic responses from both models on the same set of IRF graphs:

```
. ivsvar gmm fedfunds ip_growth (inflation = oil_inst), lags(1/12)
(output omitted)
. irf create model2, step(36)
irfname model2 not found in ivsvarirf.irf
(file ivsvarirf.irf updated)
. irf graph sirf, impulse(inflation)
```


Graphs by irfname, impulse variable, and response variable
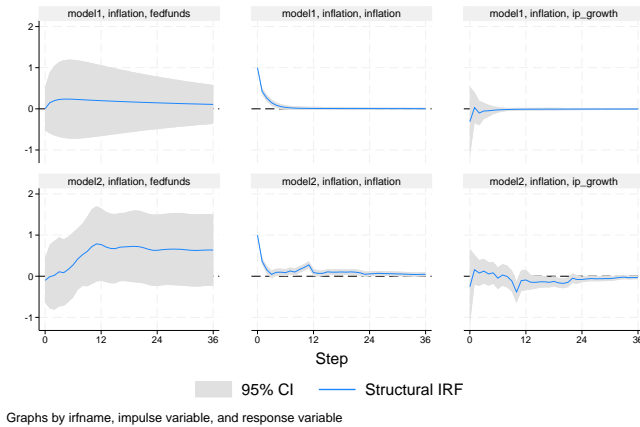
Each row contains the dynamic responses from one model. The top row reproduces the IRF graphs for the 2-lag model. The bottom row produces the IRF graphs for the 12-lag model. With these additional lags, the interest rate response is positive one year after the shock and does not return to steady state even 36 months after the shock. In the two-lag model shown on the top row, the inflation response quickly returns to zero, whereas in the many-lag model, there is additional persistence in the inflation response. Industrial production declines on impact as in the two-lag model, with additional declines in the months following the shock.

◁

Multiple target shocks

So far we have considered a VAR model with a single instrument and a single target shock. The `ivsvar mdist` estimator is available for models with multiple instruments and multiple target shocks. As before, there is a VAR model:

$$\mathbf{y}_t = \mathbf{A}_1\mathbf{y}_{t-1} + \cdots + \mathbf{A}_p\mathbf{y}_{t-p} + \mathbf{u}_t$$

As before, we assume the reduced-form residuals $\mathbf{u}_t$ can be written as linear combinations of underlying independent shocks $\mathbf{e}_t$ by

$$\mathbf{u}_t = \mathbf{B}\mathbf{e}_t$$

For convenience, we split the k shocks into two groups: r target shocks and $(k-r)$ nontarget shocks.

$$\mathbf{u}_t = \mathbf{B}_1\mathbf{e}_{1t} + \mathbf{B}_2\mathbf{e}_{2t}$$

There is an $r \times 1$ vector of instruments $\mathbf{z}_t$ that is related to the target shocks $\mathbf{e}_{1t}$ (the relevance assumption) and is unrelated to the nontarget shocks $\mathbf{e}_{2t}$ (the exclusion assumption), satisfying

$$\mathbf{z}_t = \mathbf{P_z}\mathbf{e}_{1t} + \boldsymbol{\omega}_t$$

where $\mathbf{P_z}$ is a $r \times r$ matrix describing the relationship of the instruments to the target shocks and $\boldsymbol{\omega}_t$ is noise.

▷ Example 3: Multiple target shocks

We fit an instrumental-variables SVAR model with two instruments for two target shocks. The instruments are the monetary instrument and the oil price instrument; the target shocks are the interest rate and inflation shocks. Because we have two instruments and two shocks, one restriction is required to estimate the remaining parameters. This could be a restriction on the impact coefficients $\mathbf{B}_1$ or a restriction on how the instruments are related to the target shocks. If each instrument is correlated only with one target shock, then the $\mathbf{P_z}$ matrix will be diagonal, giving us two restrictions. We can relax this assumption, allowing one of the two instruments to additionally be related to a second target shock. Setting up the matrix

```
. matrix P = (., 0 \ ., .)
```

specifies that the first instrument is related only to the first target shock and the second instrument is potentially related to both target shocks. Thus, order matters when specifying the instruments and target shocks.

```
. ivsvar mdist ip_growth (fedfunds inflation = money_inst oil_inst),
> lags(1/12) peq(P)

Estimating short-run parameters:
Iteration 0:  Distance criterion =   6.333e-32
Iteration 1:  Distance criterion =   5.009e-32

Refining estimates:
Iteration 0:  Distance criterion =   1.919e-33
Iteration 1:  Distance criterion =   1.906e-33

Instrumental-variables SVAR                         Number of obs = 468
Endogenous sample:  1955m8 thru 2019m12
Instrument sample:  1969m1 thru 2007m12

 ( 1)  [e.inflation]money_inst = 0
```

Effect	Coefficient	Std. err.	z	P>|z|	[95% conf. interval]	
e.fedfunds						
ip_growth	.162055	.0627292	2.58	0.010	.0391079	.285002
fedfunds	.4161643	.014041	29.64	0.000	.3886445	.4436841
inflation	.0319081	.0179481	1.78	0.075	-.0032696	.0670858
e.inflation						
ip_growth	-.0980082	.1484195	-0.66	0.509	-.3889051	.1928888
fedfunds	-.0313886	.0297882	-1.05	0.292	-.0897724	.0269952
inflation	.2118086	.0084356	25.11	0.000	.1952751	.2283421
e.fedfunds						
money_inst	.1626293	.0126147	12.89	0.000	.1379049	.1873537
oil_inst	.0470263	.2422618	0.19	0.846	-.4277982	.5218507
e.inflation						
money_inst	0	(constrained)				
oil_inst	1.138002	.2256836	5.04	0.000	.6956705	1.580334

```
Wald test of instrument relevance: chi2(6) = 218.1      Prob > chi2 = 0.000
Note: Underlying VAR fit with 12 lags.
Dependent variables: ip_growth fedfunds inflation
Instrumented shocks: fedfunds inflation
Instruments: money_inst oil_inst
```

Before we discuss the estimation output, notice from the header output that the endogenous variables are measured from 1955m8 and the instruments are measured from 1969m1. The endogenous sample and instrument sample may differ, which is useful when the instrument is measured on a shorter sample than the endogenous variables (or vice versa).

The estimation output is separated into blocks. Each block is the impact effect of one shock either on the endogenous variables or on the instruments. In the above output, the first two blocks are columns of the $\mathbf{B}_1$ matrix; the coefficients are the impact effects of the shocks on the endogenous variables. The last two blocks are the columns of the $\mathbf{P}_z$ matrix; the coefficients are the impact effects of the shocks on the instruments.

From the first block, we see that on impact, an interest rate shock raises the interest rate, inflation, and industrial production growth. From the second block, we find that an inflation shock increases inflation, reduces the interest rate, and reduces industrial production growth on impact. Importantly, no elements of $\mathbf{B}_1$ needed to be constrained, unlike in a short-run SVAR model.

The third block displays the effect of an interest rate shock on the two instruments. The interest rate shock is positively related to the monetary instrument. In addition, we have allowed the interest rate shock to be correlated with the oil price instrument. This coefficient of 0.047 indicates that when there is

an interest rate shock, the oil price instrument rises by 0.047. Allowing for these cross-effects—nonzero response of some instruments to the "wrong" shock—is a useful feature of the multiple-shock approach. The final block displays the effect of an inflation shock on the two instruments. In the final block, we have assumed no influence of the inflation shock on the monetary instrument and allowed the inflation shock to influence the oil price instrument.

We can compute impulse–response functions for each instrumented shock.

```
. irf create model3, step(36)
```

Responses to an interest rate shock are

```
. irf graph sirf, irf(model3) impulse(fedfunds) response(inflation ip_growth)
```
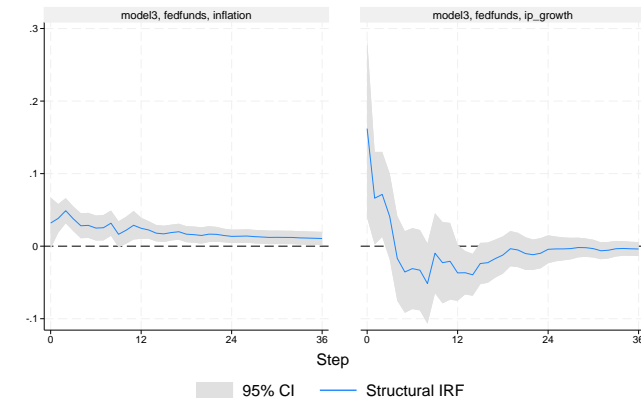

Graphs by irfname, impulse variable, and response variable

The inflation response is slightly positive for all 36 periods. Industrial production growth rises slightly initially but then falls within six months of the shock.

Responses to an inflation shock are

```
. irf graph sirf, irf(model3) impulse(inflation) response(inflation ip_growth)
```
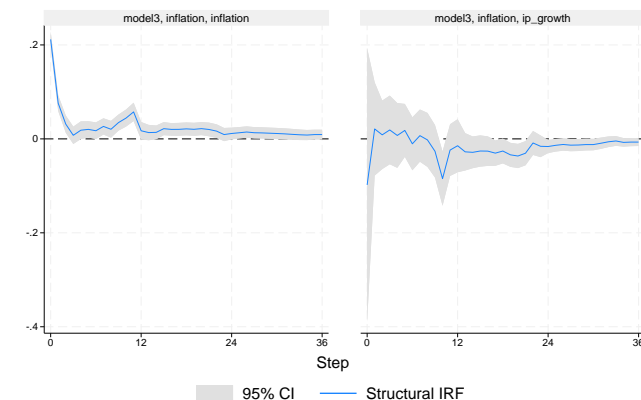

Graphs by irfname, impulse variable, and response variable

Responses here are similar to what was estimated in the `ivsvar gmm` case.

Stored results

ivsvar stores the following in e():

Scalars
e(N)	number of observations
e(neqs)	number of equations
e(k_eq)	number of equations in e(b)
e(k_dv)	number of dependent variables
e(k_exo)	number of exogenous variables
e(k_eq_var)	number of equations in underlying VAR models
e(dist)	distance
e(tmin_var)	minimum time for VAR equations
e(tmax_var)	maximum time for VAR equations
e(tmin_inst)	minimum time for instrument equations
e(tmax_inst)	maximum time for instrument equations
e(mlag_var)	highest lag in VAR
e(N_clust)	number of clusters
e(rank)	rank of e(V)
e(converged)	1 if converged, 0 otherwise

Macros
e(cmd)	ivsvar
e(cmdline)	command as typed
e(estimator)	gmm or mdist
e(tmins_var)	formatted minimum time for VAR equations
e(tmaxs_var)	formatted maximum time for VAR equations
e(tmins_inst)	formatted minimum time for instrument equations
e(tmaxs_inst)	formatted maximum time for instrument equations
e(clustvar)	name of cluster variable
e(vce)	*vcetype* specified in vce()
e(vcetype)	title used to label Std. err.
e(lags_var)	lags in model
e(endog_var)	names of endogenous variables
e(exog_var)	names of exogenous variables, if specified
e(instruments)	names of instruments
e(nocons_var)	nocons, if noconstant specified
e(eqnames_var)	names of equations
e(tsfmt)	format for the current time variable
e(timevar)	name of time variable
e(title)	title in estimation output
e(properties)	b V
e(predict)	program used to implement predict
e(marginsnotok)	predictions disallowed by margins

Matrices
e(b)	coefficient vector
e(Cns)	constraints matrix
e(Sigma)	$\widehat{\Sigma}$ matrix
e(V)	variance–covariance matrix of the estimators
e(Cholesky)	Cholesky factor of Sigma matrix
e(B1)	estimated $\mathbf{B}_1$ matrix
e(P)	estimated $\mathbf{P}$ matrix, if the minimum distance estimator was used

Functions
e(sample)	marks estimation sample

In addition to the above, the following is stored in `r()`:

Matrices
    `r(table)`   matrix containing the coefficients with their standard errors, test statistics, p-values, and confidence intervals

Note that results stored in `r()` are updated when the command is replayed and will be replaced when any r-class command is run after the estimation command.

Methods and formulas

Methods and formulas are presented under the following headings:

> *GMM*
> *Minimum distance*

GMM

A VAR(p) model without exogenous variables can be written

$$\mathbf{y}_t = \mathbf{A}_1 \mathbf{y}_{t-1} + \cdots + \mathbf{A}_p \mathbf{y}_{t-p} + \mathbf{u}_t$$

where $\mathbf{y}_t$ is a $k \times 1$ vector of endogenous variables, each of $(\mathbf{A}_1, \ldots, \mathbf{A}_p)$ is a $k \times k$ matrix of coefficients, and $\mathbf{u}_t$ is a $k \times 1$ vector of disturbances. These are referred to as the VAR equations. The VAR disturbances are related to the underlying shocks via

$$\mathbf{u}_t = \mathbf{B}\mathbf{e}_t$$

where $\mathbf{B}$ is a $k \times k$ matrix of impact effects and $\mathbf{e}_t$ is a $k \times 1$ vector of shocks. There is one target shock whose effects we are interested in, and we wish to identify the column of $\mathbf{B}$ associated with the target shock. Without loss of generality, let the target shock be $e_{1,t}$.

There is an instrument z_t with the relevance and exclusion restrictions

$$E(z_t e_{1t}) = p_z$$
$$E(z_t e_{jt}) = 0 \quad \forall j \neq 1$$

These conditions together imply the moment condition

$$E(\mathbf{u}_t z_t) = \mathbf{B}_1 p_z$$

that is, it identifies the column of $\mathbf{B}$ associated with the target shock $e_{1,t}$ up to scale p_z.

It is possible to partial out a constant term and lags of the instrument, leading to the instrument equation,

$$z_t = \gamma_0 + \gamma_1 z_{t-1} + \cdots + \gamma_l z_{t-l} + v_t$$

where the γ_i are coefficients. In this case, the moment condition relating the instruments and the VAR residuals is stated in terms of the instrument residuals,

$$E(\mathbf{u}_t v_t) = \mathbf{B}_1 p_z$$

`ivsvar gmm` estimates the VAR equations, any instrument equations, and the instrument moment conditions jointly. The moment condition is then rescaled so that the impact effect of the target shock is 1 for a prespecified variable.

Minimum distance

Same as above, let $\mathbf{y}_t$ be a $k \times 1$ collection of endogenous variables with the VAR(p) representation

$$\mathbf{y}_t = \mathbf{A}_1 \mathbf{y}_{t-1} + \cdots + \mathbf{A}_p \mathbf{y}_{t-p} + \mathbf{u}_t$$

The VAR residuals continue to be related to the underlying shocks via

$$\mathbf{u}_t = \mathbf{B}\mathbf{e}_t$$

where $\mathbf{B}$ is a $k \times k$ matrix and $\mathbf{e}_t$ is a $k \times 1$ vector of shocks. We partition the shocks into $(\mathbf{e}_{1,t}, \mathbf{e}_{2,t})$,

$$\mathbf{u}_t = \mathbf{B}_1 \mathbf{e}_{1,t} + \mathbf{B}_2 \mathbf{e}_{2,t}$$

where $\mathbf{e}_{1,t}$ is a $g \times 1$ collection of target shocks and $\mathbf{e}_{2,t}$ is a $(k-g) \times 1$ collection of nontarget shocks. Hence, $\mathbf{B}_1$ is $k \times g$ and $\mathbf{B}_2$ is $k \times (k-g)$. The parameters of interest are $\mathbf{B}_1$, which are the columns of $\mathbf{B}$ corresponding to the target shocks.

We have an $r \times 1$ collection of instruments $\mathbf{z}_t$ associated with the target shocks $\mathbf{e}_{1,t}$. The instruments are related to the target shocks by an $r \times g$ matrix $\mathbf{P_z}$ and are not related to nontarget shocks. This relationship is parameterized by

$$\mathbf{z}_t = \mathbf{P_z}\mathbf{e}_{1t} + \omega_t$$

where ω_t is an $r \times 1$ vector of white-noise disturbance terms. Lags of $\mathbf{z}_t$ can be added to the right-hand side of this equation if desired. This equation is the analogue of the relevance and exclusion restrictions in the GMM case. The matrix $\mathbf{P_z}$ measures the strength of association between the instruments and the target shocks and is estimated.

The minimum distance estimator estimates the structural parameters $(\mathbf{B}_1, \mathbf{P_z})$. The estimator minimizes the distance between reduced-form covariances and covariances implied by the model.

A $\mathbf{B}$-type SVAR model implies a relationship between the covariance matrix of the VAR residuals, $\mathbf{\Sigma_u}$, and the impact matrix $\mathbf{B}$,

$$\mathbf{\Sigma_u} = \mathbf{B}\mathbf{B}'$$

This equation provides a mapping between the k^2 structural parameters in $\mathbf{B}$ and the $k(k+1)/2$ unique elements of $\mathbf{\Sigma_u}$. With $k(k-1)/2$ additional restrictions in place, the parameters in $\mathbf{B}$ can be estimated so that the mapping holds as closely as possible.

The instrumental-variables SVAR model implies an analogous set of mappings. Let $\mathbf{\Sigma_{zu}} = \mathbf{Z}'\mathbf{U}/T$ and let $\mathbf{\Sigma_{uz}} = \mathbf{\Sigma_{zu}}'$. Then the relationships are

$$\mathbf{\Sigma_{zu}} = \mathbf{P_z}\mathbf{B}_1' \tag{2}$$

and

$$\mathbf{\Sigma_{zu}}\mathbf{\Sigma_u}^{-1}\mathbf{\Sigma_{uz}} = \mathbf{P_z}\mathbf{P_z}' \tag{3}$$

The left-hand side of each expression (2) and (3) consists of reduced-form parameters that can be computed from the instruments $\mathbf{z}_t$ and the VAR residuals $\mathbf{u}_t$. The right-hand side of each expression (2) and (3) consists of structural parameters to be estimated. $\mathbf{\Sigma_{zu}}$ is an $r \times k$ matrix; $\mathbf{\Sigma_{zu}}\mathbf{\Sigma_u}^{-1}\mathbf{\Sigma_{uz}}$ is an $r \times r$ symmetric matrix.

Let

$$\boldsymbol{\theta} = \begin{bmatrix} \text{vec}(\mathbf{B}_1) \\ \text{vec}(\mathbf{P_z}) \end{bmatrix}$$

collect all structural parameters to be estimated. $\boldsymbol{\theta}$ is a $kg + rg \times 1$ column vector, so let $n_\theta = kg + rg$ denote the number of structural parameters. Analogously, let

$$\boldsymbol{\pi} = \begin{bmatrix} \text{vec}(\boldsymbol{\Sigma_{zu}}) \\ \text{vech}(\boldsymbol{\Sigma_{zu}}\boldsymbol{\Sigma_u^{-1}}\boldsymbol{\Sigma_{uz}}) \end{bmatrix}$$

collect all the reduced-form parameters. $\boldsymbol{\pi}$ has $kr + r(r+1)/2$ elements, so let $n_\pi = kr + r(r+1)/2$. Finally, let $\mathbf{V}_\pi$ be the variance–covariance matrix of $\boldsymbol{\pi}$.

For any candidate $\boldsymbol{\theta}$, define the function $\mathbf{f}(\boldsymbol{\theta})$ as

$$\mathbf{f}(\boldsymbol{\theta}) = \begin{bmatrix} \text{vec}(\mathbf{P_z}\mathbf{B}_1') \\ \text{vech}(\mathbf{P_z}\mathbf{P_z}') \end{bmatrix}$$

which maps $\boldsymbol{\theta}$ into $\boldsymbol{\pi}$. Then the distance function $d(\boldsymbol{\theta})$ is defined as

$$d(\boldsymbol{\theta}) = \{\boldsymbol{\pi} - \mathbf{f}(\boldsymbol{\theta})\}'\mathbf{V}_\pi^{-1}\{\boldsymbol{\pi} - \mathbf{f}(\boldsymbol{\theta})\}$$

and $\hat{\boldsymbol{\theta}}$ minimizes the distance function.

Standard errors for $\hat{\boldsymbol{\theta}}$ are computed via

$$\mathbf{V}_{\hat{\boldsymbol{\theta}}} = \left(\mathbf{F}_{\boldsymbol{\theta}}'\mathbf{V}_\pi^{-1}\mathbf{F}_{\boldsymbol{\theta}}\right)^{-1}$$

where

$$\mathbf{F}_{\boldsymbol{\theta}} = \begin{bmatrix} \mathbf{0}_{r(r+1)/2\times kg} & 2\mathbf{D}_r^+(\mathbf{P_z}\otimes\mathbf{I}_r) \\ (\mathbf{I}_k\otimes\mathbf{P_z})\mathbf{K}_{kg} & \mathbf{B}_1\otimes\mathbf{I}_r \end{bmatrix}$$

is the $n_\pi \times n_\theta$ derivative matrix of $\mathbf{f}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$. In this expression, $\mathbf{D}_r^+$ is the Moore–Penrose inverse of the duplication matrix, and $\mathbf{K}$ is the commutation matrix as defined in Magnus and Neudecker (2019, 54–55).

If qmatrix() is specified, then the $n_\pi \times n_\pi$ symmetric weight matrix $\mathbf{Q}$ is used in place of $\mathbf{V}_\pi$ in the distance function and in the variance calculation. The variance formula for $\hat{\boldsymbol{\theta}}$ takes the expression

$$\mathbf{V}_{\hat{\boldsymbol{\theta}}} = (\mathbf{F}_{\boldsymbol{\theta}}'\mathbf{Q}\mathbf{F}_{\boldsymbol{\theta}})^{-1}\mathbf{F}_{\boldsymbol{\theta}}'\mathbf{Q}\mathbf{V}_\pi\mathbf{Q}\mathbf{F}_{\boldsymbol{\theta}}(\mathbf{F}_{\boldsymbol{\theta}}'\mathbf{Q}\mathbf{F}_{\boldsymbol{\theta}})^{-1}$$

There must be at least as many reduced-form parameters as there are structural parameters to be estimated; $n_\pi \geq n_\theta$. With more than one target shock, n_θ exceeds n_π, and constraints must be placed on the elements of $\mathbf{B}_1$ or $\mathbf{P_z}$. There must be $g(g-1)/2$ such constraints.

References

Amisano, G., and C. Giannini. 1997. *Topics in Structural VAR Econometrics*. 2nd ed, revised and enlarged. Heidelberg: Springer.

Andrews, D. W. K. 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59: 817–858. https://doi.org/10.2307/2938229.

Angelini, G., and L. Fanelli. 2019. Exogenous uncertainty and the identification of structural vector autoregressions with external instruments. *Journal of Applied Econometrics* 34: 951–971. https://doi.org/10.1002/jae.2736.

Becketti, S. 2020. *Introduction to Time Series Using Stata*. Rev. ed. College Station, TX: Stata Press.

Gallant, A. R. 1987. *Nonlinear Statistical Models*. New York: Wiley. https://doi.org/10.1002/9780470316719.

Gertler, M., and P. Karadi. 2015. Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics* 7: 44–76. https://doi.org/10.1257/mac.20130329.

Hamilton, J. D. 1994. *Time Series Analysis*. Princeton, NJ: Princeton University Press. https://doi.org/10.2307/j.ctv14jx6sm.

———. 2003. What is an oil shock? *Journal of Econometrics* 113: 363–398. https://doi.org/10.1016/S0304-4076(02)00207-5.

Lütkepohl, H. 2005. *New Introduction to Multiple Time Series Analysis*. New York: Springer.

Magnus, J. R., and H. Neudecker. 2019. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. 3rd ed. Hoboken, NJ: Wiley.

Mertens, K., and M. O. Ravn. 2013. The dynamic effects of personal and corporate income tax changes in the United States. *American Economic Review* 103: 1212–1247. https://doi.org/10.1257/aer.103.4.1212.

Montiel Olea, J. L., J. H. Stock, and M. W. Watson. 2021. Inference in structural vector autoregressions identified with an external instrument. *Journal of Econometrics* 225: 74–87. https://doi.org/10.1016/j.jeconom.2020.05.014.

Newey, W. K., and K. D. West. 1994. Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61: 631–653. https://doi.org/10.2307/2297912.

Romer, C. D., and D. H. Romer. 2004. A new measure of monetary shocks: Derivation and implications. *American Economic Review* 94: 1055–1084. https://doi.org/10.1257/0002828042002651.

Stock, J. H., and M. W. Watson. 2001. Vector autoregressions. *Journal of Economic Perspectives* 15: 101–115. https://doi.org/10.1257/jep.15.4.101.

———. 2012. Disentangling the channels of the 2007–09 recession. *Brookings Papers on Economic Activity* 1: 81–135. https://doi.org/10.1353/eca.2012.0005.

Wieland, J. F., and M. Yang. 2020. Financial dampening. *Journal of Money, Credit and Banking* 52: 79–113. https://doi.org/10.1111/jmcb.12681.

Also see

For suggested citations, see the FAQ on citing Stata documentation.