xtvar — Panel-data vector autoregressive models

Description

`xtvar` fits vector autoregressive (VAR) models to panel data. Similar to VAR models for time-series data, `xtvar` models each dependent variable as a function of its own lags, the lags of all other dependent variables, and a panel-level fixed effect. Other explanatory variables can be added to the model as well; these variables can be predetermined, fully exogenous, or endogenous.

Quick start

Fit a panel-data VAR model with dependent variables `y1`, `y2`, and `y3` using `xtset` data

```
xtvar y1 y2 y3
```

Same as above, but use three lags of the dependent variables instead of one lag

```
xtvar y1 y2 y3, lags(3)
```

Same as above, but limit the number of lags of the dependent variables used as instruments to two

```
xtvar y1 y2 y3, lags(3) maxldep(2)
```

Same as above, but include exogenous variables `x1` and `x2`

```
xtvar y1 y2 y3, lags(3) maxldep(2) exogenous(x1 x2)
```

Same as above, but use first differences instead of the default forward-orthogonal deviations

```
xtvar y1 y2 y3, lags(3) maxldep(2) exogenous(x1 x2) fd
```

Same as above, but use the one-step generalized method of moments (GMM) estimator instead of the default two-step estimator

```
xtvar y1 y2 y3, lags(3) maxldep(2) exogenous(x1 x2) fd onestep
```

Menu

Statistics > Longitudinal/panel data > Dynamic panel data (DPD) > Panel-data VAR estimation

Syntax

`xtvar` *depvarlist* [*if*] [*in*] [, *options*]

options	Description
Model	
`lags(#)`	specify the number of lags for the dependent variables; default is `lags(1)`
`fodeviation`	use forward-orthogonal deviations to remove fixed effects; the default
`fd`	use first differences to remove fixed effects
`minldep(#)`	specify minimum number of lags of dependent, endogenous, and predetermined variables to use as instruments; default is `minldep(1)`
`maxldep(#)`	specify maximum lags of dependent, endogenous, and predetermined variables to use as instruments; default is all lags
`collapse`	collapse moment conditions from all time periods within each panel
Additional regressors	
`exogenous(`*varlist*`)`	specify strictly exogenous regressors
`endogenous(`*varlist*`)`	specify endogenous variables
`predetermined(`*varlist*`)`	specify predetermined variables
SE/Robust	
`vce(`*vcetype*`)`	*vcetype* may be `wcrobust`, `robust`, `bootstrap`, or `jackknife`; default is `vce(wcrobust)` for the two-step estimator and `vce(robust)` for the one-step estimator
Reporting	
`level(#)`	set confidence level; default is `level(95)`
display_options	control columns and column formats, row spacing, and line width
GMM	
`winitial(xt` | `identity)`	specify initial weight matrix; default is `winitial(xt)`
`onestep`	use the one-step GMM estimator rather than the two-step GMM estimator
`coeflegend`	display legend instead of statistics

You must `xtset` your data before using `xtvar`; see [XT] **xtset**.

depvarlist and *varlist* may contain time-series operators; see **[U] 11.4.4 Time-series varlists**.

`by`, `collect`, 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

`lags(#)` specifies the lags of the dependent variables to be included in the model. The default is `lags(1)`.

`fodeviation` requests that the fixed effects be removed from the model by applying the forward-orthogonal-deviation (FOD) transformation to all variables in the model. This is the default.

`fd` requests that the fixed effects be removed from the model by taking first-difference (FD) transformation of all variables in the model.

`minldep(#)` specifies the minimum number of lags of the dependent, endogenous, and predetermined variables that need to be available for use as instruments for an observation to be included in the estimation sample.

`maxldep(#)` specifies the maximum number of lags of the dependent, endogenous, and predetermined variables to use as instruments. The default is to use all available lags of all of these variables as instruments. To request the default explicitly, specify `maxldep(.)`.

The number of lags of the predetermined variables, $\mathbf{v}_{it}$, that will be used as instruments is actually $\#+1$. This is because for time t, $\mathbf{v}_{i,t-1}$ is a valid instrument, but the same is not true for the dependent variables or endogenous variables.

`collapse` requests that `xtvar` use a version of the moment conditions that sums across t within each panel. The standard GMM estimator can be severely biased when the number of moment conditions (which are based on the number of instruments) is large, as often happens with dynamic panel-data models. Collapsing the moment conditions ameliorates that bias at the expense of yielding a less-efficient estimator.

Additional regressors

`exogenous(`*varlist*`)` specifies a list of strictly exogenous variables to be included as regressors in the model.

`endogenous(`*varlist*`)` specifies a set of endogenous variables to be included in the model. By default, all available lags for these endogenous variables will be used as instruments; additionally, at least one lag of the endogenous variables must be available for an observation to be included in the estimation sample. You can use the `maxldep()` option to specify the maximum number of lags to use as instruments and the `minldep()` option to specify the minimum number of lags that need to be available to use as instruments for an observation to be included in the estimation sample.

`predetermined(`*varlist*`)` specifies a set of predetermined variables to be included in the model. A predetermined variable is a variable that at time t is affected by the error terms in previous time periods but not affected by the error terms in the current time period, t. By default, all available lags for these predetermined variables will be used as instruments; additionally, at least one lag of the predetermined variables must be available for an observation to be included in the estimation sample. You can use the `maxldep()` option to specify the maximum number of lags to use as instruments and the `minldep()` option to specify the minimum number of lags that need to be available to use as instruments for an observation to be included in the estimation sample.

SE/Robust

vce(*vcetype*) specifies the type of standard error reported, which includes types that are robust to some kinds of misspecification (wcrobust, robust) and that use bootstrap or jackknife methods (bootstrap, jackknife); see [XT] ***vce_options***.

vce(wcrobust), the default with the two-step estimator, requests the Windmeijer (2005) robust variance–covariance estimator, which corrects for the downward bias of the usual GMM cluster–robust variance–covariance estimator. This option is not available if you request the one-step estimator with the onestep option.

vce(robust), the default with the one-step estimator, requests the robust (sandwich) GMM variance–covariance estimator (VCE), which allows for intragroup correlation at the panel level. This is the default if you specify the onestep option. In small samples, this performs well with the one-step GMM estimator but is downward biased with the two-step estimator.

vce(bootstrap) and vce(jackknife) request VCEs based on the bootstrap or jackknife, respectively, where sampling is done at the panel level.

Reporting

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

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

GMM

winitial(xt | identity) specifies the initial weight matrix.

winitial(xt), the default, specifies an initial weight matrix based on the instruments specified with the model; additionally, this matrix assumes the idiosyncratic errors are homoskedastic. The exact form of the matrix depends on whether the FOD or FD transform is used to remove the model's fixed effects. See *Estimators* in *Methods and formulas* for more details.

winitial(identity) requests that the identity matrix be used as the initial weight matrix.

onestep specifies that the one-step estimator be calculated; the default is to use the two-step estimator.

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

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

Remarks and examples

Remarks are presented under the following headings:

Introduction

In a panel-data VAR model, each dependent variable is modeled as a function of its own lags, the lags of all other dependent variables, a panel-level fixed effect, and possibly other additional covariates. Therefore, panel-data VAR models combine elements of dynamic panel-data estimators and elements of time-series VAR models. To better understand the estimators implemented by `xtvar`, you should familiarize yourself with the Arellano and Bond (1991) estimator for dynamic panels (described in [XT] **xtabond**) as well as time-series VARs (described in [TS] **var**).

Below, we provide a short formal introduction to panel-data VAR models and their estimation; you can find more technical details in *Methods and formulas*. Then we provide a practical guide on how to fit a panel-data VAR model using `xtvar` and how to specify options that address common issues that arise with these models. We demonstrate how to check whether your model satisfies the moment conditions that are necessary for correct specification. We then examine the options available to reduce the number of moment conditions in the model and discuss in which cases you might want to use them. Additionally, we present the postestimation tools available after `xtvar` that can be used for diagnostics, testing, and interpretation. Many of these commands are ones you would use after fitting a VAR model using `var`. For instance, we discuss impulse–response functions (IRFs), which can be obtained by using the `irf` commands, to see the effects of a shock on an endogenous variable on itself or other endogenous variables; Granger causality tests, which can be performed by using `vargranger`; model selection that can be performed by using `xtvarsoc`; and lag-exclusion restrictions that can be tested by using `varwle`.

Panel-data VAR model formulation

Here we introduce basic aspects of a panel-data VAR model. As we mentioned above, this model combines elements of a dynamic panel-data model and elements of a VAR model. A benefit of a panel-data VAR model over a dynamic panel-data model is that we can model multiple dependent variables and their intertemporal relation. A benefit over a VAR model is that we can account for time-invariant heterogeneity. From both frameworks, we get the benefit of being able to explore dynamic behavior.

Panel-data VAR models express each of a set of K variables as a linear function of p of its own lags, p lags of the other $K-1$ variables, and, optionally, other covariates. The most general panel-data VAR model with p lags that `xtvar` accommodates is

$$\mathbf{y}_{it} = \mathbf{A}_1\mathbf{y}_{i,t-1} + \mathbf{A}_2\mathbf{y}_{i,t-2} + \cdots + \mathbf{A}_p\mathbf{y}_{i,t-p} + \mathbf{B}\mathbf{x}_{it} + \mathbf{C}\mathbf{w}_{it} + \mathbf{D}\mathbf{v}_{it} + \mathbf{u}_i + \boldsymbol{\epsilon}_{it} \tag{1}$$

where

$\mathbf{y}_{it}$ is a $K \times 1$ vector of dependent variables;

$\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_p$ are $K \times K$ matrices of parameters;

$\mathbf{x}_{it}$ is an $M_1 \times 1$ vector of strictly exogenous variables;

$\mathbf{B}$ is a $K \times M_1$ matrix of parameters;

$\mathbf{w}_{it}$ is an $M_2 \times 1$ vector of endogenous variables;

$\mathbf{C}$ is a $K \times M_2$ matrix of parameters;

$\mathbf{v}_{it}$ is an $M_3 \times 1$ vector of predetermined variables;

$\mathbf{D}$ is a $K \times M_3$ matrix of parameters;

$\mathbf{u}_i$ is a $K \times 1$ fixed-effects vector;

$\boldsymbol{\epsilon}_{it}$ is a $K \times 1$ vector of serially uncorrelated idiosyncratic errors;

$i = 1, \ldots, N$ denotes the ith panel; and

$t = 1, \ldots, T$ denotes the tth time period.

Note that (1) does not include a constant term because it cannot be identified together with the fixed effects $\mathbf{u}_i$.

Strictly exogenous variables, $\mathbf{x}_{it}$, are uncorrelated with past, present, or future realizations of the error term. Endogenous regressors, $\mathbf{w}_{it}$, on the other hand, may be correlated with the present and past realizations of the error term. Predetermined variables, $\mathbf{v}_{it}$, represent an intermediate case; they are not correlated with the present period's error term but may be correlated with past realizations of it. Both endogenous and predetermined variables are assumed to be uncorrelated with the future realizations of the error term. The sets of exogenous, endogenous, and predetermined variables determine the instrument set to be used and the complexity of your model.

To simplify our discussion, suppose that the only covariates in the model are the lags of the dependent variables. Thus, we have

$$\mathbf{y}_{it} = \mathbf{A}_1 \mathbf{y}_{i,t-1} + \mathbf{A}_2 \mathbf{y}_{i,t-2} + \cdots + \mathbf{A}_p \mathbf{y}_{i,t-p} + \mathbf{u}_i + \boldsymbol{\epsilon}_{it} \tag{2}$$

To fit this model, we first need to remove the fixed-effect term $\mathbf{u}_i$. `xtvar` provides two ways to do that: using either the FD transformation or the FOD transformation. See *Eliminating the fixed effect* in *Methods and formulas* for full descriptions of these transformations.

Applying the FD transformation to (2), we obtain

$$(\mathbf{y}_{it} - \mathbf{y}_{i,t-1}) = \mathbf{A}_1 (\mathbf{y}_{i,t-1} - \mathbf{y}_{i,t-2}) + \cdots + \mathbf{A}_p (\mathbf{y}_{i,t-p} - \mathbf{y}_{i,t-p-1}) + (\boldsymbol{\epsilon}_{it} - \boldsymbol{\epsilon}_{i,t-1})$$

Using tildes to represent transformed variables, we can write this more succinctly as

$$\tilde{\mathbf{y}}_{it} = \mathbf{A}_1 \tilde{\mathbf{y}}_{i,t-1} + \mathbf{A}_2 \tilde{\mathbf{y}}_{i,t-2} + \cdots + \mathbf{A}_p \tilde{\mathbf{y}}_{i,t-p} + \tilde{\boldsymbol{\epsilon}}_{it} \tag{3}$$

While we used the FD transformation, the discussion that follows also applies if $\tilde{\mathbf{y}}$ and $\tilde{\boldsymbol{\epsilon}}$ were produced by the FOD transformation.

Note that the first period at which we can evaluate (3) is $t = p + 2$.

We have removed the fixed-effect term, but the transformed regressor $\tilde{\mathbf{y}}_{i,t-1}$ is still endogenous because it is correlated with the transformed error $\tilde{\boldsymbol{\epsilon}}_{i,t-1}$ (through its dependence on $\boldsymbol{\epsilon}_{i,t-1}$). As shown by Nerlove (1967) and Nickell (1981) in the single-equation context, failing to account for this endogeneity may result in considerable bias, specially when T is small.

We must therefore find suitable instruments, $\mathbf{z}_{it}$, that are correlated with our endogenous regressors but not with our error term. We use the moment conditions $E\{\text{vec}(\mathbf{z}_{it}\tilde{\boldsymbol{\epsilon}}'_{it})\} = \mathbf{0}$ and GMM to estimate our parameters. The asymptotics of the GMM estimator rely on the number of panels going to infinity, so it is important that the number of panels in your data is not too small. For more background on GMM, Hayashi (2000), Cameron and Trivedi (2022), and Wooldridge (2010) provide textbook treatments. The monograph by Hall (2005) contains more detail. Also see [R] **gmm**. For more on the specific application of the GMM estimator in the context of panel-data VAR models, see *A concise representation of the GMM estimator* in *Methods and formulas*.

As in the dynamic panel literature (see Arellano and Bond [1991]), we can use the lags of the dependent variables as instruments for our endogenous regressors. Specifically, for endogenous regressor $\tilde{\mathbf{y}}_{i,t-1}$, we can use $\mathbf{y}_{i,t-2}, \mathbf{y}_{i,t-3}, \ldots, \mathbf{y}_{i,1}$, as instruments. Similarly, for endogenous regressor $\tilde{\mathbf{y}}_{i,t-2}$, we can use lags 3 and greater of $\mathbf{y}_{i,t}$ as instruments, and so on.

The total number of moment conditions in our GMM estimator grows rapidly with the T and K. In fact, the number of moment conditions in this simplified model with no additional covariates equals $(K^2/2)(T-p-1)(T+p-2)$. This multiplicity of moment conditions (and the multiplicity of instruments from which they are derived) poses serious challenges to estimation via GMM because some of the instruments may be only weakly relevant. See Han and Phillips (2006), Stock and Wright (2000), and Roodman (2009a).

`xtvar` has tools to ameliorate the issues arising from instrument proliferation. We can use the `maxldep()` and `collapse` options. The `maxldep()` option caps the total number of lags used as instruments for each endogenous regressor. So, for instance, specifying `maxldep(2)` requests that at most two lags be used as instruments for each endogenous regressor. The `collapse` option, on the other hand, reduces the number of moment conditions by taking averages of individual moment conditions and specifying the GMM estimator in terms of the averaged conditions instead of the individual ones. We will illustrate how to use the `maxldep()` and `collapse` options in the examples below.

So far, we have mentioned issues that are inherent to dynamic panel-data models. We also need to address the VAR dimension. We need to select the number of lags to include in our model, check that we have a stationary process, verify that the relationships between dependent variables are meaningful and have predictive power, and estimate the effects of shocks over time. We can do this with `xtvar` and its postestimation tools as we will demonstrate in examples below.

Fitting a panel-data VAR model with xtvar

▷ Example 1: Our first panel VAR

Blomquist and Dahlberg (1999) used a dataset consisting of 265 Swedish municipalities observed over 9 years. For each municipality, the dataset has variables recording expenditures, revenues, and grants from the central government. We want to fit a panel-data VAR to see how grants from the central government affect municipalities' expenditures and revenues over time. We first load in the dataset and describe its panel structure:

```
. use https://www.stata-press.com/data/r19/swedishgov
(1979-1987 Swedish municipality data)

. xtdescribe

  idcode:  114, 115, ..., 2584                                 n =        265
    year:  1979, 1980, ..., 1987                               T =          9
           Delta(year) = 1 year
           Span(year)  = 9 periods
           (idcode*year uniquely identifies each observation)

Distribution of T_i:   min      5%      25%       50%       75%     95%     max
                         9       9        9         9         9       9       9

     Freq.  Percent     Cum. |  Pattern
 ----------------------------+-----------
       265    100.00   100.00 |  111111111
 ----------------------------+-----------
       265    100.00          |  XXXXXXXXX
```

The data have already been `xtset` to specify that `idcode` is the panel identifier and `year` is the time variable. The output from `xtdescribe` shows that we have nine years of data and that there are no gaps in our dataset. `xtvar` supports datasets with gaps, but we defer that discussion until later.

We fit a panel-data VAR with dependent variables `expenditures`, `revenue`, and `grants` as follows:

```
. xtvar expenditure revenues grants

Panel-data vector autoregression                    Number of obs    = 1,855
Group variable: idcode                              Number of groups =   265
Time variable: year                                 Obs per group:
                                                                 min =     7
Number of moment conditions = 252                                avg =   7.0
                                                                 max =     7

Fixed-effects transform: FOD
Two-step results
                               (Std. err. adjusted for 265 clusters in idcode)
------------------------------------------------------------------------------
             |               WC robust
             | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
expenditures |
expenditures |
         L1. |   .2839341     .06484     4.38   0.000     .1568501    .4110182
             |
    revenues |
         L1. |  -.0451041   .0622281    -0.72   0.469     -.167069    .0768607
             |
      grants |
         L1. |   -1.68128   .2770326    -6.07   0.000    -2.224254   -1.138307
-------------+----------------------------------------------------------------
revenues     |
expenditures |
         L1. |   .2568554   .0781264     3.29   0.001     .1037304    .4099804
             |
    revenues |
         L1. |   .0598285   .0709236     0.84   0.399    -.0791791    .1988361
             |
      grants |
         L1. |   -2.24419   .2805223    -8.00   0.000    -2.794003   -1.694376
-------------+----------------------------------------------------------------
grants       |
expenditures |
         L1. |   .0164546   .0165141     1.00   0.319    -.0159124    .0488215
             |
    revenues |
         L1. |  -.0404274   .0143271    -2.82   0.005     -.068508   -.0123468
             |
      grants |
         L1. |   .3179538   .0506388     6.28   0.000     .2187037    .4172039
------------------------------------------------------------------------------
Hansen's test of overid. restrictions: chi2(243) = 264.16  Prob > chi2 = 0.168
GMM-type instruments: L(2/.).(expenditures revenues grants)
```

The first part of the output summarizes the panel structure of the estimation sample. The right-hand side shows that we have 265 groups (or panels) and 7 observations per group. The left-hand side shows that we used the FOD transformation and the default two-step GMM estimator with 252 moment conditions.

The coefficient table shows the estimated coefficients, standard errors, and related information. Our panel-data VAR includes three dependent variables, so there are three equations displayed in the table. We also have, by default, one lag of each dependent variable, so for each equation we see the coefficients for the first lag of each of the three variables.

Standard errors are labeled `WC robust`, meaning that they are based on the correction proposed in Windmeijer (2005) and thus robust to arbitrary within-panel correlation. We recommend using these standard errors with the two-step GMM estimator.

The footer of the table provides two valuable pieces of information: Hansen's J test of overidentifying restrictions and the instruments used to fit the model. A rejection of Hansen's test indicates that the moment conditions necessary for the correct specification of the model do not seem to hold. In this case, the test did not reject the null hypothesis that the moment conditions of our model hold.

◁

Modifying lags

While Hansen's J test did not indicate a problem with our model, we have reasons to believe that the effect of government grants on expenditures and revenues may last for more than one period. This would imply that our model with just one lag is misspecified. In the next examples, we explore models with more lags to address this issue, and we use postestimation tools to select the optimal number of lags and test their relevance in the model.

▷ Example 2: Adding lags and restricting the number used as instruments

To allow for dependence among dependent variables beyond the first lag, we can specify the `lags()` option; here we use two lags of the dependent variables by specifying `lags(2)`.

In example 1, we used all available lags of the dependent variables as instruments. This resulted in 252 moment conditions. As we mentioned before, having a proliferation of instruments and moment conditions makes our estimates less reliable. To reduce the number of moment conditions in the model, we specify the `maxldep(2)` option and fit our model using at most two lags as instruments for each endogenous regressor.

```
. xtvar expenditures revenues grants, lags(2) maxldep(2)

Panel-data vector autoregression                    Number of obs     = 1,590
Group variable: idcode                              Number of groups =    265
Time variable: year                                 Obs per group:
                                                                  min =     6
Number of moment conditions = 108                                 avg =   6.0
                                                                  max =     6

Fixed-effects transform: FOD
Two-step results
                               (Std. err. adjusted for 265 clusters in idcode)
```

	Coefficient	WC robust std. err.	z	P>|z|	[95% conf.	interval]
expenditures						
expenditures						
L1.	.1956019	.1147648	1.70	0.088	-.0293329	.4205367
L2.	.0017664	.100328	0.02	0.986	-.1948728	.1984056
revenues						
L1.	-.163357	.1162282	-1.41	0.160	-.39116	.0644459
L2.	-.3363544	.1003698	-3.35	0.001	-.5330756	-.1396332
grants						
L1.	-4.08135	.6900914	-5.91	0.000	-5.433904	-2.728796
L2.	-1.883438	.2732505	-6.89	0.000	-2.418999	-1.347877
revenues						
expenditures						
L1.	.1709229	.1220747	1.40	0.161	-.0683391	.4101848
L2.	.0525276	.1051278	0.50	0.617	-.1535191	.2585742
revenues						
L1.	-.092228	.1237745	-0.75	0.456	-.3348216	.1503656
L2.	-.32843	.0984127	-3.34	0.001	-.5213154	-.1355446
grants						
L1.	-4.7028	.6957627	-6.76	0.000	-6.06647	-3.33913
L2.	-2.054873	.2618687	-7.85	0.000	-2.568126	-1.54162
grants						
expenditures						
L1.	.0162825	.018789	0.87	0.386	-.0205433	.0531083
L2.	.0180168	.0164781	1.09	0.274	-.0142796	.0503132
revenues						
L1.	-.0281669	.0177173	-1.59	0.112	-.0628921	.0065583
L2.	-.0162105	.0161942	-1.00	0.317	-.0479506	.0155296
grants						
L1.	.2331196	.0762458	3.06	0.002	.0836807	.3825586
L2.	.1016583	.0487391	2.09	0.037	.0061315	.1971852

```
Hansen's test of overid. restrictions: chi2(90) = 228.48   Prob > chi2 = 0.000
GMM-type instruments: L(2/3).(expenditures revenues grants)
```

The bottom of the output now indicates that we have instrumented each of the endogenous regressors using lags two and three (`L(2/3)`) of the dependent variables when both are available. Otherwise, we instrumented the endogenous regressor using just lag two. The number of moment conditions is now 108 versus 252 in example 1. We have made progress in reducing the number of moment conditions. However, Hansen's J test provides strong evidence that our moment conditions are not valid. So we would hesitate to use these results without further investigation. You can verify that when we fit the model specifying `maxldep(1)` or `maxldep(3)`, we continue to reject the null hypothesis that our moment conditions are valid.

One possibility is that our lag specification of the model is still inadequate. If we do not include enough lags in the model, then the errors could be serially correlated. In that case, lags two and greater of the dependent variables are no longer valid instruments. Let's explore a model with four lags of the dependent variables as regressors and a maximum of two lags of the dependent variables to be used as instruments.

```
. xtvar expenditures revenues grants, lags(4) maxldep(2)

Panel-data vector autoregression                      Number of obs    = 1,060
Group variable: idcode                                Number of groups =   265
Time variable: year                                   Obs per group:
                                                                   min =     4
Number of moment conditions = 72                                   avg =   4.0
                                                                   max =     4

Fixed-effects transform: FOD
Two-step results
                               (Std. err. adjusted for 265 clusters in idcode)
------------------------------------------------------------------------------
             |              WC robust
             | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
expenditures |
expenditures |
         L1. |   .3043156   .2596238     1.17   0.241    -.2045376    .8131689
         L2. |    .176059   .2198797     0.80   0.423    -.2548973    .6070153
         L3. |   .0807466   .2439179     0.33   0.741    -.3973237     .558817
         L4. |    .362827   .4003709     0.91   0.365    -.4218855    1.147539
             |
    revenues |
         L1. |  -.2788411   .2972961    -0.94   0.348    -.8615308    .3038486
         L2. |  -.2349415   .2596951    -0.90   0.366    -.7439345    .2740515
         L3. |  -.2040393   .2356327    -0.87   0.387    -.6658708    .2577923
         L4. |  -.5232187   .3816921    -1.37   0.170    -1.271321     .224884
             |
      grants |
         L1. |   1.012279   .9715139     1.04   0.297     -.891853    2.916412
         L2. |   .3069275   .4503541     0.68   0.496    -.5757503    1.189605
         L3. |   1.000217   .5569212     1.80   0.072    -.0913287    2.091762
         L4. |    .077072   1.548516     0.05   0.960    -2.957963    3.112107
```

```
─────────────┬──────────────────────────────────────────────────────────────────
revenues     │
expenditures │
         L1. │   .5287675   .2030516     2.60   0.009     .1307937    .9267412
         L2. │   .3753373   .1738501     2.16   0.031     .0345973    .7160773
         L3. │   .2828897   .1838684     1.54   0.124    -.0774858    .6432651
         L4. │    .693354   .3521214     1.97   0.049     .0032087    1.383499
             │
    revenues │
         L1. │  -.4533828   .2293747    -1.98   0.048     -.902949   -.0038167
         L2. │  -.4150536    .207767    -2.00   0.046    -.8222694   -.0078378
         L3. │  -.4150709   .1865944    -2.22   0.026    -.7807891   -.0493527
         L4. │  -.7522478   .3355086    -2.24   0.025    -1.409833   -.0946629
             │
      grants │
         L1. │   .0457337   .7650976     0.06   0.952    -1.45383     1.545298
         L2. │  -.1807682   .3181763    -0.57   0.570    -.8043823     .442846
         L3. │   .5877744   .5370399     1.09   0.274    -.4648045    1.640353
         L4. │  -.5295232   1.330482    -0.40   0.691    -3.137221    2.078174
─────────────┼──────────────────────────────────────────────────────────────────
grants       │
expenditures │
         L1. │  -.0747673   .0932016    -0.80   0.422    -.2574391    .1079044
         L2. │   -.039655   .0806699    -0.49   0.623    -.1977651     .118455
         L3. │  -.1083643   .0865062    -1.25   0.210    -.2779134    .0611848
         L4. │  -.0165543   .1265477    -0.13   0.896    -.2645833    .2314747
             │
    revenues │
         L1. │   .0551585   .1044118     0.53   0.597    -.1494848    .2598018
         L2. │   .0471306   .0934546     0.50   0.614    -.1360369    .2302982
         L3. │   .1051582   .0851519     1.23   0.217    -.0617365    .2720529
         L4. │  -.0509242   .1199238    -0.42   0.671    -.2859705     .184122
             │
      grants │
         L1. │  -.2410357   .2761211    -0.87   0.383    -.7822232    .3001517
         L2. │  -.0788769    .141957    -0.56   0.578    -.3571075    .1993537
         L3. │   .2370528   .2222193     1.07   0.286    -.1984889    .6725946
         L4. │   .2673538   .4635082     0.58   0.564    -.6411055    1.175813
─────────────┴──────────────────────────────────────────────────────────────────
Hansen's test of overid. restrictions: chi2(36) = 38.80    Prob > chi2 = 0.345
GMM-type instruments: L(2/3).(expenditures revenues grants)
```

As we can see from Hansen's J test, we cannot reject the null that the moment conditions in our model hold. We were previously concerned about both the lag structure and the number of instruments. Hansen's J test suggests that the lag structure was the more relevant concern in our case.

◁

Reducing moment conditions by collapsing the instrument matrix

Collapsing the instrument matrix (and thus collapsing the moment conditions) is another alternative to address the issues that come with instrument proliferation. We explore it in the next example.

▷ Example 3: Collapsing the instrument matrix

We refit our municipal finance model with two lags. This time we use all available lags as instruments but request the collapsed version of the instrument matrix by specifying the `collapse` option. This reduces the number of moment conditions because the GMM estimation now uses average moment conditions across time instead of individual moment conditions.

```
. xtvar expenditures revenues grants, lags(2) collapse

Panel-data vector autoregression                    Number of obs     = 1,590
Group variable: idcode                              Number of groups =    265
Time variable: year                                 Obs per group:
                                                                 min =      6
Number of moment conditions = 63 (collapsed)                     avg =    6.0
                                                                 max =      6
Fixed-effects transform: FOD
Two-step results
                              (Std. err. adjusted for 265 clusters in idcode)

                           WC robust
             Coefficient  std. err.      z    P>|z|     [95% conf. interval]

expenditures
expenditures
         L1.    .1900148   .1513004    1.26   0.209    -.1065285    .4865581
         L2.    .0327313    .147236    0.22   0.824    -.2558458    .3213085

    revenues
         L1.   -.2920254   .1556137   -1.88   0.061    -.5970226    .0129719
         L2.   -.4337591   .1341645   -3.23   0.001    -.6967167   -.1708014

      grants
         L1.   -5.062357   .9626468   -5.26   0.000     -6.94911   -3.175604
         L2.   -2.221608   .4108095   -5.41   0.000     -3.02678   -1.416436

revenues
expenditures
         L1.    .1919241   .1543727    1.24   0.214    -.1106408     .494489
         L2.     .095686   .1481062    0.65   0.518    -.1945967    .3859688

    revenues
         L1.   -.2275803   .1544383   -1.47   0.141    -.5302737    .0751132
         L2.   -.4275827   .1291448   -3.31   0.001    -.6807018   -.1744635

      grants
         L1.   -5.526232   .9128089   -6.05   0.000    -7.315304   -3.737159
         L2.   -2.382141   .3992335   -5.97   0.000    -3.164624   -1.599658

grants
expenditures
         L1.    .0107939   .0220194    0.49   0.624    -.0323634    .0539512
         L2.    .0176607   .0170148    1.04   0.299    -.0156878    .0510092

    revenues
         L1.   -.0191644   .0198971   -0.96   0.335     -.058162    .0198332
         L2.   -.0103149    .016148   -0.64   0.523    -.0419645    .0213346

      grants
         L1.    .3128186   .0756299    4.14   0.000     .1645867    .4610504
         L2.    .1347204   .0584898    2.30   0.021     .0200826    .2493582

Hansen's test of overid. restrictions: chi2(45) = 211.25   Prob > chi2 = 0.000
GMM-type instruments: L(2/.).(expenditures revenues grants)
```

The number of moment conditions after collapsing is 63, which is considerably smaller than with the previous examples. In the header of the output, the moment conditions are labeled "`(collapsed)`" to indicate that `xtvar` used a collapsed instrument matrix as we requested.

Despite the significant reduction in the number of moment conditions, we see that Hansen's J test still suggests our moment conditions are not valid. As we mentioned at the end of the previous example, the real solution is to increase the number of lags we include in the model.

◁

Lag-order selection

So far, we have explored how we can use different options to modify the model we fit and what this implies regarding the number of parameters, the instruments in our model, and Hansen's J test. Rather than fitting many models manually and comparing results, we can use the `xtvarsoc` command to help select the number of lags to satisfy the moment conditions.

▷ Example 4: Reconsidering the lag order

After fitting a model with `xtvar`, we can use `xtvarsoc` to obtain model- and moment-selection criteria (MMSC) to help determine the correct lag length in our model. MMSC are an adaption of the Akaike (1973) information criterion (AIC), Schwarz's (1978) Bayesian information criterion (BIC), and Hannan and Quinn (1979) information criterion (HQIC) for panel-data VAR models. Panel-data VARs are fit using GMM, and no distributional assumptions are made; this precludes the computation of a likelihood function and the standard information criteria.

Andrews and Lu (2001) developed variants of the AIC, BIC, and HQIC for use with dynamic panel-data models. They refer to their statistics as MMSC (rather than using lag-order selection criteria terminology used with time-series VARs) because their statistics can be used not only to select a lag length but also to select a set of instruments conditional on lag length. We will refer to the three of Andrews and Lu's statistics implemented by `xtvarsoc` as MMSC-AIC, MMSC-BIC, and MMSC-HQIC. These statistics are based on Hansen's J test statistic rather than the maximized likelihood function value, and they apply penalty factors for the sample size and number of overidentifying restrictions. Here we use `xtvarsoc` to obtain the MMSC statistics for a range of panel-data VAR models based on the instrument specification of our model with two lags and a maximum of two lags used as instruments. We will request that `xtvarsoc` consider models with up to four lags. We type

```
. xtvar expenditures revenues grants, lags(2) maxldep(2)
  (output omitted)
. xtvarsoc, maxlag(4)
Model- and moment-selection criteria
```

Lag	N	MC	Hansen's J	df	p	MMSC-AIC	MMSC-BIC	MMSC-HQIC
1	1060	72	206.44	63	0.000	80.443	-232.42*	-38.129
2	1060	72	182.98	54	0.000	74.977	-193.19	-26.656
3	1060	72	103.72	45	0.000	13.72	-209.75	-70.974
4	1060	72	38.80	36	0.345	-33.199*	-211.98	-100.95*

```
* indicates minimum value within column.
Note: Maximum lag order specified exceeds lags used at estimation; using
      reduced sample.
GMM-type instruments: L(2/3).(expenditures revenues grants)
```

xtvarsoc computed the MMSC for models with from one to four lags. Your first inclination may be to look at the MMSC statistics to find the best model. Before you do that, though, you need to make sure that the numbers are comparable and relevant. In particular, you should look at the sample size, number of moment conditions, and Hansen's J test. Once those check out, then you can search for the best model on the basis of the MMSC statistics.

We first verify that the sample size is the same across models. The sample size reported in each row is 1,060. Our data have 265 panels, implying $1060/265 = 4$ observations per panel. There is a note near the bottom of the output indicating xtvarsoc used a reduced sample; we used 2 lags in our xtvar specification and are using xtvarsoc for models with up to 4 lags. For MMSC comparisons to be valid, the statistics must be computed on the same sample of data. xtvarsoc first fits a model using the highest number of lags requested and then uses that model's estimation sample for models with fewer lags. That generally results in a constant sample size across lags, though we cannot rule out the possibility that an odd pattern of missing data could cause the number of observations to change across different model lags. xtvarsoc will issue an error message if it detects that the sample size is not constant across estimates.

When comparing models to determine the optimal lag length, you must also make sure that the same set of instruments is being used across models. The final line of the output shows we are using lags 2 and 3 of our left-hand-side variables as instruments, just as we used when fitting our original model. The column labeled MC, which stands for moment conditions, indicates that each of the 4 candidate models here were fit using 72 moment conditions. Our original model used 108 moment conditions, while the models fit here used 72; the difference stems from the smaller sample size here because of our request to consider a model with 4 lags.

The number of moment conditions must be the same across models with different lag lengths when choosing the number of lags. In [XT] **xtvar postestimation**, we show how to use the MMSC statistics to select the optimal set of instruments instead of the optimal number of lags. There we will require the number of lags in different models to be the same. Either the number of lags or the number of moment conditions can change across models to make valid comparisons but not both at the same time. xtvarsoc will exit with an error message if it detects that comparisons will not be valid.

The column marked Hansen's J contains Hansen's J test statistic for each model, and the columns marked df and p contain, respectively, the corresponding degrees of freedom and p values. The degrees of freedom change across models because the number of estimated parameters changes; that this column's values change does not invalidate MMSC comparisons. In this example, we reject the null hypothesis that our moment conditions are valid for models with one, two, and three lags. Based on Hansen's J test, we are left with just one model: the one with four lags.

Having verified that the sample size and the number of moment conditions remain constant across models and having tested for instrument validity, now we are ready to consider the MMSC statistics. As with standard information criteria, a model with a lower value is to be preferred over a model with a higher value. Here the MMSC-AIC and MMSC-HQIC select a model with four lags, which is fortuitous because that model is the only one for which the moment conditions are valid based on Hansen's J test. The MMSC-BIC selects a model with one lag, though we would probably reject that model because of Hansen's J test. Andrews and Lu (2001) performed Monte Carlo analyses and recommend the MMSC-BIC as the best all-around criterion to use, assuming the orthogonality conditions are valid.

◁

Including endogenous covariates

▷ Example 5: Endogenous covariates

Suppose we are primarily interested in the relationship between grants and expenditures, but we suspect that revenues are also endogenously determined. In other words, there is a dynamic relationship between grants and expenditures, while revenues have a contemporaneous effect. An alternative to fitting a three-variable panel-data VAR model is to fit a two-variable model and include revenues as an additional endogenous covariate. Below, we fit this model using the endogenous() option and again requesting that at most two lags of each variable be used as instruments.

```
. xtvar expenditures grants, lags(2) maxldep(2) endogenous(revenues)

Panel-data vector autoregression                   Number of obs     = 1,590
Group variable: idcode                             Number of groups  =   265
Time variable: year                                Obs per group:
                                                                 min =     6
Number of moment conditions = 72                                 avg =   6.0
                                                                 max =     6

Fixed-effects transform: FOD
Two-step results
                              (Std. err. adjusted for 265 clusters in idcode)
------------------------------------------------------------------------------
             |               WC robust
             | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
expenditures |
expenditures |
         L1. |   -.036028   .0217366    -1.66   0.097    -.0786309    .0065749
         L2. |  -.0580125   .0204245    -2.84   0.005    -.0980438   -.0179813
             |
      grants |
         L1. |   .6735404   .2300296     2.93   0.003     .2226906     1.12439
         L2. |   .2240255   .1219378     1.84   0.066    -.0149682    .4630192
             |
    revenues |   .9932527   .0296293    33.52   0.000     .9351804    1.051325
-------------+----------------------------------------------------------------
grants       |
expenditures |
         L1. |  -.0068636   .0076561    -0.90   0.370    -.0218693     .008142
         L2. |    .003658   .0068504     0.53   0.593    -.0097685    .0170846
             |
      grants |
         L1. |   .3318416   .0841429     3.94   0.000     .1669245    .4967586
         L2. |     .16419    .045025     3.65   0.000     .0759426    .2524374
             |
    revenues |   .0083887   .0098906     0.85   0.396    -.0109966     .027774
------------------------------------------------------------------------------
Hansen's test of overid. restrictions: chi2(62) = 142.48   Prob > chi2 = 0.000
Added endogenous: revenues
GMM-type instruments: L(2/3).(expenditures grants revenues)
```

The coefficients on revenues are 0.99 and 0.01 in the equations for expenditures and grants, respectively. When revenues are treated as endogenous covariates, the results of this model indicate that they have a positive effect on expenditures but are not relevant for grants.

◁

Lag exclusion tests

▷ Example 6: Wald lag exclusion tests

After we fit a panel-data VAR model, one hypothesis of interest is that all the endogenous variables at a given lag are jointly zero. `varwle` reports Wald tests of this hypothesis for each equation and for all equations jointly. We refit our panel-data VAR model using four lags of the dependent variables as covariates, which is the only specification we have fit so far in which the validity of the moment conditions has not been rejected by Hansen's J test.

```
. xtvar expenditures revenues grants, lags(4) maxldep(2)
  (output omitted)
. varwle
```

Equation: expenditures

lag	chi2	df	Prob > chi2
1	6.100499	3	0.107
2	4.914952	3	0.178
3	10.01386	3	0.018
4	33.88895	3	0.000

Equation: revenues

lag	chi2	df	Prob > chi2
1	10.6371	3	0.014
2	5.669409	3	0.129
3	14.89268	3	0.002
4	21.80127	3	0.000

Equation: grants

lag	chi2	df	Prob > chi2
1	4.001778	3	0.261
2	3.773334	3	0.287
3	2.394958	3	0.495
4	46.76529	3	0.000

Equation: All

lag	chi2	df	Prob > chi2
1	22.40912	9	0.008
2	14.22638	9	0.114
3	21.47419	9	0.011
4	89.77261	9	0.000

The first block of output refers to the equation for `expenditures`. Line 1 of this output reports a Wald test for the first lag of the 3 dependent variables. Here the χ^2 statistic is 6.10 ($p \approx 0.107$), so we cannot reject the null hypothesis that the 3 coefficients are jointly 0.

The first line in the first block is equivalent to typing

```
. test [expenditures]_b[L1.expenditures]   ///
       [expenditures]_b[L1.revenues]       ///
       [expenditures]_b[L1.grants]
```

but `varwle` does all the work for us and assembles the results into nice tables. The remaining lines in the first block test the same set of coefficients on the model variables at lags two through four. The next two blocks perform similar tests for the other two equations.

The final block is arguably more useful. It tests all the coefficients of all the variables for all equations at a particular lag. Referring back to (1), each line of the final block contains a test of the hypothesis that $\mathbf{A}_\ell = \mathbf{0}$ for $\ell = 1, 2, 3,$ or 4. If you fit a model with multiple lags and find that the last lag's coefficient matrix provides evidence to support the null hypothesis, you could consider refitting your model without that lag.

What is perhaps surprising about our results is that for `expenditures` the first two lags provide evidence to support the null and similarly for the first three lags of `grants`. However, for all three variables, the fourth lag's coefficients do not provide evidence to support the null hypothesis, perhaps suggesting a slow response. In terms of the $\mathbf{A}$ coefficient matrices, we find that our estimates $\widehat{\mathbf{A}}_1$, $\widehat{\mathbf{A}}_3$, and $\widehat{\mathbf{A}}_4$ do not provide evidence to support the null hypothesis.

◁

Granger causality test

Granger causality tests are a popular tool in the VAR analyst's toolkit. Granger causality tests, in the simplest case, test whether lags of one variable are useful in predicting the values of another variable. More formally, variable x Granger-causes variable y if for all $s > 0$ the mean squared error (MSE) of a forecast of y_{t+s} based on $(y_{t-1}, y_{t-2}, \ldots, x_{t-1}, x_{t-2}, \ldots)$ is lower than a forecast using only $(y_{t-1}, y_{t-2}, \ldots)$. See Hamilton (1994, chap. 11). The extension to variables x and z jointly Granger-causing variable y is immediate.

Just as with time-series VARs, we can conduct Granger causality tests after panel-data VARs by performing Wald tests of the joint significance of the parameters associated with the lags of one or more variables in an equation. In our example, to see if `revenues` Granger-causes `expenditures`, we would test the hypothesis that the coefficients on `L1.revenues`, `L2.revenues`, `L3.revenues`, and `L4.revenues` in the equation for `expenditures` are jointly equal to zero. `vargranger` makes that, and much more, very easy.

▷ Example 7: Granger causality

Typing `vargranger` after fitting our four-lag model, we obtain

```
. vargranger
```

Granger causality Wald tests

Equation	Excluded	chi2	df	Prob > chi2
expenditures	revenues	5.0979	4	0.277
expenditures	grants	5.0767	4	0.280
expenditures	ALL	34.47	8	0.000
revenues	expenditures	9.3224	4	0.054
revenues	grants	3.4496	4	0.486
revenues	ALL	20.979	8	0.007
grants	expenditures	5.5802	4	0.233
grants	revenues	7.5613	4	0.109
grants	ALL	59.35	8	0.000

Consider the first line of output, which refers to the equation for `expenditures`. Under column `Excluded`, we see the `revenues` variable. This line of output is therefore a test of whether `revenues` Granger-causes `expenditures`. In other words, by including `revenues` in the equation for `expenditures`, do we obtain a lower forecast MSE for `expenditures` than if we had not included `revenues`? Here the χ^2 statistic is 5.10 with 4 degrees of freedom. Because the corresponding p-value is large (0.277), we cannot reject the null hypothesis that `revenues` does not Granger-causes `expenditures`. If we were building a forecast model of `expenditures`, this line of output suggest that we would not lower our forecast MSE by including `revenues`.

Similarly, the second line of output provides no evidence that `grants` Granger-causes `expenditures`, either.

The results in the third line stand in contrast to the first two lines' results. Shown in column `Excluded` is the keyword `ALL`, which represents all variables in the equation for `expenditures` other than (lags of) `expenditures` itself. This line contains a Wald test that the coefficients on the lags of both `revenues` and `grants` are all jointly equal to zero. It is a test of whether `revenues` and `grants` jointly Granger-causes `expenditures`. Here we reject the null hypothesis of no Granger causality and conclude that `revenues` and `grants` do Granger-causes `expenditures`.

These results suggest neither `revenues` nor `grants` Granger-causes `expenditures`, conditional on the other variable being in the model. However, the two variables do jointly Granger-causes `expenditures`.

The last three lines of output, pertaining to `grants`, show a similar pattern to those for `expenditures`.

◁

Verifying the stability condition of the VAR

▷ Example 8: Checking stability

For the relationships implied by our estimates to be meaningful, we need to verify that we have a stationary process. To check whether our fitted VAR represents a stable dynamic process, we use

```
. varstable, graph
  (output omitted)
```

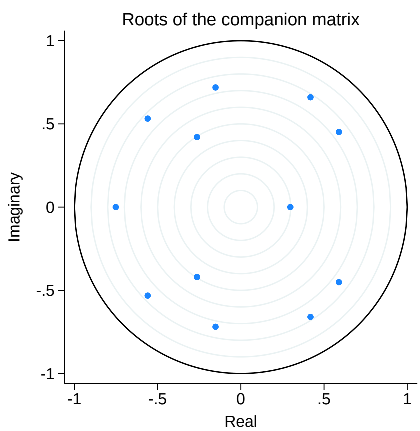

Figure 1. Checking the stability of our panel-data VAR model

Our panel-data VAR model contains 3 variables and 4 lags, so the companion matrix will have $3 \times 4 = 12$ eigenvalues, some of which are conjugate pairs. Looking at figure 1, we see that all the eigenvalues lie within the unit circle. This indicates that our panel-data VAR model does satisfy the stability condition. We could have alternatively displayed the values and modulus of the eigenvalues in a table by simply omitting the `graph` option after `varstable`.

Because our model represents a stable dynamic process, we can proceed to compute and analyze IRFs.

◁

IRFs

An IRF measures the effect across time of a shock to an endogenous variable on itself or another endogenous variable. After fitting a panel-data VAR using `xtvar`, you can obtain simple IRFs, cumulative IRFs, orthogonalized IRFs, and cumulative orthogonalized IRFs. A full description of IRFs and how you obtain them in Stata is in [TS] **irf**. Here we cover the highlights to get you started.

In Stata, an IRF set is a special dataset that contains one or more IRF results. An IRF result refers to the IRFs, cumulative IRFs, and other statistics as well as their standard errors and is obtained by calling `irf create` one time. Stata has several commands for manipulating IRF sets and results within IRF sets. For our purposes, though, we can accomplish everything we will do just by using `irf create` and specifying appropriate options.

▷ Example 9: Simple IRFs

We first create the IRF results from our current four-lag model that is still in memory, and we store them under the name `four_lags`. We will store those results in an IRF set called `example1`, and we will create the IRFs for nine steps:

```
. irf create four_lags, set(example1) step(9)
(file example1.irf created)
(file example1.irf now active)
(file example1.irf updated)
```

If IRF set `example1` already contained results stored under the name `four_lags`, the previous command would have issued an error message. If that happens and you wish to replace the existing results with that name, you can specify `replace` within the `set()` option.

With an IRF result available, we can plot the IRFs:

```
. irf graph irf, irf(four_lags)
```

When we used `irf create`, we made `irf` set `example1` active, and Stata remembered that action. Therefore, when we use the `irf graph` command without the `set()` option, we use this currently active IRF set and obtain the following graph:

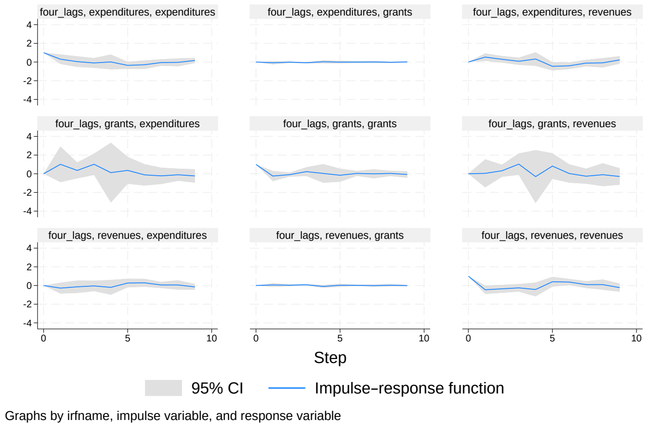

Figure 2. IRFs of our panel-data VAR model

The first row of graphs shows the effect of a shock to expenditures on expenditures themselves, grants, and revenues. In the second row, we see the effects of a shock to grants on each variable. And in the last row, we see the effects of a shock to revenues on each variable.

◁

Simple IRFs are often uninformative because they assume that a shock happens to a single variable in isolation. In the vast majority of applications, shocks affect multiple variables at the same time. One solution is to use orthogonalized IRFs, and the most common way to orthogonalize shocks is with the Cholesky decomposition of the error covariance matrix. See Hamilton (1994), Lütkepohl (2005), and Kilian and Lütkepohl (2017) for more information.

Getting Cholesky-orthogonalized IRFs after fitting a panel-data VAR requires no additional work: earlier when we called `irf create` before plotting simple IRFs, it automatically created the orthogonalized IRFs as well as the cumulative variants of both of those types of IRFs. To view the orthogonalized IRFs, we just tell `irf graph` or a similar command to plot the orthogonalized IRFs instead.

▷ Example 10: Orthogonalized IRFs

To create tables or graphs of orthogonalized IRFs, we simply change the *stat* from `irf` to `oirf` in the table or graph specification. Here we use the `irf cgraph` command to combine graphs of two of the orthogonalized IRFs of interest. We plot the effects of an orthogonalized shock to `grants` on `revenues` and `expenditures`:

```
. irf cgraph (four_lags grants revenues oirf)
>            (four_lags grants expenditures oirf)
```

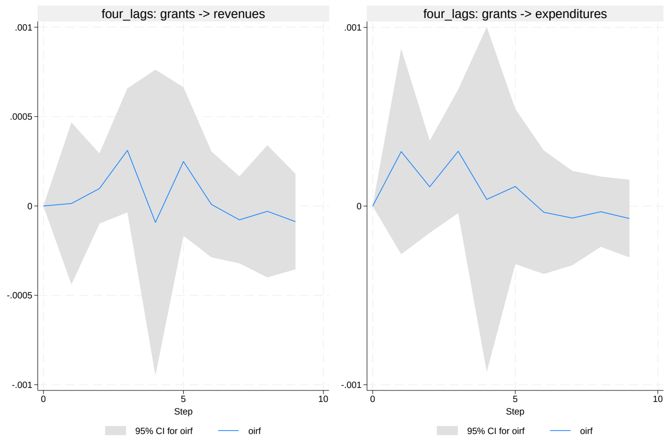

Figure 3. Orthogonalized IRFs of our panel-data VAR model

The orthogonalized IRF graphs both have wide confidence intervals that include 0. We do not have evidence that a random shock to `grants` has an effect on `revenues` or `expenditures`.

◁

Stored results

xtvar stores the following in e():

Scalars

e(N)	number of observations
e(n_zc)	number of columns of instrument matrix
e(J)	Hansen's J statistic
e(p_J)	p-value for Hansen test
e(df_J)	degrees of freedom for Hansen test
e(rank_weight)	rank of final weight matrix
e(k)	number of parameters
e(k_eq)	number of equations
e(N_clust)	number of clusters
e(N_g)	number of groups
e(g_min)	smallest group size
e(g_max)	largest group size
e(g_avg)	average group size
e(tmin)	first period in estimation sample
e(tmax)	last period in estimation sample
e(maxldep_act)	largest lag order actually used for instruments
e(maxldep)	maximum number of instruments requested
e(minldep)	minimum number of instruments requested
e(lags)	number of lags in model

Macros

e(cmd)	xtvar
e(cmdline)	command as typed
e(depvar)	names of dependent variables
e(vce)	*vcetype* specified in vce()
e(vcetype)	title used to label Std. err.
e(winit_type)	type of initial weight matrix used
e(transform)	specified transform for removing fixed effects
e(collapse)	collapse, if specified
e(predetermined)	predetermined variables, if specified
e(endogenous)	additional endogenous variables, if specified
e(exogenous)	exogenous variables, if specified
e(onestep)	onestep, if specified
e(ivar)	group variable specified in xtset
e(tvar)	time variable specified in xtset
e(tsfmt)	display format for time variable
e(datasignaturevars)	variables used in calculation of checksum
e(datasignature)	checksum from datasignature
e(properties)	b V
e(predict)	program used to implement predict
e(eqnames)	names of equations

Matrices

e(b)	coefficient vector
e(V)	variance–covariance matrix of the estimators

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:

Introduction

The panel-data VAR model with p lags, K regressands, M_1 strictly exogenous variables $\mathbf{x}_{it}$, M_2 additional endogenous variables $\mathbf{w}_{it}$, and M_3 predetermined variables $\mathbf{v}_{it}$ is

$$\mathbf{y}_{it} = \mathbf{A}_1\mathbf{y}_{i,t-1} + \mathbf{A}_2\mathbf{y}_{i,t-2} + \cdots + \mathbf{A}_p\mathbf{y}_{i,t-p} + \mathbf{B}\mathbf{x}_{it} + \mathbf{C}\mathbf{w}_{it} + \mathbf{D}\mathbf{v}_{it} + \mathbf{u}_i + \boldsymbol{\epsilon}_{it} \tag{4}$$

where i indexes panels 1 through N, t indexes time from 1 to T, $\mathbf{u}_i$ is a $K \times 1$ fixed-effect vector, and $\boldsymbol{\epsilon}_{it}$ is a $K \times 1$ vector of idiosyncratic i.i.d shocks. $\mathbf{A}_1, \dots, \mathbf{A}_p$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ are parameter matrices to be estimated.

Fully exogenous variables are uncorrelated with past, present, or future realizations of the error term. Therefore, they satisfy the condition $E\{\text{vec}(\mathbf{x}_{it}\boldsymbol{\epsilon}'_{is})\} = \mathbf{0}$ for all s and t.

Endogenous variables, on the other hand, are not correlated with future realizations of the error term but may be correlated with present or past realizations of it. This means that $E\{\text{vec}(\mathbf{w}_{it}\boldsymbol{\epsilon}'_{is})\} = \mathbf{0}$ for $s > t$, but possibly $E\{\text{vec}(\mathbf{w}_{it}\boldsymbol{\epsilon}'_{is})\} \neq \mathbf{0}$ for $s \leq t$.

Predetermined variables represent an intermediate case. They are not correlated with future or present realizations of the error term but may be correlated with past realizations of it. Therefore, they satisfy $E\{\text{vec}(\mathbf{v}_{it}\boldsymbol{\epsilon}'_{is})\} = \mathbf{0}$ for $s \geq t$, but possibly $E\{\text{vec}(\mathbf{v}_{it}\boldsymbol{\epsilon}'_{is})\} \neq \mathbf{0}$ for $s < t$.

Eliminating the fixed effect

The first task in estimating the panel-data VAR model is to eliminate the fixed effect term $\mathbf{u}_i$ using either the FD transformation or the FOD transformation.

For an arbitrary variable h_{it}, the FD transformation is

$$h_{it}^{\text{FD}} = h_{it} - h_{i,t-1}$$

for $t = 2, \dots, T$. Note that, if h_{it} is missing, then both h_{it}^{FD} and $h_{i,t+1}^{\text{FD}}$ will be missing.

The FOD transformation, introduced by Arellano and Bover (1995), is defined as follows. Let $\Omega(i,t)$ denote the set of periods for which $h_{i\tau}$ is nonmissing for $\tau = t+1,\dots,T$. Let $\#_{\Omega(i,t)}$ denote the number of elements in set $\Omega(i,t)$. Then

$$h_{it}^{\text{FOD}} = \sqrt{\frac{\#_{\Omega(i,t)}}{\#_{\Omega(i,t)}+1}}\left(h_{it} - \frac{1}{\#_{\Omega(i,t)}}\sum_{\tau\in\Omega(i,t)} h_{i\tau}\right)$$

and exists for $t = 1,\dots,T-1$. To make the formulas provided below work regardless of which transform is used, we use the common practice of storing the value h_{it}^{FOD} in period $t+1$.

We adopt the tilde diacritic to represent a variable that has undergone one of these transformations: $\widetilde{h}_{it} = h_{it}^{\text{FD}}$ or $\widetilde{h}_{it} = h_{it}^{\text{FOD}}$ depending on whether the `fd` or `fod` option was specified with `xtvar`. For vectors like $\mathbf{y}_{it}$ or $\mathbf{x}_{it}$, we apply the transform to each element within the vector individually.

Applying either the FD or the FOD transformation has the effect of eliminating $\mathbf{u}_i$ from (4) as follows:

$$\tilde{\mathbf{y}}_{it} = \mathbf{A}_1\tilde{\mathbf{y}}_{i,t-1} + \mathbf{A}_2\tilde{\mathbf{y}}_{i,t-2} + \cdots + \mathbf{A}_p\tilde{\mathbf{y}}_{i,t-p} + \mathbf{B}\tilde{\mathbf{x}}_{it} + \mathbf{C}\widetilde{\mathbf{w}}_{it} + \mathbf{D}\tilde{\mathbf{v}}_{it} + \tilde{\boldsymbol{\epsilon}}_{it}$$

The FOD transformation is often preferred over the FD transformation because it usually results in fewer observations being dropped because of gaps. The following data illustrate this:

	t	x	x_fd	x_fod
1.	1	1	.	.
2.	2	2	1	-1.5652476
3.	3	.	.	-.8660254
4.	4	4	.	.
5.	5	5	1	-.70710678

We have five periods `t`, and variable `x` runs from 1 to 5 but is missing when `t` = 3. Variable `x_fd` represents the FD transformation applied to `x`. Period 1 is missing because we do not have `x` at `t` = 0. `x_fd` is also missing for both periods `t` = 3 and `t` = 4 because `x` is missing at `t` = 3. Now consider variable `x_fod`, the FOD transform of `x`. Period `t` = 1 is again missing, but the gap in `x` at `t` = 3 causes us to lose only one observation for `x_fod` at `t` = 4.

Besides the greater number of observations, Hayakawa (2009) provides Monte Carlo evidence showing that the GMM estimator with the FOD transformation yields an estimate of the autoregression parameter with lower bias than the estimator with the FD transformation in the single equation case.

Constructing the instrument matrix

The second task in estimating the panel-data VAR model is to construct the instrument matrix Z_i for the transformed variables.

To simplify the exposition, suppose first that the only covariates in the model are the lags of the dependent variables. Therefore, after applying either the FOD or the FD transformation, we have

$$\tilde{\mathbf{y}}_{it} = \mathbf{A}_1\tilde{\mathbf{y}}_{i,t-1} + \mathbf{A}_2\tilde{\mathbf{y}}_{i,t-2} + \cdots + \mathbf{A}_p\tilde{\mathbf{y}}_{i,t-p} + \tilde{\boldsymbol{\epsilon}}_{it}$$

Note that each matrix of coefficients in the model is of dimension $K \times K$, and therefore there are pK^2 parameters to estimate in total.

To estimate these parameters consistently, we need to instrument the endogenous regressors with the lags of the dependent variables. Start with endogenous regressor $\tilde{\mathbf{y}}_{i,t-1}$. Given the assumptions in the model, we can instrument this endogenous regressor using $\mathbf{z}_{it} = [\mathbf{y}'_{i,t-2}, \mathbf{y}'_{i,t-3}, \ldots, \mathbf{y}'_{i1}]'$ as instruments. This implies that the moment conditions associated with this endogenous regressor are given by $E\{\text{vec}(\mathbf{z}_{it}\tilde{\boldsymbol{\epsilon}}'_{it})\} = \mathbf{0}$. Because vector $\mathbf{z}_{it}$ has $R_t = K(t-2)$ rows and the error vector $\tilde{\boldsymbol{\epsilon}}_{it}$ has K elements, we have a total of $KR_t = K^2(t-2)$ moment conditions associated with endogenous regressors $\tilde{\mathbf{y}}_{i,t-1}$. For the next endogenous regressor, $\tilde{\mathbf{y}}_{i,t-2}$, we proceed similarly. We use as instruments $\mathbf{z}_{i,t-1} = [\mathbf{y}'_{i,t-3}, \mathbf{y}'_{i,t-4}, \ldots, \mathbf{y}'_{i1}]'$ and thus obtain $KR_{t-1} = K^2(t-3)$ moment conditions. Proceeding in the same way for all the endogenous regressors, we obtain a total of $\sum_{t=p+2}^{T} KR_t = (K^2/2)(T-p-1)(T+p-2)$ moment conditions.

In tables 1a, 1b, and 1c, we show the number of parameters, instruments, and moment conditions for models with a varying number of variables, lags, and panel sizes. As you can see, even when the dataset has just 7 observations per panel, the number of moment conditions can become unwieldy, especially as the number of variables in the model increases. When the dataset has 10 or more observations per panel, tables 1a, 1b, and 1c lay bare the need to control the number of instruments.

Table 1a. Parameters, instruments, and moment conditions for $K = 2$

	$T=5$		$T=7$			$T=10$			$T=15$		
Lag order p	2	3	2	3	4	2	3	4	2	3	4
Parameters	8	12	8	12	16	8	12	16	8	12	16
Instruments	10	6	28	24	19	70	66	60	180	176	170
Moments	20	12	56	48	36	140	132	120	360	352	340

Table 1b. Parameters, instruments, and moment conditions for $K = 3$

	$T=5$		$T=7$			$T=10$			$T=15$		
Lag order p	2	3	2	3	4	2	3	4	2	3	4
Parameters	18	27	18	27	36	18	27	36	18	27	36
Instruments	15	9	42	36	27	105	99	90	270	264	255
Moments	45	27	126	108	81	315	297	270	810	792	765

Table 1c. Parameters, instruments, and moment conditions for $K = 4$

	$T=5$		$T=7$			$T=10$			$T=15$		
Lag order p	2	3	2	3	4	2	3	4	2	3	4
Parameters	32	48	32	48	64	32	48	64	32	48	64
Instruments	20	12	56	48	36	140	132	120	360	352	340
Moments	80	48	224	192	144	560	528	480	1,440	1,408	1,360

The panel-data VAR estimator is implemented at the panel level. Therefore, our next step is to consider the appropriate instrument matrix for an entire panel i.

Define $\boldsymbol{\xi}_i \equiv [\widetilde{\boldsymbol{\epsilon}}_{i,p+2}, \widetilde{\boldsymbol{\epsilon}}_{i,p+3}, \ldots, \widetilde{\boldsymbol{\epsilon}}_{iT}]'$ as the $(T-p-1) \times K$ matrix of errors for panel i. The $T-p-1$ rows of the matrix correspond to the number of periods at which we can evaluate the model after transforming our variables. The K columns correspond to the errors for each of the K equations. Thus, we can write the moment conditions for an entire panel as $E\{\text{vec}(\mathbf{Z}_i'\boldsymbol{\xi}_i)\} = \mathbf{0}$, where we define $\mathbf{Z}_i$ as the block-diagonal matrix

$$\mathbf{Z}_i = \begin{pmatrix} \mathbf{z}_{i,p+2}' & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{z}_{i,p+3}' & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{z}_{i,p+4}' & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{z}_{iT}' \end{pmatrix} \tag{5}$$

and where $\mathbf{0}$ represents a vector of 0s, the length of which is implied by its position in the matrix. Definition (5) makes $\mathbf{Z}_i$ look relatively compact, but it conceals just how many columns it contains—and hence how many moment conditions our estimator will use. We denote the number of rows of $\mathbf{Z}$ as Z_r and the number of columns as Z_c.

To get an idea of how large $\mathbf{Z}_i$ can be, let us look at $\mathbf{Z}_i$ for a bivariate panel-data VAR(2) model where we use first differencing to remove the fixed effects and further assume that $T = 6$. The first period at which we can evaluate the model is $t = 4$, so the matrix $\mathbf{Z}_i$ will have $Z_r = (6 - 4 + 1) = 3$ rows. Recalling that we used R_t to denote the number of rows in $\mathbf{z}_{it}$, we will obtain $Z_c = R_4 + R_5 + R_6 = 18$ columns. Then

$$\mathbf{Z}_i = \begin{pmatrix} \mathbf{y}_{i2}' & \mathbf{y}_{i1}' & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{y}_{i3}' & \mathbf{y}_{i2}' & \mathbf{y}_{i1}' & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{y}_{i4}' & \mathbf{y}_{i3}' & \mathbf{y}_{i2}' & \mathbf{y}_{i1}' \end{pmatrix} \tag{6}$$

With these definitions, $E\{\text{vec}(\mathbf{Z}_i'\boldsymbol{\xi}_i)\} = \mathbf{0}$ is equivalent to $E\{\text{vec}(\mathbf{z}_{it}\widetilde{\boldsymbol{\epsilon}}_{it}')\} = \mathbf{0}$ for $t = 4$, 5, and 6. There are a total of $KZ_c = 2 \times 18 = 36$ scalar moment conditions.

Returning to the general case, the number of rows of $\mathbf{Z}_i$ is $Z_r = T - p - 1$, the number of periods at which we can evaluate the model. The number of columns depends on the panel dimension T, the model's lag order p, and K, the number of variables that make up $\mathbf{y}_{it}$. This works out to be $Z_c = \sum_{t=p+2}^{T} R_t = (K/2)(T-p-1)(T+p-2)$ columns. The total number of moment conditions therefore equals $KZ_c = (K^2/2)(T-p-1)(T+p-2)$. Notice that the number of instruments is quadratic in T. Adding more observations to each panel causes the number of instruments to grow rapidly.

`xtvar` provides two ways to control the number of instruments and hence moment conditions in a panel-data VAR model. One method controls the number of lags of the dependent variables that are used as instruments (see *Restricting instrument lags*). The other method, which has come to be known as "collapsing", results in a GMM estimator that replaces the moment conditions $E\{\text{vec}(\mathbf{z}_{it}\widetilde{\boldsymbol{\epsilon}}_{it}')\} = 0$ with an alternative that averages those moment conditions over time (see *Collapsing the instrument matrix*).

Dealing with gaps and missing data

Here we show how `xtvar` creates $\mathbf{Z}_i$ when there are gaps in the data. Gaps pose two problems: first, the regression equations cannot be evaluated when either left-hand-side or right-hand-side variables are missing; and second, gaps reduce the number of available instruments. Whether the gaps are caused by some periods being absent from the data or because some variables contain missing values, the solution in both cases is essentially a strategic placement of zeros in $\mathbf{Z}_i$.

Consider data in which we have three panels and six time periods. Suppose we want to fit a VAR(1) model with two variables, y1 and y2, using these data.

	id	t	y1	y2
1.	1	1	1	2
2.	1	2	2	3
3.	1	3	3	6
4.	1	4	4	4
5.	1	5	5	7
6.	1	6	6	9
7.	2	1	1	1
8.	2	2	2	6
9.	2	4	4	7
10.	2	5	5	6
11.	2	6	6	8
12.	3	1	1	3
13.	3	2	2	5
14.	3	3	3	5
15.	3	4	4	6
16.	3	5	5	5
17.	3	6	.a	.b

The first panel is balanced; we have valid data for all six time periods. The second panel has a gap: no data are recorded for $t = 3$. Finally, the third panel contains missing values for y1 and y2 at $t = 6$.

The maximum time period in our data is $T = 6$, and we are considering a panel VAR model with one lag. Therefore, each $\mathbf{Z}_i$ will have 4 rows, corresponding to $t = 3$, 4, 5, and 6. The number of columns is $(K/2)(T - p - 1)(T + p - 2) = 1(6 - 2)(6 - 1) = 20$. $\mathbf{Z}_1$ will be just as (5) shows,

$$\mathbf{Z}_1 = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 3 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 6 & 2 & 3 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 4 & 3 & 6 & 2 & 3 & 1 & 2 \end{pmatrix}$$

Turning to the second panel, that y1 and y2 are missing for $t = 3$ implies that we will have fewer available instruments in periods $t = 5$ and $t = 6$. The missing values also affect the time periods at which we can evaluate the regression equations, and if we cannot evaluate the regression equations for a certain period, we make the corresponding row in $\mathbf{Z}_i$ contain all zeros. After applying the FOD transformation to y1 and y2, our data for the second panel look like this:

	id	t	y1	y2	fod_y1	fod_y2
1.	2	1	1	1	.	.
2.	2	2	2	6	-4.011887	-6.326437
3.	2	4	4	7	.	.
4.	2	5	5	6	-2.236068	-1.341641
5.	2	6	6	8	-1.732051	-2.886751

After taking the FOD transform, we cannot evaluate the regression equations for $t = 3$, 4, or 5. We can at least evaluate them when $t = 6$, though in that period we cannot use the values of y1 and y2 at $t = 3$ as instruments, and we must put 0s where they would go if they were not missing. We have

$$\mathbf{Z}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 7 & 0 & 0 & 2 & 6 & 1 & 1 \end{pmatrix}$$

Finally, when a gap occurs at the end of the panel, as in our third panel, we cannot evaluate the regression equations for the last period, so we put a vector of 0s in the row of $\mathbf{Z}_3$ corresponding to $t = 6$. Earlier periods, however, are not affected and so we have

$$\mathbf{Z}_3 = \begin{pmatrix} 1 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 5 & 1 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 5 & 2 & 5 & 1 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Missing values and gaps cause more problems for panel-data VARs than for many other estimators. The value at time t is used at a minimum for both the regression equation for time t and $t+1$, because we use lags as regressors; in short, a missing value for one time period could easily result in many observations being omitted from the regression equation. To make matters worse, the FOD and FD transformations can result in the loss of additional periods from our regression equation. Finally, when we have missing values and gaps, we need to consider how they affect the instruments available for later periods as well. xtvar does its best to accommodate gaps, but the nature of the estimator limits the tolerance of gaps.

Restricting instrument lags

One way to reduce the dimensionality of $\mathbf{Z}_i$ is to use a fixed number of lags to instrument each endogenous regressor, rather than using all the available lags. If we use a maximum of $l^{\max}$ lags of the dependent variables as instruments for each endogenous regressor, then the panel-level instrument matrix $\mathbf{Z}_i$ will still have $Z_r = T - p - 1$ rows. However, now the number of columns is at most $Z_c = Kl^{\max}(T - p - 1)$, which is linear in T: for each time period we add, the number of instruments increases by just $Kl^{\max}$.

Table 2a. Using two lags as instruments with $K = 2$

	$T=5$		$T=7$			$T=10$			$T=15$		
Lag order p	2	3	2	3	4	2	3	4	2	3	4
Parameters	8	12	8	12	16	8	12	16	8	12	16
Instruments	8	4	16	12	8	28	24	20	48	44	40
Moments	16	8	32	24	16	56	48	40	96	88	80

Table 2b. Using two lags as instruments with $K = 3$

	$T=5$		$T=7$			$T=10$			$T=15$		
Lag order p	2	3	2	3	4	2	3	4	2	3	4
Parameters	18	27	18	27	36	18	27	36	18	27	36
Instruments	12	6	24	18	12	42	36	30	72	66	60
Moments	36	18	72	54	36	126	108	90	216	198	180

Table 2c. Using two lags as instruments with $K = 4$

	$T=5$		$T=7$			$T=10$			$T=15$		
Lag order p	2	3	2	3	4	2	3	4	2	3	4
Parameters	32	48	32	48	64	32	48	64	32	48	64
Instruments	16	8	32	24	16	56	48	40	96	88	80
Moments	64	32	128	96	64	224	192	160	384	352	320

In tables 2a, 2b, and 2c, we show the number of parameters, instruments, and moment conditions for the same constellation of models as in tables 1a, 1b, and 1c, except that here we use just two lags of the dependent variables as instruments. We have certainly made progress limiting the proliferation of instruments. Models with three or four variables and just two or three lags still have many moment conditions in comparison with the number of parameters; but we have only one case in which the ratio exceeds 10. When $T = 5$, we can no longer fit panel-data VAR(3) models when we use $l^{\max} = 2$ because we do not have enough instruments. In these cases, though, even when using all available instruments, the models are just identified, so Hansen's J statistic would not be available to test the validity of the moment conditions. If we do limit ourselves to just two moment conditions, then the panel-data VAR(4) model is just identified when $T = 7$.

You use the `maxldep()` option with `xtvar` to limit the number of lags of $\mathbf{y}_{it}$ to use as instruments. For example, if you specify `maxldep(3)`, then `xtvar` will use $\mathbf{y}_{i,t-2}$, $\mathbf{y}_{i,t-3}$, and $\mathbf{y}_{i,t-4}$ as instruments.

Collapsing the instrument matrix

The second method provided by `xtvar` to control the number of moment conditions has come to be known as "collapsing" the instrument matrix, and it amounts to taking averages of individual moment conditions across time. You can request this method by specifying the `collapse` option with `xtvar`. Roodman (2009b) introduced the concept of collapsing in his community-contributed `xtabond2` command, and he mentions several other authors who have implemented this method of reducing the number of moment conditions.

Consider the $\mathbf{Z}_i$ matrix shown in (6) for a bivariate panel-data VAR(2) model with $T = 6$. When we collapse an instrument matrix, we first move the leading $\mathbf{0}$ vectors in each row to the end of the row, and then we remove columns consisting of all 0s. The matrix $\mathbf{Z}_i$ in that example becomes

$$\mathbf{Z}_i^C = \begin{pmatrix} \mathbf{y}'_{i2} & \mathbf{y}'_{i1} & \mathbf{0} & \mathbf{0} \\ \mathbf{y}'_{i3} & \mathbf{y}'_{i2} & \mathbf{y}'_{i1} & \mathbf{0} \\ \mathbf{y}'_{i4} & \mathbf{y}'_{i3} & \mathbf{y}'_{i2} & \mathbf{y}'_{i1} \end{pmatrix} \tag{7}$$

The number of rows still equals 3 because $t = 4$ is the first period in which we can evaluate a panel-data VAR(2) model and $T = 6$. However, now the number of columns equals $K(6-2) = 4K$ because at period $T = 6$ we have 4 available lags to use as instruments in that period.

When the "collapsing" procedure is applied to a general K-variable panel VAR(p) model with T observations per panel, the number of rows remains the same, but the number of columns reduces to $Z_c^C = K(T-p)$. As with the case where we limit the number of lags used as instruments, by collapsing we again have an instrument matrix for which the number of columns is linear in T rather than quadratic.

When we collapse the instrument matrix, we are changing the moment conditions used to identify the parameters. In the full panel-data VAR(p) model, we wrote the moment conditions as $E\{\text{vec}(\mathbf{z}_{it}\widetilde{\boldsymbol{\epsilon}}'_{it})\} = \mathbf{0}$ or, equivalently, $E(y_{i,\tau}^{(k)}\widetilde{\epsilon}_{it}^{(k')}) = 0$ for all k and k' from 1 to K and for all $\tau < t-1$. Each dependent variable lagged two periods or more is hypothesized to be uncorrelated with the time-t transformed error term.

Examining (7) in the context of the panel-level moment conditions $E(\mathbf{Z}_i^C\widetilde{\boldsymbol{\epsilon}}_t) = \mathbf{0}$ reveals that our collapsed instrument matrix implies something different. Performing the necessary multiplication, we see that the first K moment conditions can be written as $E(\sum_{t=p+2}^{T} y_{i,t-2}^{(k)}\epsilon_{i,t}^{(k')}) = 0$ for all k and k' from 1 to K. The total number of moment conditions when we use $\mathbf{Z}_i^C$ rather than $\mathbf{Z}_i$ is $KZ_c^C = K^2(T-2)$.

With the uncollapsed instrument matrix, the GMM estimator tries to make each term within that summation as close to zero as possible. Because $E(\widetilde{\boldsymbol{\epsilon}}_i) = \mathbf{0}$, our collapsed moment conditions still imply the terms within that summation have expectation zero. However, now we only ask the GMM estimator only to make sums (averages) across those terms as close to zero as possible. In that sense, the moment conditions with the collapsed instrument matrix are "weaker" because having an average of terms equal to zero is much easier than having each term equal to zero.

Adding other covariates

`xtvar` allows three types of additional covariates. Strictly exogenous regressors $\mathbf{x}_{it}$ are uncorrelated with past, present, and future realizations of the error term. Other regressors $\mathbf{w}_{it}$ may be endogenous, meaning that they are correlated with past and present realizations of the error. `xtvar` also accommodates an intermediate class of variables $\mathbf{v}_{it}$ that is predetermined, meaning that the current value of $\mathbf{v}_{it}$ is affected by errors from previous time periods but not errors from the current period.

Exogenous regressors

Letting the $M_1 \times 1$ vector $\mathbf{x}_{it}$ represent strictly exogenous variables, we can write our panel-data VAR model with such covariates as

$$\mathbf{y}_{it} = \mathbf{A}_1\mathbf{y}_{i,t-1} + \mathbf{A}_2\mathbf{y}_{i,t-2} + \cdots + \mathbf{A}_p\mathbf{y}_{i,t-p} + \mathbf{B}\mathbf{x}_{it} + \mathbf{u}_i + \boldsymbol{\epsilon}_{it} \tag{8}$$

Strictly exogenous variables are uncorrelated with past, present, or future realizations of the error term. They satisfy the condition $E\{\text{vec}(\mathbf{x}_{it}\boldsymbol{\epsilon}'_{is})\} = \mathbf{0}$ for all s and t. After applying the FD or FOD transform to remove the fixed-effect term from (8), we can use the transformed variant of $\mathbf{x}_{it}$ as its own instrument. The full-instrument matrix $\mathbf{Z}_i$ analogous to the one shown in (5) is now

$$\mathbf{Z}_i = \begin{pmatrix} \mathbf{z}'_{i,p+2} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \tilde{\mathbf{x}}'_{i,p+2} \\ \mathbf{0} & \mathbf{z}'_{i,p+3} & \mathbf{0} & \ldots & \mathbf{0} & \tilde{\mathbf{x}}'_{i,p+3} \\ \mathbf{0} & \mathbf{0} & \mathbf{z}'_{i,p+4} & \ldots & \mathbf{0} & \tilde{\mathbf{x}}'_{i,p+4} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{z}'_{iT} & \tilde{\mathbf{x}}'_{iT} \end{pmatrix} \tag{9}$$

Including $\mathbf{x}_{it}$ adds M_1 columns to $\mathbf{Z}_i$ regardless of K, p, or T.

Endogenous regressors

Endogenous regressors affect $\mathbf{y}_{it}$, and your theory implies that those variables may be correlated with the present and past realizations of the error term. These variables are represented by $M_2 \times 1$ vector $\mathbf{w}_{it}$, and we have that $E\{\text{vec}(\mathbf{w}_{it}\boldsymbol{\epsilon}'_{is})\} \neq \mathbf{0}$ for $s \leq t$. Under this definition, $\mathbf{w}_{it}$ and $\boldsymbol{\epsilon}_{it}$ may be contemporaneously correlated, and past periods' errors may affect $\mathbf{w}_{it}$ as well. Adding these variables to our model, we have

$$\mathbf{y}_{it} = \mathbf{A}_1\mathbf{y}_{i,t-1} + \mathbf{A}_2\mathbf{y}_{i,t-2} + \cdots + \mathbf{A}_p\mathbf{y}_{i,t-p} + \mathbf{B}\mathbf{x}_{it} + \mathbf{C}\mathbf{w}_{it} + \mathbf{u}_i + \boldsymbol{\epsilon}_{it}$$

Predetermined regressors

Predetermined regressors are not correlated with the present period's error term but are correlated with past realizations of the error term. We represent these regressors with the $M_3 \times 1$ vector $\mathbf{v}_{it}$, and we have that $E\{\text{vec}(\mathbf{v}_{it}\boldsymbol{\epsilon}'_{is})\} \neq \mathbf{0}$ for $s < t$. Including these variables completes our model:

$$\mathbf{y}_{it} = \mathbf{A}_1\mathbf{y}_{i,t-1} + \mathbf{A}_2\mathbf{y}_{i,t-2} + \cdots + \mathbf{A}_p\mathbf{y}_{i,t-p} + \mathbf{B}\mathbf{x}_{it} + \mathbf{C}\mathbf{w}_{it} + \mathbf{D}\mathbf{v}_{it} + \mathbf{u}_i + \boldsymbol{\epsilon}_{it}$$

By taking the first difference of this equation, you will see that the transformed regressor $\mathbf{v}_{it} - \mathbf{v}_{i,t-1}$ will be correlated with the transformed error term $\tilde{\boldsymbol{\epsilon}}_{it} = \boldsymbol{\epsilon}_{it} - \tilde{\boldsymbol{\epsilon}}_{i,t-1}$. However, $\mathbf{v}_{i,t-1}$ is a valid instrument for the transformed regressor because $\mathbf{v}_{i,t-1}$ is not correlated with $\boldsymbol{\epsilon}_{i,t-1}$. In addition, $\mathbf{v}_{i,t-2}, \mathbf{v}_{i,t-3}, \ldots$ are valid instruments. We can continue to use the definition of $\mathbf{Z}_i$ as in (5) or (9) if we again redefine $\mathbf{z}_{it}$. Now we have

$$\mathbf{z}_{it} = \left[\mathbf{y}'_{i,t-2}, \mathbf{y}'_{i,t-3}, \ldots \mathbf{y}'_{i1}, \ldots \mathbf{w}'_{i,t-2}, \mathbf{w}'_{i,t-3}, \ldots \mathbf{w}'_{i1}, \mathbf{v}'_{i,t-1}, \mathbf{v}'_{i,t-2}, \ldots \mathbf{v}'_{i1}\right]'$$

We have more lags to use as instruments for predetermined variables versus endogenous ones; each additional predetermined variable adds $(1/2)(T-p-1)(T+p)$ columns to $\mathbf{Z}_i$.

When you specify the maxldep(#) option to limit the number of lags of the dependent variables to use as instruments, the number of lags of predetermined variables used as instruments will equal $\# + 1$ because $\mathbf{v}_{i,t-1}$ is a valid instrument while $\mathbf{y}_{i,t-1}$ and $\mathbf{w}_{i,t-1}$ are not.

The number of instruments revisited

Earlier in *Constructing the instrument matrix*, we showed that for a panel-data VAR with K dependent variables and p lags and no additional covariates, the number of columns in $\mathbf{Z}_i$ was given by

$$Z_c = \text{cols}\left(\mathbf{Z}_i\right) = \frac{K}{2}\left(T-p-1\right)\left(T+p-2\right)$$

unless you use collapse, maxldep(), or minldep().

Taking stock of the full model with additional covariates, we see $\mathbf{Z}_i$ can have many columns, as shown below,

$$\begin{aligned}
\mathbf{Z}_c = & \underbrace{\frac{K}{2}\left(T-p-1\right)\left(T+p-2\right)}_{\text{Lags of } \mathbf{y}_{\text{it}}} + \underbrace{M_1}_{\text{cols}(\mathbf{x}_{\text{it}})} \\
& + \underbrace{\frac{M_2}{2}\left(T-p-1\right)\left(T+p-2\right)}_{\text{Lags of } \mathbf{w}_{\text{it}}} + \underbrace{\frac{M_3}{2}\left(T-p-1\right)\left(T+p\right)}_{\text{Lags of } \mathbf{v}_{\text{it}}} \\
= & \frac{M_2+K}{2}\left(T-p-1\right)\left(T+p-2\right) + \frac{M_3}{2}\left(T-p-1\right)\left(T+p\right) + M_1 \qquad (10)
\end{aligned}$$

and the number of moment conditions is KZ_c. When a model includes additional endogenous or predetermined variables, you likely need to specify the maxldep() or collapse option to keep the number of instruments and moment conditions manageable. If you specify the maxldep() option, then the number of columns of $\mathbf{Z}_i$ is limited to

$$Z_c = M_1 + \left\{l^{\max}\left(K+M_2+M_3\right)+M_3\right\}\left(T-p-1\right)$$

The maxldep() option reduces the growth rate of instruments from quadratic to linear, which is often the difference between being able to fit a model and not being able to fit it.

A concise representation of the GMM estimator

We now write our model at the panel level as

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\theta} + \boldsymbol{e}_i \qquad (11)$$

for regressand vector $\boldsymbol{y}_i$, regressor matrix $\boldsymbol{X}_i$, parameter vector $\boldsymbol{\theta}$, and residual vector $\boldsymbol{e}_i$. Additionally, we write the instrument matrix of the model as $\boldsymbol{Z}_i$. The definition of these matrices is given below. The model in (11) does not contain the panel-level fixed effects because all the matrices in the equation are defined in terms of transformed variables.

Writing the model as such simplifies deriving our estimator because, after all, (11) is just a linear regression equation we are to fit using GMM. Therefore, all the usual single-equation results regarding the consistency and asymptotic normality of GMM estimators apply in our case as long as our VAR process is stationary.

Before we can define the matrices used in (11), we need to create some other matrices. For each matrix composed of variables, the rows of the matrix will represent the periods in which we can evaluate our regression equation.

Define the $(T-p-1) \times K$ matrix

$$\widetilde{\boldsymbol{Y}}_i = [\tilde{\mathbf{y}}_{i,p+2}, \tilde{\mathbf{y}}_{i,p+2}, \ldots, \tilde{\mathbf{y}}_{i,T}]$$

Earlier, we had defined

$$\boldsymbol{\xi}_i \equiv [\tilde{\boldsymbol{\epsilon}}_{i,p+2}, \tilde{\boldsymbol{\epsilon}}_{i,p+3}, \ldots, \tilde{\boldsymbol{\epsilon}}_{iT}]'$$

as the $(T-p-1) \times K$ matrix of residuals, and we retain that definition here.

Define

$$\widetilde{\mathbf{X}}_i = \begin{pmatrix} \tilde{\mathbf{y}}'_{i,p+1} & \widetilde{\mathbf{y}'}_{i,p} & \cdots & \widetilde{\mathbf{y}'}_{i,2} & \widetilde{\mathbf{x}'}_{i,p+2} & \widetilde{\mathbf{w}'}_{i,p+2} & \tilde{\mathbf{v}}'_{i,p+2} \\ \tilde{\mathbf{y}}'_{i,p+2} & \widetilde{\mathbf{y}'}_{i,p+1} & \cdots & \widetilde{\mathbf{y}'}_{i,3} & \widetilde{\mathbf{x}'}_{i,p+3} & \widetilde{\mathbf{w}'}_{i,p+3} & \tilde{\mathbf{v}}'_{i,p+3} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \tilde{\mathbf{y}}'_{i,T-1} & \widetilde{\mathbf{y}'}_{i,T-2} & \cdots & \widetilde{\mathbf{y}'}_{i,T-p} & \widetilde{\mathbf{x}'}_{i,p+3} & \widetilde{\mathbf{w}'}_{i,p+3} & \tilde{\mathbf{v}}'_{i,p+3} \end{pmatrix}$$

with dimensions $(T-p-1) \times (pK+M)$, where $M = M1 + M2 + M3$.

By concatenating the individual parameter matrices horizontally, define

$$\boldsymbol{\Theta} = \left[\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_p, \mathbf{B}, \mathbf{C}, \mathbf{D}\right]$$

Matrix $\boldsymbol{\Theta}$ has dimensions $K \times (pK+M)$.

Finally, the instrument matrix $\mathbf{Z}_i$ is defined in (9) using the $\mathbf{z}_{it}$ defined in (12); its dimensions are $(T-p-1) \times Z_c$, where Z_c is the long expression in (10).

Now we can define the matrices in (11) as follows: $\boldsymbol{y}_i = \text{vec}(\widetilde{\boldsymbol{Y}}_i)$ is the $K(T-p-1) \times 1$ regressand vector; $\boldsymbol{X}_i = \widetilde{\boldsymbol{X}}_i \otimes I_K$ is the $K(T-p-1) \times K(pK+M)$ regressor matrix; $\boldsymbol{\theta} = \text{vec}(\boldsymbol{\Theta})$ is the $K(pK+M) \times 1$ vector of parameters; and $\boldsymbol{e}_i = \text{vec}(\boldsymbol{\xi}_i)$ is the $K(T-p-1) \times 1$ vector of errors. Additionally, the $K(T-p-1) \times KZ_c$ instrument matrix is given by $\boldsymbol{Z}_i = \mathbf{Z}_i \otimes I_K$.

Finally, we define

$$\boldsymbol{Q}_{Z'X} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{Z}'_i \boldsymbol{X}_i$$

and

$$\boldsymbol{Q}_{Z'y} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{Z}'_i \boldsymbol{y}_i$$

Estimators

`xtvar` estimates the parameters of model (4) using the GMM. By default, `xtvar` uses the two-step GMM estimator unless you specify the `onestep` option, in which case the one-step GMM estimator is used. The two-step estimator makes use of results from the one-step estimator, so we describe the one-step estimator first.

One-step estimator

Equation (11) is a single-equation linear regression model with endogenous covariate $\boldsymbol{X}_i$ and corresponding instrument matrix $\mathbf{Z}_i$, so the one-step estimator $\widehat{\boldsymbol{\theta}}_1$ is

$$\widehat{\boldsymbol{\theta}}_1 = \left(\boldsymbol{Q}'_{Z'X}\mathbf{W}_0\boldsymbol{Q}_{Z'X}\right)^{-1}\boldsymbol{Q}'_{Z'X}\mathbf{W}_0\boldsymbol{Q}_{Z'y}$$

where $\mathbf{W}_0$ is an initial weight matrix with dimensions $KZ_c \times KZ_c$, which we describe next.

You control the initial weight matrix $\mathbf{W}_0$ with the `winitial()` option. If you specify `winitial(identity)` then `xtvar` sets $\mathbf{W}_0$ to the identity matrix of appropriate size. We generally do not recommend using the identity matrix as the initial weight matrix for the GMM estimator, as better alternatives exist; we include it as an option here for compatibility reasons.

Alternatively, you may specify `winitial(xt)`, which creates an initial weight matrix appropriate for dynamic panel models. We have

$$\mathbf{W}_0 = N\left(\sum_i \mathbf{Z}_i'\boldsymbol{\Omega}\mathbf{Z}_i\right)^{-1} \otimes \mathbf{I}_K$$

If `fod` is specified, then

$$\boldsymbol{\Omega} = \mathbf{I}_{T-p-1}$$

If `fd` is specified, then

$$\boldsymbol{\Omega} = \begin{pmatrix} -2 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & \ddots & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & 0 & 0 \\ 0 & 0 & 0 & \ddots & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 \end{pmatrix}$$

These initial weight matrices assume that for residual vector $\boldsymbol{\epsilon}_{it}$ elements ϵ_{itj} and ϵ_{itk} are independently and identically distributed for $j \neq k$. Matrix $\boldsymbol{\Omega}$ is proportional to the covariance matrix of the transformed residual $\widetilde{\epsilon}_{it}$. (The constant of proportionality cancels out of all the formulas we use involving $\boldsymbol{\Omega}$.) The form of $\boldsymbol{\Omega}$ in the case of using first differences to remove the fixed effects follows from the fact that even though we assume $\boldsymbol{\epsilon}_{it}$ is independent across t (and i), by taking first differences, the transformed residual will follow a first-order moving average process with coefficient $\theta = -(1/2)$.

If you specify the `onestep` option, then `xtvar` reports a cluster–robust variance–covariance matrix of $\widehat{\boldsymbol{\theta}}_1$ that allows for arbitrary within-cluster correlation. We have

$$\text{Var}(\widehat{\boldsymbol{\theta}}_1) = \left(\boldsymbol{Q}'_{Z'X}\mathbf{W}_0\boldsymbol{Q}_{Z'X}\right)^{-1}\boldsymbol{Q}'_{Z'X}\mathbf{W}_0\widehat{\mathbf{S}}_1\mathbf{W}_0\boldsymbol{Q}_{Z'X}\left(\boldsymbol{Q}'_{Z'X}\mathbf{W}_0\boldsymbol{Q}_{Z'X}\right)^{-1}$$

where

$$\widehat{\mathbf{S}}_1 = \frac{1}{N}\sum_i \mathbf{Z}_i'\widehat{\mathbf{e}}_{1i}\widehat{\mathbf{e}}_{1i}'\mathbf{Z}_i \tag{12}$$

and $\widehat{\mathbf{e}}_{1i}$ are the residuals from the first step.

Two-step estimator

To compute the two-step estimator, we first obtain the one-step estimated parameter vector $\widehat{\boldsymbol{\theta}}_1$, use it to compute the residuals $\widehat{\mathbf{e}}_{1i}$, and use those to compute $\widehat{\mathbf{S}_1}$ as shown in (12). Hansen (1982) shows that to obtain the optimal GMM estimator, we should use as a weight matrix the inverse of the VCE of the moment conditions. $\widehat{\mathbf{S}_1}$ is a consistent estimator of that VCE, so we use as our second-step weight matrix $\widehat{\mathbf{W}_1} = \widehat{\mathbf{S}_1^{-1}}$. The two-step estimator is then

$$\widehat{\boldsymbol{\theta}}_2 = \left(\boldsymbol{Q}'_{Z'X}\widehat{\mathbf{W}_1}\boldsymbol{Q}_{Z'X}\right)^{-1}\boldsymbol{Q}'_{Z'X}\widehat{\mathbf{W}_1}\boldsymbol{Q}_{Z'y}$$

The default VCE is the Windmeijer (2005) robust VCE given by

$$\text{Var}(\widehat{\boldsymbol{\theta}}_2) = \frac{1}{N}V_2 + \frac{1}{N}(DV_2 + V_2D') + D\,\text{Var}(\widehat{\boldsymbol{\theta}}_1)D'$$

where

$$\begin{aligned}
V_2 &= \left(\boldsymbol{Q}'_{Z'X}\widehat{\mathbf{W}_1}\boldsymbol{Q}_{Z'X}\right)^{-1} \\
\frac{\partial S(\theta)}{\partial \theta_j} &= -\frac{1}{N}\sum_i \left(\boldsymbol{Q}^{[j]}_{Z'X}\widehat{\mathbf{e}}'_{1i}\mathbf{Z}_i + \widehat{\mathbf{e}}_{1i}\mathbf{Z}'_i\boldsymbol{Q}^{[j]'}_{Z'X}\right) \\
D^{[j]} &= -V_2\boldsymbol{Q}'_{Z'X}\widehat{\mathbf{W}_1}\frac{\partial S(\theta)}{\partial \theta_j}\widehat{\mathbf{W}_1}\bar{\mathbf{g}}\left(\widehat{\boldsymbol{\theta}}_2\right) \\
\bar{\mathbf{g}}\left(\widehat{\boldsymbol{\theta}}_2\right) &= \frac{1}{N}\sum_i \mathbf{Z}'_i\widehat{\mathbf{e}}_{2i}
\end{aligned}$$

and $[j]$ corresponds to the jth column of a matrix and $\widehat{\mathbf{e}}_{2i}$ are the residuals of the two-step estimator.

Hansen's J statistic

Hansen's J statistic is N times the GMM objective function as in

$$J = N\left(\frac{1}{N}\sum_i \mathbf{Z}'_i\widehat{\mathbf{e}}_{1i}\right)'\widehat{\mathbf{W}}\left(\frac{1}{N}\sum_i \mathbf{Z}'_i\widehat{\mathbf{e}}_{1i}\right)$$

where $\widehat{\mathbf{W}}$ is the weight matrix used, either the one-step or two-step weight matrix.

Acknowledgment

We thank Brian Poi of Poi Consulting LLC for his advice on and review of the `xtvar` command.

References

Akaike, H. 1973. "Information theory and an extension of the maximum likelihood principle". In *Second International Symposium on Information Theory*, edited by B. N. Petrov and F. Csaki, 267–281. Budapest: Akailseoniai–Kiudo.

Andrews, D. W. K., and B. Lu. 2001. Consistent model and moment selection procedures for GMM estimation with application to dynamic panel data models. *Journal of Econometrics* 101: 123–164. https://doi.org/10.1016/S0304-4076(00)00077-4.

Arellano, M., and S. Bond. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58: 277–297. https://doi.org/10.2307/2297968.

Arellano, M., and O. Bover. 1995. Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics* 68: 29–51. https://doi.org/10.1016/0304-4076(94)01642-D.

Blomquist, S., and M. Dahlberg. 1999. Small sample properties of LIML and jackknife IV estimators: Experiments with weak instruments. *Journal of Applied Econometrics* 14: 69–88. https://doi.org/10.1002/(SICI)1099-1255(199901/02)14:1<C69::AID-JAE521>3.0.CO;2-7.

Cameron, A. C., and P. K. Trivedi. 2022. *Microeconometrics Using Stata*. 2nd ed. College Station, TX: Stata Press.

Hall, A. R. 2005. *Generalized Method of Moments*. Oxford: Oxford University Press.

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

Han, C., and P. C. B. Phillips. 2006. GMM with many moment conditions. *Econometrica* 74: 147–192. https://doi.org/10.1111/j.1468-0262.2006.00652.x.

Hannan, E. J., and B. G. Quinn. 1979. The determination of the order of an autoregression. *Journal of the Royal Statistical Society*, B ser., 41: 190–195. https://doi.org/10.1111/j.2517-6161.1979.tb01072.x.

Hansen, L. P. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50: 1029–1054. https://doi.org/10.2307/1912775.

Hayakawa, K. 2009. First difference or forward orthogonal deviation—which transformation should be used in dynamic panel data models?: A simulation study. *Economics Bulletin* 29: 2008–2017.

Hayashi, F. 2000. *Econometrics*. Princeton, NJ: Princeton University Press.

Kilian, L., and H. Lütkepohl. 2017. *Structural Vector Autoregressive Analysis*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108164818.

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

Nerlove, M. 1967. Experimental evidence on the estimation of dynamic economic relations from a time series of cross-section. *Economics Studies Quarterly* 18: 42–74. https://doi.org/10.11398/economics1950.18.3_42.

Nickell, S. J. 1981. Biases in dynamic models with fixed effects. *Econometrica* 49: 1417–1426. https://doi.org/10.2307/1911408.

Roodman, D. 2009a. A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics* 71: 135–158. https://doi.org/10.1111/j.1468-0084.2008.00542.x.

———. 2009b. How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal* 9: 86–136.

Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6: 461–464. https://doi.org/10.1214/aos/1176344136.

Stock, J. H., and J. H. Wright. 2000. GMM with weak identification. *Econometrica* 68: 1055–1096. https://doi.org/10.1111/1468-0262.00151.

Windmeijer, F. 2005. A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics* 126: 25–51. https://doi.org/10.1016/j.jeconom.2004.02.005.

Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, MA: MIT Press.

Also see

[XT] **xtvar postestimation** — Postestimation tools for xtvar

[XT] **xtset** — Declare data to be panel data

[TS] **forecast** — Econometric model forecasting

[TS] **var intro** — Introduction to vector autoregressive models

[TS] **var** — Vector autoregressive models

[U] 20 Estimation and postestimation commands

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