

Estimation of Panel Vector Autoregression in Stata: a Package of Programs

Michael R.M. Abrigo and Inessa Love

(February 2015)

Abstract. Panel vector autoregression (VAR) models have been increasingly used in applied research. While programs specifically designed to estimate time-series VAR models are often included as standard features in most statistical packages, panel VAR model estimation and inference are often implemented with general-use routines that require some programming dexterity. In this paper, we briefly discuss model selection, estimation and inference of panel VAR models in a generalized method of moments (GMM) framework, and present a set of Stata programs to conveniently execute them. We illustrate the `pvar` package of programs by using standard Stata datasets.

Estimation of panel vector autoregression in Stata: A package of programs

Michael R.M. Abrigo^{*1} and Inessa Love²

(February 2015)

1. Introduction

Time-series vector autoregression (VAR) models originated in the macroeconometrics literature as an alternative to multivariate simultaneous equation models (Sims, 1980). All variables in a VAR system are typically treated as endogenous, although identifying restrictions based on theoretical models or on statistical procedures may be imposed to disentangle the impact of exogenous shocks onto the system. With the introduction of VAR in panel data settings (Holtz-Eakin, Newey and Rosen, 1988), panel VAR models have been used in multiple applications across fields.

In this paper, we give a brief overview of panel VAR model selection, estimation and inference in a generalized method of moments (GMM) framework, and provide a package of Stata programs, which we illustrate using the US National Longitudinal Survey and Lutkepohl's (1993) West Germany data. An early paper that used panel VAR in Stata was Love and Zicchino (2006), who made the programs available informally to other researchers.³ This paper introduces an updated package of programs with additional functionality, including sub-routines to implement Granger (1969) causality tests, and optimal moment and model selection following Andrews and Lu (2001).

* Corresponding author: Michael R.M. Abrigo, email: abrigomm@hawaii.edu.

¹ Graduate student, Department of Economics, University of Hawai'i at Manoa (USA) and Research specialist, Philippine Institute for Development Studies (Philippines).

² Associate Professor, Department of Economics, University of Hawai'i at Manoa (USA).

³ As of February 2015, Love and Zicchino (2006) has been cited in 445 research papers, most of which use the early version of the package of programs to estimate panel VAR models. For example, these programs have been used in studies recently published in *The American Economic Review* (Head, Lloyd-Ellis and Sun, 2014), *Applied Economics* (Mora and Logan, 2012), *Journal of Macroeconomics* (Carpenter and Demiralp, 2012) and *The Journal of Economic History* (Neumann, Fishback and Kantor, 2010), among others.

2. Panel vector autoregression

We consider a k -variate panel VAR of order p with panel-specific fixed effects represented by the following system of linear equations:

$$\begin{aligned} Y_{it} &= Y_{it-1}A_1 + Y_{it-2}A_2 + \dots + Y_{it-p+1}A_{p-1} + Y_{it-p}A_p + X_{it}B + u_{it} + e_{it} \\ i &\in \{1, 2, \dots, N\}, t \in \{1, 2, \dots, T_i\} \end{aligned} \quad (1)$$

where Y_{it} is a $(1 \times k)$ vector of dependent variables; X_{it} is a $(1 \times l)$ vector of exogenous covariates; u_i and e_{it} are $(1 \times k)$ vectors of dependent variable-specific fixed-effects and idiosyncratic errors, respectively. The $(k \times k)$ matrices $A_1, A_2, \dots, A_{p-1}, A_p$ and the $(l \times k)$ matrix B are parameters to be estimated. We assume that the innovations have the following characteristics: $E[e_{it}] = \mathbf{0}$, $E[e'_{it}e_{it}] = \Sigma$ and $E[e'_{it}e_{is}] = \mathbf{0}$ for all $t > s$.

The parameters above may be estimated jointly with the fixed effects or, alternatively, independently of the fixed effects after some transformation, using equation-by-equation ordinary least squares (OLS). With the presence of lagged dependent variables in the right-hand side of the system of equations, however, estimates would be biased even with large N (Nickell, 1981). Although the bias approaches zero as T gets larger, simulations by Judson and Owen (1999) find significant bias even when $T = 30$.

2.1. GMM estimation

Various estimators based on GMM have been proposed to calculate consistent estimates of the above equation, especially in fixed T and large N settings.⁴ With our assumption that errors are serially uncorrelated, the first-difference transformation may be consistently estimated equation-by-equation

⁴ Other methods include analytical bias correction for the least squares dummy variable model, e.g. Kiviet (1995), and Bun and Carree (2005), bias correction based on bootstrap methods, e.g. Everaert and Pozzi (2007), among others. See Canova and Ciccarelli (2013) for a survey of panel VAR models.

by instrumenting lagged differences with differences and levels of Y_{it} from earlier periods as proposed by Anderson and Hsiao (1982). This estimator, however, poses some problems. The first-difference transformation magnifies the gap in unbalanced panels. For instance, if some Y_{it-1} are not available, then the first-differences at time t and $t - 1$ are likewise missing. Also, the necessary time periods each panel is observed gets larger with the lag order of the panel VAR. As an example, for a second-order panel VAR, instruments in levels require that $T_i \geq 5$ realizations are observed for each panel.

Arellano and Bover (1995) proposed forward orthogonal deviation as an alternative transformation, which does not share the weaknesses of the first-difference transformation. Instead of using deviations from past realizations, it subtracts the average of all available future observations, thereby minimizing data loss. Potentially, only the most recent observation is not used in estimation. Since past realizations are not included in this transformation, they remain as valid instruments. For instance, in a second-order panel VAR only $T_i \geq 4$ realizations are necessary to have instruments in levels.

We can improve efficiency by including a longer set of lags as instruments. This, however, has the unattractive property of reducing observations especially with unbalanced panels or with missing observations, in general. As a remedy, Holtz-Eakin, Newey and Rosen (1988) proposed creating instruments using observed realizations, with missing observations substituted with zero, based on the standard assumption that the instrument list is uncorrelated with the errors.

While equation-by-equation GMM estimation yields consistent estimates of panel VAR, estimating the model as a system of equations may result to efficiency gains (Holtz-Eakin, Newey and Rosen, 1988).

Suppose the common set of $L \geq kp + l$ instruments is given by the row vector Z_{it} , where $X_{it} \in Z_{it}$, and equations are indexed by a number in superscript. Consider the following transformed panel VAR model based on equation (1) but represented in a more compact form:

$$Y_{it}^* = \overline{Y_{it}^*}A + e_{it}^* \quad (2)$$

$$\begin{aligned}
Y_{it}^* &= [y_{it}^{1*} \quad y_{it}^{2*} \quad \dots \quad y_{it}^{k-1*} \quad y_{it}^{k*}] \\
\overline{Y}_{it}^* &= [Y_{it-1}^* \quad Y_{it-2}^* \quad \dots \quad Y_{it-p+1}^* \quad Y_{it-p}^* \quad X_{it}^*] \\
e_{it}^* &= [e_{it}^{1*} \quad e_{it}^{2*} \quad \dots \quad e_{it}^{k-1*} \quad e_{it}^{k*}] \\
A' &= [A_1' \quad A_2' \quad \dots \quad A_{p-1}' \quad A_p' \quad B']
\end{aligned}$$

where the asterisk denotes some transformation of the original variable. If we denote the original variable as m_{it} , then the first difference transformation imply that $m_{it}^* = m_{it} - m_{it-1}$, while for the forward orthogonal deviation $m_{it} = (m_{it} - \overline{m}_{it})\sqrt{T_{it}/(T_{it} + 1)}$, where T_{it} is the number of available future observations for panel i at time t , and \overline{m}_{it} is its average.

Suppose we stack observations over panels then over time. The GMM estimator is given by

$$A = (\overline{Y}^*{}' Z \widehat{W} Z' \overline{Y}^*)^{-1} (\overline{Y}^*{}' Z \widehat{W} Z' Y^*) \quad (3)$$

where \widehat{W} is a $(L \times L)$ weighting matrix assumed to be non-singular, symmetric and positive semi-definite. Assuming that $E[Z'e] = \mathbf{0}$ and $\text{rank } E[\overline{Y}^*{}' Z] = kp + l$, the GMM estimator is consistent. The weighting matrix \widehat{W} may be selected to maximize efficiency (Hansen, 1982).⁵

Joint estimation of the system of equations makes cross-equation hypothesis testing straightforward. Wald tests about the parameters may be implemented based on the GMM estimate of A and its covariance matrix. Granger causality tests, with the hypothesis that all coefficients on the lag of variable m are jointly zero in the equation for variable n , may likewise be carried out using this test.

⁵ Roodman (2009) provides an excellent discussion of GMM estimation in a dynamic panel setting and its applications using Stata. Readers are encouraged to read his paper for a more detailed discussion of this topic.

2.2. Model Selection

Panel VAR analysis is predicated upon choosing the optimal lag order in both panel VAR specification and moment condition. Andrews and Lu (2001) proposed consistent moment and model selection criteria (MMSC) for GMM models based on Hansen's (1982) J statistic of over-identifying restrictions. Their proposed MMSC are analogous to various commonly used maximum likelihood-based model selection criteria, namely the Akaike information criteria (AIC) (Akaike, 1969), the Bayesian information criteria (BIC) (Schwarz, 1978; Rissanen, 1978; Akaike, 1977), and the Hannan-Quinn information criteria (HQIC) (Hannan and Quinn, 1979).

Applying Andrews and Lu's MMSC to the GMM estimator in (3), their proposed criteria select the pair of vectors (p, q) that minimizes

$$MMSC_{BIC,n}(k, p, q) = J_n(k^2 p, k^2 q) - (|q| - |p|)k^2 \ln n \quad (4)$$

$$MMSC_{AIC,n}(k, p, q) = J_n(k^2 p, k^2 q) - 2k^2(|q| - |p|) \quad (5)$$

$$MMSC_{HQIC,n}(p, q) = J_n(k^2 p, k^2 q) - Rk^2(|q| - |p|) \ln \ln n, \quad R > 2 \quad (6)$$

where $J_n(k, p, q)$ is the J statistic of over-identifying restriction for a k -variate panel VAR of order p and moment conditions based on q lags of the dependent variables with sample size n .

By construction, the above MMSC are available only when $q > p$. As an alternative criterion, the overall coefficient of determination (CD) may be calculated even with just-identified GMM models. Suppose we denote the $(k \times k)$ unconstrained covariance matrix of the dependent variables by Ψ . CD captures the proportion of variation explained by the panel VAR model, and may be calculated as

$$CD = 1 - \frac{\det(\Sigma)}{\det(\Psi)} \quad (7)$$

2.3. Impulse Response

Without loss of generality, we drop the exogenous variables in our notation and focus on the autoregressive structure of the panel VAR in equation (1). Lutkepohl (2005) and Hamilton (1994) both show that a VAR model is stable if all moduli of the companion matrix \bar{A} are strictly less than one, where the companion matrix is formed by

$$\bar{A} = \begin{bmatrix} A_1 & A_2 & \cdots & A_p & A_{p-1} \\ I_k & \mathbf{0}_k & \cdots & \mathbf{0}_k & \mathbf{0}_k \\ \mathbf{0}_k & I_k & \cdots & \mathbf{0}_k & \mathbf{0}_k \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0}_k & \mathbf{0}_k & \cdots & I_k & \mathbf{0}_k \end{bmatrix} \quad (8)$$

Stability implies that the panel VAR is invertible and has an infinite-order vector moving-average (VMA) representation, providing known interpretation to estimated impulse-response functions and forecast-error variance decompositions. The simple impulse-response function Φ_i may be computed by rewriting the model as an infinite vector moving-average, where Φ_i are the VMA parameters.

$$\Phi_i = \begin{cases} I_k & , \quad i = 0 \\ \sum_{j=1}^i \Phi_{t-j} A_j & , \quad i = 1, 2, \dots \end{cases} \quad (9)$$

The simple IRFs have no causal interpretation, however. Since the innovations e_{it} are correlated contemporaneously, a shock on one variable is likely to be accompanied by shocks in other variables, as well. Suppose we have a matrix P , such that $P'P = \Sigma$. Then P may be used to orthogonalize the innovations as $e_{it}P^{-1}$ and to transform the VMA parameters into the orthogonalized impulse-responses $P\Phi_i$. The matrix P effectively imposes identification restrictions on the system of dynamic equations. Sims (1980) proposed the Cholesky decomposition of Σ to impose a recursive structure on a VAR. The decomposition however is not unique, but depends on the ordering of variables in Σ .

Impulse-response function confidence intervals may be derived analytically based on the asymptotic distribution of the panel VAR parameters and the cross-equation error variance-covariance matrix. Alternatively, the confidence interval may likewise be estimated using Monte Carlo simulation, and bootstrap resampling methods.⁶

2.4. Forecast-error variance decomposition

The h -step ahead forecast-error can be expressed as

$$\mathbf{Y}_{it+h} - E[\mathbf{Y}_{it+h}] = \sum_{i=0}^{h-1} \mathbf{e}_{i(t+h-i)} \Phi_i \quad (10)$$

where \mathbf{Y}_{it+h} is the observed vector at time $t + h$ and $E[\mathbf{Y}_{it+h}]$ is the h -step ahead predicted vector made at time t . Similar to impulse-response functions, we orthogonalize the shocks using the matrix \mathbf{P} to isolate each variable's contribution to the forecast-error variance. The orthogonalized shocks $\mathbf{e}_{it}\mathbf{P}^{-1}$ have a covariance matrix \mathbf{I}_k , which allows straightforward decomposition of the forecast-error variance. More specifically, the contribution of a variable m to the h -step ahead forecast-error variance of variable n may be calculated as

$$\sum_{i=0}^{h-1} \theta_{mn}^2 = \sum_{i=1}^{h-1} (\mathbf{i}'_n \mathbf{P} \Phi_i \mathbf{i}_m)^2 \quad (11)$$

where \mathbf{i}_s is the s -th column of \mathbf{I}_k . In application, the contributions are often normalized relative to the h -step ahead forecast-error variance of variable n

$$\sum_{i=0}^{h-1} \theta_n^2 = \sum_{i=1}^{h-1} \mathbf{i}'_n \Phi_i' \Sigma \Phi_i \mathbf{i}_n \quad (12)$$

⁶ See for instance Lutkepohl (2005) for details applied in time-series VAR.

Similar to impulse-response functions, confidence intervals may be derived analytically or estimated using various resampling techniques.

3. Stata syntax

Model selection, estimation and inference about the panel vector autoregression model above can be implemented with the new Stata commands `pvar`, `pvarsoc`, `pvargranger`, `pvarstable`, `pvarirf` and `pvarfevd`. The syntax and outputs are closely patterned after Stata's built-in `var` commands for ease of use in switching between panel and time series VAR. We describe the commands' syntax in this section and provide examples in section 4.

3.1. `pvar`

`pvar` estimates panel vector autoregression models by fitting a multivariate panel regression of each dependent variable on lags of itself, lags of all other dependent variables and exogenous variables, if any. The estimation is by generalized method of moments (GMM). The command is implemented using the interactive version of Stata's `gmm` with analytic derivatives.

Syntax

```
pvar depvarlist [if] [in] [, options]
```

Options

`lags(#)` specifies the maximum lag order `#` to be included in the model. The default is to use the first lag of each variable in `depvarlist`.

`exog(varlist)` specifies a list of exogenous variables to be included in the panel VAR.

`fod` and `fd` specifies how the panel-specific fixed effects will be removed. `fod` specifies that the panel fixed-effects be removed using forward orthogonal deviation or Helmert transformation. By default, the first # lags of transformed `depvarlist` in the model are instrumented by the same lags in level (i.e. untransformed). `fod` is the default option. `fd` specifies that the panel-specific fixed effects be removed using first difference instead of forward orthogonal deviations. By default, the first # lags of transformed (i.e. differenced) `depvarlist` in the model are instrumented by the `#+1` to `2#+1` lags of `depvarlist` in levels (i.e. untransformed).

`td` subtracts from each variable in the model its cross-sectional mean before estimation. This could be used to removed time fixed effects from all the variables prior to any other transformation.

`instlags(numlist)` overrides the default lag orders of `depvarlist` used as instruments in the model (see `fod` and `fd` options above which describe which lags are used as default). Instead the `numlist`-th lags are used as instruments.

`gmmstyle` specifies that "GMM-style" instruments as proposed by Holtz-Eakin, Newey and Rosen (1988) be used. For each instrument based on lags of `depvarlist`, missing values are substituted with zero. Observations with no valid instruments are excluded. This option is available only with `instlags()`.

`gmmopts(options)` overrides the default `gmm` options run by `pvar`. Each equation in the model may be accessed individually using the variable names in `depvarlist` as equation names.

`vce(vcetype[, independent])` specifies the type of standard error reported, which includes types that are robust to some types of misspecification, that allow for intragroup correlation, and that use bootstrap or jackknife methods.

`overid` specifies that Hansen's J statistic of over-identifying restriction be reported. This option is available only for over-identified systems.

`level (#)` specifies the confidence level, as a percentage, to be used for reporting confidence intervals.

The default is `level(95)` or as set by `set level`.

`noprnt` suppresses printing of the coefficient table.

Saved Results

`pvar` saves the following in `e()`:

Scalars

<code>e(N)</code>	number of observations
<code>e(n)</code>	number of panels
<code>e(tmin)</code>	first time period in sample
<code>e(tmax)</code>	last time period in sample
<code>e(tbar)</code>	average time periods among panels
<code>e(mlag)</code>	maximum lag order in panel VAR
<code>e(N_clust)</code>	number of clusters
<code>e(Q)</code>	criterion function
<code>e(J)</code>	Hansen's J chi-squared statistic
<code>e(J_df)</code>	J statistic degrees of freedom
<code>e(rank)</code>	rank of $e(V)$
<code>e(ic)</code>	number of iterations used by iterative GMM estimator
<code>e(converged)</code>	1 if converged, 0 otherwise

Macros

<code>e(cmd)</code>	<code>pvar</code>
<code>e(cmdline)</code>	command as typed

`e(depvar)` names of dependent variables
`e(exog)` names of exogenous variables, if specified
`e(clustvar)` name of cluster variable
`e(instr)` instruments
`e(eqnames)` equation names
`e(timevar)` name of time variable
`e(panelvar)` name of panel VAR iable
`e(properties) b V`

Matrices

`e(b)` coefficient vector
`e(V)` Variance-covariance matrix of the estimator
`e(Sigma)` Variance-covariance matrix of the model residuals
`e(W)` weight matrix used for final round of estimation
`e(init)` initial values of the estimators

Functions

`e(sample)` mark estimation sample

3.2. `pvarsoc`

`pvarsoc` provides various summary measures to aid in panel VAR model selection. It reports the model overall coefficient of determination, Hansen's (1982) J statistic and corresponding p-value, and moment model selection criteria developed by Andrews and Lu (2001) based on the J statistic. Andrew and Lu's criteria are all based on Hansen's J statistic, which requires the number of moment conditions to be greater than the number of endogenous variables in the model. `pvarsoc` uses the estimation sample

of the least restrictive panel VAR model, i.e. with the highest lag order used, for all models that would be estimated by the program.

Syntax

```
pvarsoc depvarlist [if] [in] [, options]
```

Options

`maxlag(#)` specifies the maximum lag order for which the statistics are obtained.

`pinstag(numlist)` specifies that the *numlist*-th lag from the highest lag order of *depvarlist* specified in the panel VAR model implemented using `pvar` be used. This option cannot be specified with the option `pvaropts(instlag(numlist))`.

`pvaropts(options)` passes arguments to `pvar`. All arguments specified in *options* are passed to and used by `pvar` in estimation.

Saved Results

`pvarsoc` saves the following in `r()`:

Scalars

<code>r(N)</code>	number of observations
<code>r(n)</code>	number of panels
<code>r(tmin)</code>	first time period in sample
<code>r(tmax)</code>	last time period in sample
<code>r(tbar)</code>	average time periods among panels
<code>r(maxlag)</code>	maximum lag order in panel VAR

Macros

<code>r(endog)</code>	names of endogenous variables
-----------------------	-------------------------------

`r(exog)` names of exogenous variables, if specified

Matrices

`r(stats)` CD, J and p-value, MBIC, MAIC, and MQIC

3.3. `pvargranger`

The post-estimation command `pvargranger` performs Granger causality Wald tests for each equation of the underlying panel VAR model. It provides a convenient alternative to Stata's built-in `test` command.

Syntax

```
pvargranger [, estimates(estname)]
```

Options

`estimates(estname)` requests that `pvargranger` use the previously obtained set of panel VAR estimates saved as `estname`. By default, `pvargranger` uses the active (i.e. the latest) results.

Saved Results

`pvargranger` saves the following in `r()`:

Matrix

`r(pgstats)` chi-squared, degrees of freedom, and p-values

3.4. `pvarstable`

The post-estimation command `pvarstable` checks the stability condition of panel VAR estimates by calculating the modulus of each eigenvalue of the estimated model. Lutkepohl (2005) and Hamilton (1994) both show that a VAR model is stable if all moduli of the companion matrix are strictly less than one. Stability implies that the panel VAR is invertible and has an infinite-order vector moving-average

representation, providing known interpretation to estimated impulse-response functions and forecast-error variance decompositions.

Syntax

```
pvarstable [, options]
```

Options

`estimates(estname)` requests that `pvarstable` use the previously obtained set of `pvar` estimates saved in `estname`. By default, `pvarstable` uses the active estimation results.

`graph` requests `pvarstable` to draw a graph of the eigenvalue of the companion matrix.

`nogrid` suppresses the polar grid circles on the plotted eigenvalues.

Saved Results

`pvarstable` saves the following in `r()`:

Matrices

<code>r(Re)</code>	real part of the eigenvalues of the companion matrix
<code>r(Im)</code>	imaginary part of the eigenvalues of the companion matrix
<code>r(Modulus)</code>	modulus of the eigenvalues of the companion matrix

3.5. pvarirf

The post-estimation command `pvarirf` calculates and plots impulse-response functions (IRF).

Confidence bands are estimated using Gaussian approximation based on Monte Carlo draws from the estimated panel VAR model. Orthogonalized IRF are based on Cholesky decomposition, and cumulative IRF may be also computed using `pvarirf`.

Syntax

`pvarirf [, options]`

Options

`step(#)` specifies the step (forecast) horizon; the default is ten periods.

`impulse(impulsevars)` and `response(endogvars)` specify the impulse and response variables.

Usually one of each is specified, and one graph is drawn. If multiple variables are specified, a separate subgraph is drawn for each impulse-response combination. If `impulse()` and `response()` are not specified, subgraphs are drawn for all combinations of impulse and response variables.

`porder(varlist)` specifies the Cholesky ordering of the endogenous variables to be used when estimating orthogonalized IRFs, as well as the order of the IRF plots. By default, the order in which the variables were originally specified on the `pvar` command is used. This allows a new set of IRFs with a different order to be produced without re-estimating the system.

`oirf` requests that orthogonalized IRFs be estimated. The default is simple IRFs.

`cumulative` computes cumulative impulse response functions. This option may be combined with `oirf`.

`mc(#)` requests that # Monte Carlo draws be used to estimate the confidence intervals of the IRFs using Gaussian approximation. The default is not to plot confidence intervals, i.e. # = 0.

`table` displays the calculated impulse-response functions as a table. The default is not to tabulate IRFs.

`level(#)` specifies the confidence level, as a percentage, to be used for computing confidence bands.

The default is `level(95)` or as set by `set level`. `level` is available only when `mc(#)` > 1 is specified.

`dots` requests the display of iteration dots. By default, one dot character is displayed for each iteration.

A red 'x' is displayed if the iteration returns an error.

`save(filename)` specifies that the calculated IRFs be saved under the name `filename`.

`byoption(by_option)` affects how the subgraphs are combined, labeled, etc. This option is

documented in [G] `by_option`.

`nodraw` suppresses the display of the estimated IRFs.

Saved Results

`pvarirf` saves the following in `r()`:

Scalars

`r(iter)` Monte Carlo iterations

`r(step)` forecast horizon

Macros

`r(porder)` Cholesky order of orthogonalized IRF

3.6. `pvarfevd`

The post-estimation command `pvarfevd` computes forecast-error variance decomposition (FEVD) based on a Cholesky decomposition of the residual covariance matrix of the underlying panel VAR model. Standard errors and confidence intervals based on Monte Carlo simulation may be optionally computed.

Caution in interpreting computed FEVD should be exercised when exogenous variables are included in the underlying panel VAR model. Contributions of exogenous variables, when included in the panel VAR model, to forecast-error variance are disregarded in calculating FEVD.

Syntax

`pvarfevd [, options]`

Options

`step(#)` specifies the step (forecast) horizon; the default is ten periods.

`impulse(impulsevars)` and `response(responsevars)` specify the impulse and response variables for which forecast-error variance decomposition are to be reported. If `impulse()` or `response()` is not specified, each endogenous variable, in turn, is used.

`porder(varlist)` specifies the Cholesky ordering of the endogenous variables to be used when estimating FEVDs. By default, the order in which the variables were originally specified on the underlying `pvar` command is used.

`mc(#)` requests that # Monte Carlo draws be used to estimate the standard errors and the percentile-based 90% confidence intervals of the FEVDs. Computed standard errors and confidence intervals are not displayed, but may be saved as a separate file.

`dots` requests the display of iteration dots. By default, one dot character is displayed for each iteration. A red 'x' is displayed if the iteration returns an error.

`save(filename)` specifies that the FEVDs be saved under the name `filename`. In addition, standard errors and percentile-based 90% confidence intervals are saved when `#>1` in option `mc(#)` is specified.

`notable` requests the table to be constructed but not displayed.

Saved Results

`pvarfevd` saves the following in `r()`:

Scalars

<code>r(iter)</code>	Monte Carlo iterations
<code>r(step)</code>	forecast horizon

Macros

<code>r(porder)</code>	Cholesky order
------------------------	----------------

4. Examples

We illustrate the use of the `pvar` suite of commands by analyzing the relationship between annual hours worked and hourly earnings, which has been previously analyzed by Holtz-Eakin, Newey and Rosen (1988) on their seminal paper on panel vector autoregression. To compare our new programs with Stata's built-in `var` suite of commands, we also applied the new `pvar` suite of commands to Lutkepohl's (1993) West Germany time series data.

4.1. National Longitudinal Survey data

We use the subsample of women aged 14-26 years in 1968 from the National Longitudinal Surveys of 1968 to 1978 available from Stata. Our subsample consists of 2,039 women who had reported wages (`wage`) and annual hours worked (`hours`) in at least three rounds of the survey, of which two are in consecutive years. Holtz-Eakin, et. al. used the same survey but with a different time period and different subsample of workers, thus results may not be directly comparable.

Model selection measures calculated using `pvarsoc` for first- to third-order panel VARs using the first four lags of `hours` and `wage` as instruments is shown below.

```

. webuse nlswork2
(National Longitudinal Survey. Young Women 14-26 years of age in 1968)

. gen wage = exp(ln_wage)

. pvarsoc wage hours, maxlag(3) pvaropts(instl(1/4))
Running panel VAR lag order selection on estimation sample
...

```

```

Selection order criteria
Sample: 72 - 73
No. of obs = 838
No. of panels = 506
Ave. no. of T = 1.656

```

lag	CD	J	J pvalue	MBIC	MAIC	MQIC
1	.9906918	11.74496	.4663737	-69.02726	-12.25504	-34.01648
2	.988392	5.395145	.7146273	-48.453	-10.60486	-25.11248
3	.9862297	3.624628	.4591831	-23.29944	-4.375372	-11.62918

Based on the three model selection criteria by Andrews and Lu (2001) and the over-all coefficient of determination, first-order panel VAR is the preferred model, since this has the smallest MBIC, MAIC and MQIC. While we also want to minimize Hansen's J statistic, it does not correct for the degrees of freedom in the model like the model and moment selection criteria by Andrews and Lu. Based on the selection criteria, we fit a first-order panel VAR model with the same specification of instruments as above using GMM estimation implemented by `pvar`.

```

. pvar wage hours, instl(1/4)
Panel vector autoregression

```

GMM Estimation

```

Final GMM Criterion Q(b) = .014
Initial weight matrix: Identity
GMM weight matrix: Robust

```

```

No. of obs = 838
No. of panels = 506
Ave. no. of T = 1.656

```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
wage						
wage L1.	.6428702	.0978213	6.57	0.000	.4511439	.8345965
hours L1.	.0170489	.0176144	0.97	0.333	-.0174747	.0515725
hours						
wage L1.	-.0575627	.5706831	-0.10	0.920	-1.176081	1.060956
hours L1.	.5834965	.1436703	4.06	0.000	.3019079	.865085

```

Instruments : 1(1/4).(wage hours)

```

Note that the 506 women included in the estimation is significantly less than the full subsample of women available in the data. By default, `pvar` drops from estimation any observation with missing data. Since hours worked and wages are not observed in all years for all women in the subsample the number of observations dropped grows with the lag order of variables included as instruments. We can improve estimation by using “GMM-style” instruments as proposed by Holtz-Eakin, et. al. Instrument lags with missing values are replaced with zeroes. This increases the estimation sample, which results to more efficient estimates.

```
. pvar wage hours, instl(1/4) gmmstyle
Panel vector autoregression

GMM Estimation
Final GMM Criterion Q(b) = .00792
Initial weight matrix: Identity
GMM weight matrix: Robust
No. of obs = 5257
No. of panels = 2039
Ave. no. of T = 2.578
```

		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
wage	wage L1.	.8062972	.079843	10.10	0.000	.6498078	.9627867
	hours L1.	.0721378	.0248413	2.90	0.004	.0234498	.1208257
hours	wage L1.	-2.280437	.3250711	-7.02	0.000	-2.917565	-1.643309
	hours L1.	-.1068443	.0947648	-1.13	0.260	-.2925799	.0788912

Instruments : 1(1/4).(wage hours)

Although Granger causality for a first-order panel VAR may be inferred from the `pvar` output above, we still perform the test using `pvargranger` as an illustration. Results of the Granger causality tests below show that wage Granger-causes hours, and hours Granger-causes wage at the usual confidence levels, similar to the findings by Holtz-Eakin, et.al.

```

. pvargranger
panel VAR-Granger causality wald test
Ho: Excluded variable does not Granger-cause Equation variable
Ha: Excluded variable Granger-causes Equation variable

```

Equation \ Excluded		chi2	df	Prob > chi2
wage	hours	8.433	1	0.004
	ALL	8.433	1	0.004
hours	wage	49.213	1	0.000
	ALL	49.213	1	0.000

Panel vector autoregression model estimates are seldom interpreted by its self. In practice, researchers are often interested in the impact of exogenous changes in each endogenous variable to other variables in the panel VAR system. Prior to estimating impulse-response functions (IRF) and forecast-error variance decompositions (FEVD), however, we first check the stability condition of the estimated panel VAR. The resulting table and graph of eigenvalues confirms that the estimate is stable.

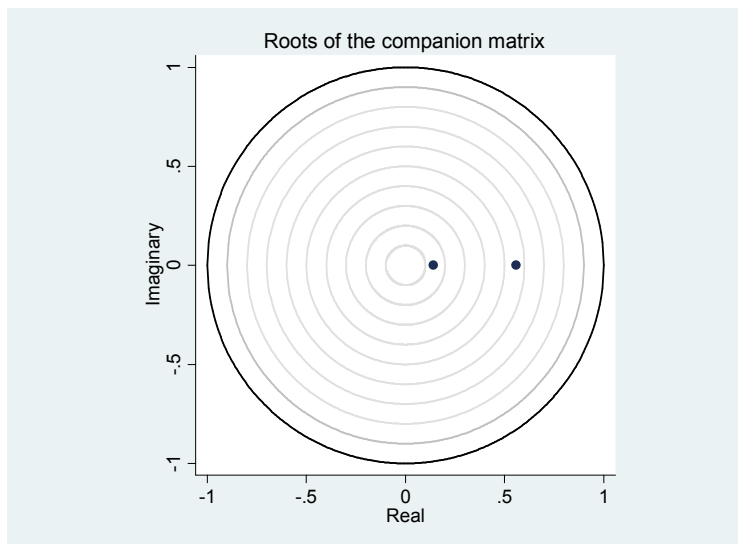
```

. pvarstable, graph
Eigenvalue stability condition

```

Eigenvalue		Modulus
Real	Imaginary	
.559372	0	.559372
.1400809	0	.1400809

All the eigenvalues lie inside the unit circle.
pVAR satisfies stability condition.



Following the theoretical exposition by Holtz-Eakin, et. al., we argue that shocks in wage levels have direct impact on contemporaneous hours worked, while current work effort affects wages only in the future. Using this causal ordering, we calculated the implied IRF using `pvarirf` and the implied FEVD using `pvarfevd`. The IRF confidence intervals are computed using 200 Monte Carlo draws based on the estimated model. Standard errors and confidence intervals for the FEVD estimates are likewise available but not shown here to save on space.

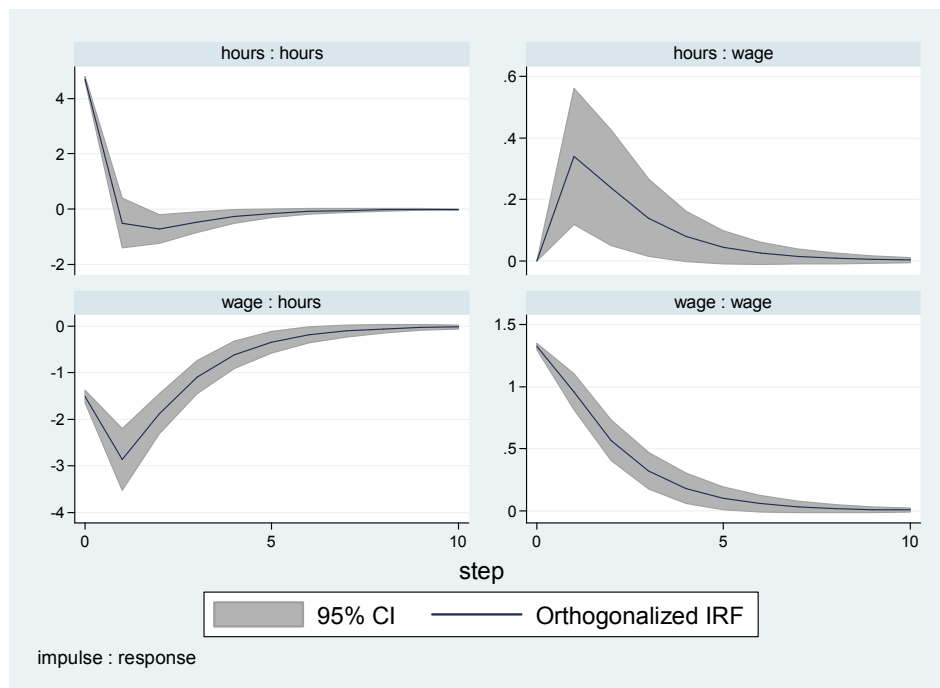
```
. pvarfevd, mc(200) save("fevd.dta")
note: label truncated to 80 characters
```

Forecast-error variance decomposition

Response variable and Forecast horizon	Impulse variable	
	wage	hours
wage		
0	0	0
1	1	0
2	.9587072	.0412928
3	.945842	.054158
4	.9419529	.0580471
5	.9407574	.0592426
6	.9403861	.0596139
7	.9402702	.0597298
8	.9402341	.059766
9	.9402227	.0597773
10	.9402192	.0597809
hours		
0	0	0
1	.0933682	.9066318
2	.3183373	.6816627
3	.3792638	.6207362
4	.3964525	.6035476
5	.4016319	.598368
6	.4032323	.5967677
7	.403731	.596269
8	.4038868	.5961132
9	.4039355	.5960644
10	.4039508	.5960492

FEVD standard errors and confidence intervals based on 200 Monte Carlo simulations are saved in file `fevd.dta`

```
. pvarirf, mc(200) oirf byopt(yrescale)
```



Based on the FEVD estimates, we see that as much as 40 percent of variation in hours worked by women in our example can be explained by their wages. On the other hand, hours worked explain only five percent of variation in future wages of women. In terms of levels, the IRF plot shows that a positive shock on real wages leads to a decreased work effort, which implies a backward bending labor supply among women in the sample. It is also noteworthy that current shock in work effort have positive yet short-lived impacts on both hours worked and wages. The effect of current shock on wages, on the other hand, has a persistent positive impact on future wages.

4.2. Lutkepohl (1993) West Germany data

We compare our `pvar` command with the built-in Stata `var` command using Lutkepohl's West Germany time series data available from Stata. The data contains first difference of the natural log of investment (`dln_inv`), of income (`dln_inc`), and of consumption (`dln_consump`) from the second quarter of 1962 up to the fourth quarter of 1982. Lutkepohl uses only observations up to the fourth quarter of 1978 in his

examples, but we use the full sample in our exposition here. We set the time-series data as a one-panel data for `pvar` to function.

```
. webuse lutkepoh12
(Quarterly SA West German macro data, Bil DM, from Lutkepoh1 1993 Table E.1)
. gen id = 1
. xtset id qtr
    panel variable: id (strongly balanced)
    time variable: qtr, 1960q1 to 1982q4
                   delta: 1 quarter
```

For simplicity, we compare VAR(1) estimates using the built-in Stata `var` (denoted `var1` in the output below), and two specifications of the new Stata command `pvar`: 1) using the default options with a one-lag instruments (`pvar1_1`) and 2) using a five-lags “GMM style” instrument set (`pvar1_5`). The VAR/panel VAR point estimates are summarized as a table below. Based on the point estimates and standard errors calculated, notice that each coefficient’s 95 percent confidence interval, i.e. roughly two standard errors on either side of the point estimate, overlap across estimators. Also, `pvar` uses one less observation than `var` because of the forward orthogonal transformation.

```

. qui var dln_inv dln_inc dln_consump, lags(1)
.   est store var1
. qui pvar dln_inv dln_inc dln_consump, lags(1)
.   est store pvar1_1
. qui pvar dln_inv dln_inc dln_consump, lags(1) instl(1/5) gmms
.   est store pvar1_5
. est table var1 pvar1_1 pvar1_5, se stat(N tmin tmax) drop(_cons)

```

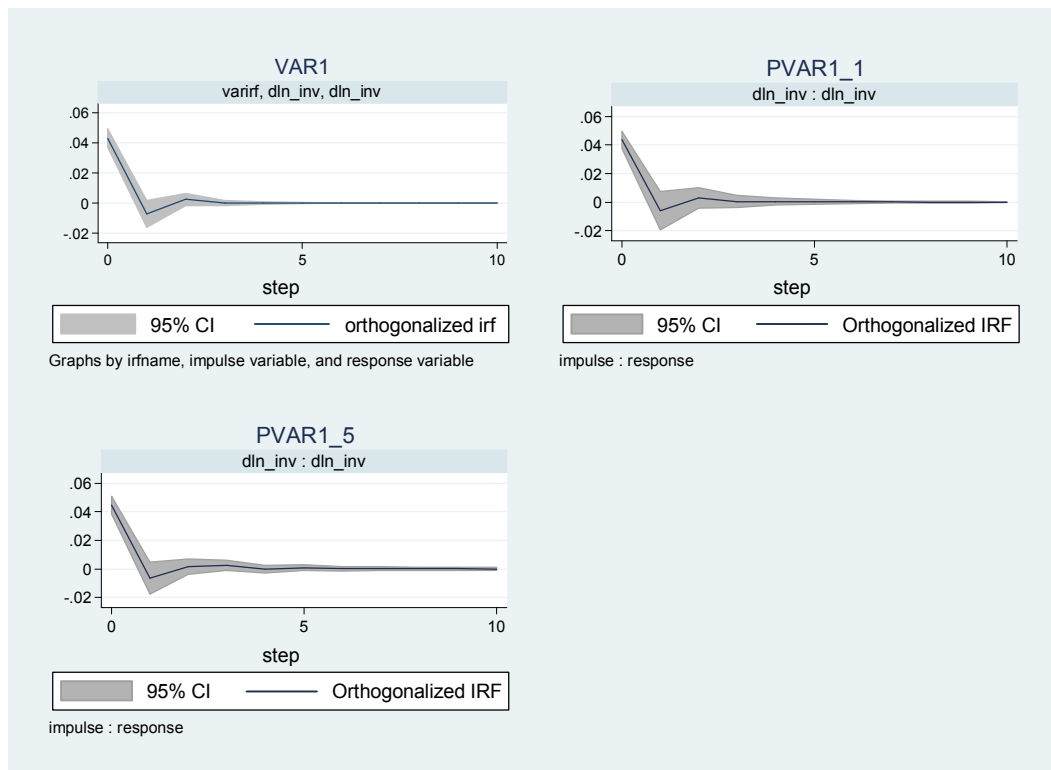
variable	var1	pvar1_1	pvar1_5
dln_inv			
dln_inv			
L1.	-.22185123 .10653512	-.21273369 .17654561	-.26849009 .11342286
dln_inc			
L1.	.41510931 .44122824	.56587455 .45532768	.08242313 .24411809
dln_consump			
L1.	.57644709 .5011716	.72224157 .52570264	1.6836959 .36813698
dln_inc			
dln_inv			
L1.	.0340018 .02821766	.02621551 .02586756	.0209324 .01613904
dln_inc			
L1.	-.00827691 .11686691	.12441837 .16229116	.47762624 .12738437
dln_consump			
L1.	.23605129 .13274394	.41862255 .14565136	.15905599 .12359724
dln_consump			
dln_inv			
L1.	-.00139181 .02590542	-.00633585 .02349897	-.00875257 .01656548
dln_inc			
L1.	.30784668 .10729047	.41644295 .14926382	.7993262 .13852001
dln_consump			
L1.	-.20676573 .12186649	-.03995971 .1430883	-.21808259 .13304683
Statistics			
N	90	89	89
tmin	2	2	2
tmax	91	90	90

Legend: b/se

Cholesky impulse-response functions and forecast-error variance decompositions can likewise be estimated using the new Stata commands `pvarirf` and `pvarfevd`, but are not shown in the interest of space.⁷ Similar to the VAR/panel VAR point estimates, the 95 percent confidence intervals of the

⁷ This paper is accompanied with a *.do file which details all commands used in this paper.

Cholesky IRFs and FEVDs overlap for the three estimators. Below, we show the response of `dln_inv` to a one standard deviation shock on `dln_inv` using the three models.



5. References

- Akaike, H. (1969). Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics*, 21, 243-247.
- Akaike, H. (1977). On entropy maximization principle. In: Krishnaiah, P.R. (Ed.), *Applications of Statistics*. Amsterdam: North-Holland.
- 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(1), 123-164.

- Arellano, M. and O. Bover (1995). Another look at the instrumental variable estimation of error-components model. *Journal of Econometrics*, 68(1), 29-51.
- Bun, M.J.G. and M.A. Carree (2005). Bias-corrected estimation in dynamic panel data models. *Journal of Business & Economic Statistics*, 23(2), 200-210.
- Canova, F. and M. Ciccarelli (2013). Panel vector autoregressive models: A survey. *Advance in Econometrics*, 32,
- Carpenter, S. and S. Demiralp (2012). Money, reserves, and the transmission of monetary policy: Does the money multiplier exist? *Journal of Macroeconomics*, 34(1), 59-75.
- Everaert, G. and L. Pozzi (2007). Bootstrap-based correction for dynamic panels. *Journal of Economic Dynamics and Control*, 31(4), 1160-1184.
- Granger, C.W.J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3), 424-438.
- Hamilton, J.D. (1994). *Time Series Analysis*. Princeton: Princeton University Press.
- Hannan, E.J. and B.G. Quinn (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society, Series B*, 41(2), 190-195.
- Hansen, L.P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4), 1029-1054.
- Head, H. H. Lloyd-Ellis and H. Sun (2015). Search, liquidity, and the dynamics of house prices and construction. *The American Economic Review*, 104(4), 1172-1210.
- Holtz-Eakin, D., W. Newey and H.S. Rosen (1988). Estimating vector autoregressions with panel data. *Econometrica*, 56(6), 1371-1395.

- Judson, R.A., and A.L. Owen. 1999. Estimating dynamic panel data models: A guide for macroeconomists. *Economics Letters*, 65(1), 9-15.
- Kiviet, J.F. (1995). On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of Econometrics*, 68(1), 53-78.
- Love, I. and L. Zicchino (2006). Financial development and dynamic investment behavior: Evidence from panel VAR. *The Quarterly Review of Economics and Finance*, 46(2), 190-210.
- Lutkepohl, H. (1993). Introduction to Multiple Time Series Analysis, 2nd Ed. New York: Springer.
- Lutkepohl, H. (2005). New Introduction to Multiple Time Series Analysis. New York: Springer.
- Mora, N. and A. Logan (2012). Shocks to bank capital: Evidence from UK banks at home and away. *Applied Economics*, 44(9), 1103-1119.
- Neumann, T.C., P.V. Fishback and S. Kantor (2010). The dynamics of relief spending and the private urban market during the New Deal. *The Journal of Economic History*, 70(1) 195-220.
- Nickell, S.J. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6), 1417-1426.
- Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, 14(5), 465-471.
- Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal*, 9(1), 86-139.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2), 461-464.
- Sims, C.A. (1980). Macroeconomics and reality. *Econometrica*, 48(1), 1-48.