xtdpd — Linear dynamic panel-data estimation

**Description**

_xtdpd_ fits a linear dynamic panel-data model where the unobserved panel-level effects are correlated with the lags of the dependent variable. The command can fit Arellano–Bond and Arellano–Bover/Blundell–Bond models like those fit by _xtabond_ and _xtdpdsys_. However, it also allows models with low-order moving-average correlation in the idiosyncratic errors or predetermined variables with a more complicated structure than allowed for _xtabond_ or _xtdpdsys_.

**Quick start**

Arellano–Bond model of y on L.y and x with the first difference of x as an instrument for the difference equation using _xtset_ data

```
xtdpd y L.y x, div(x) dgmmiv(y)
```

Add the first difference of the lag of x as an instrument for the level equation

```
xtdpd y L.y x, div(x) dgmmiv(y) lgmmiv(x)
```

Use lags 3 to 5 of x as instruments for the difference equation

```
xtdpd y L.y x, div(x) dgmmiv(y, lagrange(3 5))
```

**Menu**

Statistics > Longitudinal/panel data > Dynamic panel data (DPD) > Linear DPD estimation
Syntax

```
xtdpd  depvar  [ indepvars ]  [ if ]  [ in ],  dgmmiv(varlist [...] )  [ options ]
```

```
Model
  *dgmmiv(varlist [...] )  GMM-type instruments for the difference equation;
    can be specified more than once
  lgmmiv(varlist [...] )  GMM-type instruments for the level equation;
    can be specified more than once
  iv(varlist [...] )  standard instruments for the difference and level equations;
    can be specified more than once
  div(varlist [...] )  standard instruments for the difference equation only;
    can be specified more than once
  liv(varlist )  standard instruments for the level equation only;
    can be specified more than once
  noconstant  suppress constant term
  twostep  compute the two-step estimator instead of the one-step estimator
  hascons  check for collinearity only among levels of independent variables;
    by default checks occur among levels and differences
  fodeviation  use forward-orthogonal deviations instead of first differences
```

```
SE/Robust
  vce(vcetype)  vcetype may be gmm or robust
```

```
Reporting
  level(#)  set confidence level; default is level(95)
  artests(#)  use # as maximum order for AR tests; default is artests(2)

display_options  control spacing and line width
```

```
coeflegend  display legend instead of statistics
```

---

*dgmmiv() is required.
A panel variable and a time variable must be specified; use `xtset'; see [XT] xttset.
`depvar', `indepvars', and all `varlists' may contain time-series operators; see [U] 11.4.4 Time-series varlists.
by, `statsby', and `xi' 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

```
dgmmiv(varlist [, lagrange(flag [ llag ] ) ])  specifies GMM-type instruments for the difference equation. Levels of the variables are used to form GMM-type instruments for the difference equation. All possible lags are used, unless `lagrange(flag llag )' restricts the lags to begin with `flag' and end with `llag'. You may specify as many sets of GMM-type instruments for the difference equation as you need within the standard Stata limits on matrix size. Each set may have its own `flag' and `llag'. `dgmmiv()' is required.

lgmmiv(varlist [, lag(# ) ])  specifies GMM-type instruments for the level equation. Differences of the variables are used to form GMM-type instruments for the level equation. The first lag of the
differences is used unless `lag(#)` is specified, indicating that #th lag of the differences be used. You may specify as many sets of GMM-type instruments for the level equation as you need within the standard Stata limits on matrix size. Each set may have its own `lag`.

`iv(varlist [, nodifference])` specifies standard instruments for both the differenced and level equations. Differences of the variables are used as instruments for the differenced equations, unless `nodifference` is specified, which requests that levels be used. Levels of the variables are used as instruments for the level equations. You may specify as many sets of standard instruments for both the differenced and level equations as you need within the standard Stata limits on matrix size.

`div(varlist [, nodifference])` specifies additional standard instruments for the difference equation. Specified variables may not be included in `iv()` or in `liv()`. Differences of the variables are used, unless `nodifference` is specified, which requests that levels of the variables be used as instruments for the difference equation. You may specify as many additional sets of standard instruments for the difference equation as you need within the standard Stata limits on matrix size.

`liv(varlist)` specifies additional standard instruments for the level equation. Specified variables may not be included in `iv()` or in `div()`. Levels of the variables are used as instruments for the level equation. You may specify as many additional sets of standard instruments for the level equation as you need within the standard Stata limits on matrix size.

`noconstant`; see [R] Estimation options.

`twostep` specifies that the two-step estimator be calculated.

`hascons` specifies that xtdpd check for collinearity only among levels of independent variables; by default checks occur among levels and differences.

`fodeviation` specifies that forward-orthogonal deviations be used instead of first differences. `fodeviation` is not allowed when there are gaps in the data or when `lgmmiv()` is specified.

`vce(vcetype)` specifies the type of standard error reported, which includes types that are derived from asymptotic theory and that are robust to some kinds of misspecification; see Methods and formulas.

`vce(gmm)`, the default, uses the conventionally derived variance estimator for generalized method of moments estimation.

`vce(robust)` uses the robust estimator. For the one-step estimator, this is the Arellano–Bond robust VCE estimator. For the two-step estimator, this is the Windmeijer (2005) WC-robust estimator.

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

`artests(#)` specifies the maximum order of the autocorrelation test to be calculated. The tests are reported by `estat abond`; see [XT] xtdpd postestimation. Specifying the order of the highest test at estimation time is more efficient than specifying it to `estat abond`, because `estat abond` must refit the model to obtain the test statistics. The maximum order must be less than or equal to the number of periods in the longest panel. The default is `artests(2)`.

`display_options`: `vsquish` and `nolstretch`; see [R] Estimation options.

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

`coeflegend`; see [R] Estimation options.
Remarks and examples

If you have not read [XT] xtabond and [XT] xtdpdsys, you should do so before continuing.

Linear dynamic panel-data models include p lags of the dependent variable as covariates and contain unobserved panel-level effects, fixed or random. By construction, the unobserved panel-level effects are correlated with the lagged dependent variables, making standard estimators inconsistent. xtdpd fits a dynamic panel-data model by using the Arellano–Bond (1991) or the Arellano–Bover/Blundell–Bond (1995, 1998) estimator.

Consider the dynamic panel-data model

\[ y_{it} = \sum_{j=1}^{p} \alpha_j y_{i,t-j} + \mathbf{x}_{it} \beta_1 + \mathbf{w}_{it} \beta_2 + \nu_i + \epsilon_{it} \quad i = \{1, \ldots, N\}; \quad t = \{1, \ldots, T_i\} \quad (1) \]

where

- the \( \alpha_1, \ldots, \alpha_p \) are p parameters to be estimated,
- \( \mathbf{x}_{it} \) is a 1 \( \times \) \( k_1 \) vector of strictly exogenous covariates,
- \( \beta_1 \) is a \( k_1 \times 1 \) vector of parameters to be estimated,
- \( \mathbf{w}_{it} \) is a 1 \( \times \) \( k_2 \) vector of predetermined covariates,
- \( \beta_2 \) is a \( k_2 \times 1 \) vector of parameters to be estimated,
- \( \nu_i \) are the panel-level effects (which may be correlated with \( \mathbf{x}_{it} \) or \( \mathbf{w}_{it} \)), and
- \( \epsilon_{it} \) are i.i.d. or come from a low-order moving-average process, with variance \( \sigma^2_\epsilon \).

Building on the work of Anderson and Hsiao (1981, 1982) and Holtz-Eakin, Newey, and Rosen (1988), Arellano and Bond (1991) derived one-step and two-step GMM estimators using moment conditions in which lagged levels of the dependent and predetermined variables were instruments for the difference equation. Blundell and Bond (1998) show that the lagged-level instruments in the Arellano–Bond estimator become weak as the autoregressive process becomes too persistent or the ratio of the variance of the panel-level effect \( \nu_i \) to the variance of the idiosyncratic error \( \epsilon_{it} \) becomes too large. Building on the work of Arellano and Bover (1995), Blundell and Bond (1998) proposed a system estimator that uses moment conditions in which lagged differences are used as instruments for the level equation in addition to the moment conditions of lagged levels as instruments for the difference equation. The additional moment conditions are valid only if the initial condition \( E[\nu_i \Delta y_i] = 0 \) holds for all \( i \); see Blundell and Bond (1998) and Blundell, Bond, and Windmeijer (2000).

xtdpd fits dynamic panel-data models by using the Arellano–Bond or the Arellano–Bover/Blundell–Bond system estimator. The parameters of many standard models can be more easily estimated using the Arellano–Bond estimator implemented in xtabond or the Arellano–Bover/Blundell–Bond system estimator implemented in xtdpdsys; see [XT] xtabond and [XT] xtdpdsys. xtdpd can fit more complex models at the cost of a more complicated syntax. That the idiosyncratic errors follow a low-order MA process and that the predetermined variables have a more complicated structure than accommodated by xtabond and xtdpdsys are two common reasons for using xtdpd instead of xtabond or xtdpdsys.

The standard GMM robust two-step estimator of the VCE is known to be seriously biased. Windmeijer (2005) derived a bias-corrected robust estimator for two-step VCEs from GMM estimators known as the WC-robust estimator, which is implemented in xtdpd.

The Arellano–Bond test of autocorrelation of order \( m \) and the Sargan test of overidentifying restrictions derived by Arellano and Bond (1991) are computed by xtdpd but reported by estat abond and estat sargan, respectively; see [XT] xtdpd postestimation.
Because `xtdpd` extends `xtabond` and `xtdpdsys`, [XT] `xtabond` and [XT] `xtdpdsys` provide useful background.

> **Example 1: An Arellano–Bond estimator**

Arellano and Bond (1991) apply their new estimators and test statistics to a model of dynamic labor demand that had previously been considered by Layard and Nickell (1986), using data from an unbalanced panel of firms from the United Kingdom. All variables are indexed over the firm \( i \) and time \( t \). In this dataset, \( n_{it} \) is the log of employment in firm \( i \) inside the United Kingdom at time \( t \), \( w_{it} \) is the natural log of the real product wage, \( k_{it} \) is the natural log of the gross capital stock, and \( y_{sit} \) is the natural log of industry output. The model also includes time dummies yr1980, yr1981, yr1982, yr1983, and yr1984. To gain some insight into the syntax for `xtdpd`, we reproduce the first example from [XT] `xtabond` using `xtdpd`:

```
use https://www.stata-press.com/data/r16/abdata
.xtdpd L(0/2).n L(0/1).w L(0/2).(k ys) yr1980-yr1984 year, noconstant
> div(L(0/1).w L(0/2).(k ys) yr1980-yr1984 year) dgmmiv(n)
```

Dynamic panel-data estimation

| Coef. Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|-----------------|-------|------|----------------------|
| \( n \) L1.     | .6862261 | .1486163 | 4.62 | 0.000 | .3949435 , .9775088 |
| L2.             | -.0853582 | .0444365 | -.192 | 0.055 | -.1724523 , .0017358 |
| \( w \) L1.     | -.6078208 | .0657694 | -9.24 | 0.000 | -.7367265 , -.4789151 |
| \( k \) L1.     | .3926237 | .1092374 | 3.59 | 0.000 | .1785222 , .607251 |
| \( y_{sit} \) L1. | -.7111651 | .1844599 | -3.86 | 0.000 | -1.0727 , -.3496304 |
| \( y_{sit} \) L2. | .1057969 | .1428568 | 0.74 | 0.459 | -.1741974 , .3857912 |

Instruments for differenced equation

GMM-type: L(2/.).n
Unlike most instrumental-variables estimation commands, the independent variables in the varlist are not automatically used as instruments. In this example, all the independent variables are strictly exogenous, so we include them in \texttt{div()}, a list of variables whose first differences will be instruments for the difference equation. We include the dependent variable in \texttt{dgmmiv()}, a list of variables whose lagged levels will be used to create GMM-type instruments for the difference equation. (GMM-type instruments are discussed in a technical note below.)

The footer in the output reports the instruments used. The first line indicates that \texttt{xtdpd} used lags from 2 on back to create the GMM-type instruments described in Arellano and Bond (1991) and Holtz-Eakin, Newey, and Rosen (1988). The second line says that the first difference of all the variables included in the \texttt{div()} varlist were used as standard instruments for the difference equation.

\section*{Technical note}

GMM-type instruments are built from lags of one variable. Ignoring the strictly exogenous variables for simplicity, our model is

\[
 n_{it} = \alpha_1 n_{i t-1} + \alpha_2 n_{it-2} + \nu_i + \epsilon_{it} \tag{2}
\]

After differencing we have

\[
 \Delta n_{it} = \Delta \alpha_1 n_{i t-1} + \Delta \alpha_2 n_{it-2} + \Delta \epsilon_{it} \tag{3}
\]

Equation (3) implies that we need instruments that are not correlated with either \( \epsilon_{it} \) or \( \epsilon_{it-1} \). Equation (2) shows that L2.n is the first lag of n that is not correlated with \( \epsilon_{it} \) or \( \epsilon_{it-1} \), so it is the first lag of n that can be used to instrument the difference equation.

Consider the following data from one of the complete panels in the previous example:

\begin{verbatim}
 . list id year n L2.n dl2.n if id==140

<table>
<thead>
<tr>
<th>id</th>
<th>year</th>
<th>n</th>
<th>L2. n</th>
<th>L2D. n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1023</td>
<td>140</td>
<td>1976</td>
<td>.4324315</td>
<td>.</td>
</tr>
<tr>
<td>1024</td>
<td>140</td>
<td>1977</td>
<td>.3694925</td>
<td>.</td>
</tr>
<tr>
<td>1025</td>
<td>140</td>
<td>1978</td>
<td>.3541718</td>
<td>.4324315</td>
</tr>
<tr>
<td>1026</td>
<td>140</td>
<td>1979</td>
<td>.3632532</td>
<td>.3694926</td>
</tr>
<tr>
<td>1027</td>
<td>140</td>
<td>1980</td>
<td>.3371863</td>
<td>.3541718</td>
</tr>
<tr>
<td>1028</td>
<td>140</td>
<td>1981</td>
<td>.285179</td>
<td>.3632532</td>
</tr>
<tr>
<td>1029</td>
<td>140</td>
<td>1982</td>
<td>.1756326</td>
<td>.3371863</td>
</tr>
<tr>
<td>1030</td>
<td>140</td>
<td>1983</td>
<td>.1275133</td>
<td>.285179</td>
</tr>
<tr>
<td>1031</td>
<td>140</td>
<td>1984</td>
<td>.0889263</td>
<td>.1756326</td>
</tr>
</tbody>
</table>
\end{verbatim}

The missing values in L2D.n show that we lose 3 observations because of lags and the difference that removes the panel-level effects. The first nonmissing observation occurs in 1979 and observations on n from 1976 and 1977 are available to instrument the 1979 difference equation. The table below gives the observations available to instrument the differenced equation for the data above.
The table shows that there are a total of 27 GMM-type instruments.

The output in the example above informs us that there were a total of 41 instruments applied to the difference equation. Because there are 14 standard instruments, there must have been 27 GMM-type instruments, which matches our above calculation.

Example 2: An Arellano–Bond estimator with predetermined variables

Sometimes we cannot assume strict exogeneity. Recall that a variable $x_{it}$ is said to be strictly exogenous if $E[x_{it} \epsilon_{is}] = 0$ for all $t$ and $s$. If $E[x_{it} \epsilon_{is}] \neq 0$ for $s < t$ but $E[x_{it} \epsilon_{is}] = 0$ for all $s \geq t$, the variable is said to be predetermined. Intuitively, if the error term at time $t$ has some feedback on the subsequent realizations of $x_{it}$, $x_{it}$ is a predetermined variable. In the output below, we use xtdpd to reproduce example 6 in [XT] xtabond.
. xtdpd L(0/2).n L(0/1).(w ys) L(0/2).k yr1980-yr1984 year,
> div(L(0/1).(ys yr1980-yr1984 year) dgmmiv(n) dgmmiv(L.w L2.k, lag(1 .))
> twostep noconstant vce(robust)
Dynamic panel-data estimation
Number of obs = 611
Group variable: id
Number of groups = 140
Time variable: year
Obs per group:
    min = 4
    avg = 4.364286
    max = 6
Number of instruments = 83
    Wald chi2(15) = 958.30
    Prob > chi2 = 0.0000
Two-step results
(Std. Err. adjusted for clustering on id)

|   | WC-Robust Coef. | Std. Err. | z   | P>|z| | [95% Conf. Interval] |
|---|-----------------|-----------|-----|------|----------------------|
| n | L1. 0.8580958   | 0.1265515 | 6.78| 0.000| 0.6100594 1.106132   |
|   | L2. -0.081207   | 0.0760703 | -1.07| 0.286| -0.2303022 0.0678881 |
| w | --. -0.6910855 | 0.1387684 | -4.98| 0.000| -0.9630666 -0.4191044|
|   | L1. 0.5961712   | 0.1497338 | 3.98| 0.000| 0.3026982 0.8896441  |
| ys| --. -0.6936392 | 0.1726223 | -4.01| 0.000| -0.3548354 1.032443  |
|   | L1. -0.8773678 | 0.2183085 | -4.02| 0.000| -1.305245  -0.449491 |
| k | --. 0.4140654  | 0.1382788 | 2.99| 0.003| 0.1430439 0.6850868  |
|   | L1. -0.1537048 | 0.1220244 | -1.26| 0.208| -0.3928681 0.0854586 |
|   | L2. -0.1025833 | 0.0710886 | -1.44| 0.149| -0.2419143 0.0367477 |
| yr1980| -0.0072451 | 0.017163  | -0.42| 0.673| -0.0408839 0.0263938 |
| yr1981| -0.0609608 | 0.030207  | -2.02| 0.044| -0.1201655 -0.0017561|
| yr1982| -0.1130369 | 0.045826  | -2.49| 0.013| -0.2021812 -0.0238926|
| yr1983| -0.1335249 | 0.0600213 | -2.22| 0.026| -0.2511645 -0.0158853|
| yr1984| -0.1623177 | 0.0725434 | -2.24| 0.025| -0.3045001 -0.0201352|
| year| 0.0264501  | 0.0119329 | 2.22| 0.027| 0.003062   0.0498381 |

Instruments for differenced equation
GMM-type: L(2/).n L(1/).L.w L(1/).L2.k

The footer informs us that we are now including GMM-type instruments from the first lag of L.w on back and from the first lag of L2.k on back.

Example 3: A weaker definition of predetermined variables

As discussed in [XT] xtabond and [XT] xtdpdsys, xtabond and xtdpdsys both use a strict definition of predetermined variables with lags. In the strict definition, the most recent lag of the variable in pre() is considered predetermined. (Here specifying pre(w, lag(1, .)) to xtabond means that L.w is a predetermined variable and pre(k, lag(2, .)) means that L2.k is a predetermined variable.) In a weaker definition, the current observation is considered predetermined, but subsequent
lags are included in the model. Here \( w \) and \( k \) would be predetermined instead of \( L.w \) and \( L2.w \). The output below implements this weaker definition for the previous example.

```stata
.xtdpd L(0/2).n L(0/1).(w ys) L(0/2).k yr1980-yr1984 year,
> div(L(0/1).(ys) yr1980-yr1984 year) dgmmiv(n) dgmmiv(w k, lag(1 .))
> twostep noconstant vce(robust)
```

```
Dynamic panel-data estimation
Number of obs = 611
Group variable: id
Number of groups = 140
Time variable: year
Obs per group:
    min = 4
    avg = 4.364286
    max = 6
Number of instruments = 101
Wald chi2(15) = 879.53
Prob > chi2 = 0.0000
Two-step results
(Std. Err. adjusted for clustering on id)
```

<table>
<thead>
<tr>
<th></th>
<th>WC-Robust</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
</tr>
<tr>
<td>n</td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>.6343155</td>
</tr>
<tr>
<td>L2</td>
<td>-.0871247</td>
</tr>
<tr>
<td>w</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>-.720063</td>
</tr>
<tr>
<td>L1</td>
<td>.238069</td>
</tr>
<tr>
<td>ys</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>.5999718</td>
</tr>
<tr>
<td>L1</td>
<td>-.5674808</td>
</tr>
<tr>
<td>k</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>.3931997</td>
</tr>
<tr>
<td>L1</td>
<td>-.0019641</td>
</tr>
<tr>
<td>L2</td>
<td>-.0231165</td>
</tr>
<tr>
<td>yr1980</td>
<td>-.006209</td>
</tr>
<tr>
<td>yr1981</td>
<td>-.0398491</td>
</tr>
<tr>
<td>yr1982</td>
<td>-.0525715</td>
</tr>
<tr>
<td>yr1983</td>
<td>-.0451175</td>
</tr>
<tr>
<td>yr1984</td>
<td>-.0437772</td>
</tr>
<tr>
<td>year</td>
<td>.0173374</td>
</tr>
</tbody>
</table>

Instruments for differenced equation
GMM-type: L(2/.).n L(1/.).w L(1/.).k

As expected, the output shows that the additional 18 instruments available under the weaker definition can affect the magnitudes of the estimates. Applying the stricter definition when the true model was generated by the weaker definition yielded consistent but inefficient results; there were some additional moment conditions that could have been included but were not. In contrast, applying the weaker definition when the true model was generated by the stricter definition yields inconsistent estimates.
Example 4: A system estimator of a dynamic panel-data model

Here we use xtdpd to reproduce example 2 from [XT] xtdpdsys in which we used the system estimator to fit a model with predetermined variables.

```
.xtdpd L(0/1).n L(0/2).(w k) yr1980-yr1984 year, div(yr1980-yr1984 year)
> dgmmiv(n) dgmmiv(L2.(w k), lag(1 .)) lgmmiv(n L1.(w k)) vce(robust) hascons
```

Dynamic panel-data estimation

| Number of obs | 751 |
| Group variable: id | 140 |
| Time variable: year | |

Obs per group:

| min | 5 |
| avg | 5.364286 |
| max | 7 |

Number of instruments = 95

Wald chi2(13) = 7562.80

Prob > chi2 = 0.0000

One-step results

(Std. Err. adjusted for clustering on id)

| Coef. | Std. Err. | z | P>|z| | [95% Conf. Interval] |
|-------|-----------|---|------|----------------------|
| n L1. | 0.913278 | 0.0460602 | 19.83 | 0.000 | 0.8230017 | 1.003554 |
| w --. | -0.728159 | 0.1019044 | -7.15 | 0.000 | -0.9278879 | -0.5284301 |
| L1. | 0.5602737 | 0.1939617 | 2.89 | 0.004 | 0.1801156 | 0.9404317 |
| L2. | -0.0523028 | 0.1487653 | -0.35 | 0.725 | -0.3438775 | 0.2392718 |
| k --. | 0.4820097 | 0.0760787 | 6.34 | 0.000 | 0.3328983 | 0.6311212 |
| L1. | -0.2846944 | 0.0831902 | -3.42 | 0.001 | -0.4477442 | -0.1216446 |
| L2. | -0.1394181 | 0.0405709 | -3.44 | 0.001 | -0.2189356 | -0.0599006 |
| yr1980 | -0.0325146 | 0.0216371 | -1.50 | 0.133 | -0.0749226 | 0.0100935 |
| yr1981 | -0.0726116 | 0.0346842 | -2.10 | 0.036 | -0.1405207 | -0.0047024 |
| yr1982 | -0.0477038 | 0.0451914 | -1.06 | 0.291 | -0.1362772 | 0.0408969 |
| yr1983 | -0.0396264 | 0.058734 | -0.71 | 0.478 | -0.1491362 | 0.0698835 |
| yr1984 | -0.0810383 | 0.0736648 | -1.10 | 0.271 | -0.2254186 | 0.063342 |
| year | 0.0192741 | 0.0145326 | 1.33 | 0.185 | -0.0092092 | 0.0477574 |
| _cons | -37.34972 | 28.77747 | -1.30 | 0.194 | -93.75253 | 19.05308 |

The first lags of the variables included in lgmmiv() are used to create GMM-type instruments for the level equation. Only the first lags of the variables in lgmmiv() are used because the moment conditions using higher lags are redundant; see Blundell and Bond (1998) and Blundell, Bond, and Windmeijer (2000).
Example 5: Allowing for MA(1) errors

All the previous examples have used moment conditions that are valid only if the idiosyncratic errors are i.i.d. This example shows how to use `xtdpd` to estimate the parameters of a model with first-order moving-average [MA(1)] errors using the Arellano–Bond estimator, the Arellano–Bover/Blundell–Bond system estimator, or any other consistent GMM estimator you want to specify. For simplicity, we assume that the independent variables are strictly exogenous. Also, to highlight the fact that we can specify the instrument list flexibly, we only include the levels and first lags of the exogenous variables in the instrument list. An Arellano–Bond estimator, for instance, would have included levels and first and second lags of the exogenous variables.

We begin by noting that the Sargan test rejects the null hypothesis that the overidentifying restrictions are valid in the model with i.i.d. errors.

```
. xtdpd L(0/1).n L(0/2).(w k) yr1980-yr1984 year,
> div(L(0/1).(w k) yr1980-yr1984 year) dgmmiv(n) hascons
(output omitted)
. estat sargan
Sargan test of overidentifying restrictions
  H0: overidentifying restrictions are valid
    chi2(24)  =  49.70094
    Prob > chi2 =  0.0015
```
Assuming that the idiosyncratic errors are MA(1) implies that only lags three or higher are valid instruments for the difference equation. (See the technical note below.)

```
.xtdpd L(0/1).n L(0/2).(w k) yr1980-yr1984 year,
> div(L(0/1).(w k) yr1980-yr1984 year) dgmmiv(n, lag(3 .)) hascons
```

Dynamic panel-data estimation                  Number of obs =  751  
Group variable: id                              Number of groups =  140 
Time variable: year                             Obs per group:  
                                                        min = 5   
                                                        avg = 5.364286  
                                                        max =  7   
Number of instruments =  32                    Wald chi2(13) = 1195.04  
                                                        Prob > chi2 = 0.0000  

One-step results

| n  | Coef. Std. Err. | z    | P>|z| | [95% Conf. Interval] |
|----|-----------------|------|------|---------------------|
| n  |                 |      |      |                     |
| L1.| 0.8696303       | 0.2014473 | 4.32 | 0.000 | 0.4748008 | 1.26446 |
| w  |                 |      |      |                     |
| -- | -0.5802971      | 0.0762659 | -7.61 | 0.000 | -0.7297756 | -0.4308187 |
| L1.| 0.2918658       | 0.1543883 | 1.89  | 0.059 | -0.0107296 | 0.5944613  |
| L2.| -0.5903459      | 0.2995123 | -1.97 | 0.049 | -1.177379  | -0.0033126 |
| k  |                 |      |      |                     |
| -- | 0.3428139       | 0.0447916 | 7.65  | 0.000 | 0.2550239  | 0.4306039  |
| L1.| -0.1383918      | 0.0825823 | -1.68 | 0.094 | -0.3002502 | 0.0234665  |
| L2.| -0.0260956      | 0.1535855 | -0.17 | 0.865 | -0.327177  | 0.2749265  |
| yr1980| -0.0036873    | 0.0301587 | -0.12 | 0.903 | -0.0627973 | 0.0554226  |
| yr1981| 0.00218       | 0.0592014 | 0.04  | 0.971 | -0.1138526 | 0.1182125  |
| yr1982| 0.0782939      | 0.0897622 | 0.87  | 0.383 | -0.0976367 | 0.2542246  |
| yr1983| 0.1734231      | 0.1308914 | 1.32  | 0.185 | -0.0831193 | 0.4299655  |
| yr1984| 0.2400685      | 0.1734456 | 1.38  | 0.166 | -0.0989787 | 0.5800157  |
| year| -0.0354681     | 0.0309963 | -1.14 | 0.253 | -0.0962198 | 0.0258236  |
|_cons| 73.13706       | 62.61443 | 1.17  | 0.243 | -49.58496  | 195.8591   |

Instruments for differenced equation

GMM-type: L(3/.).n

Instruments for level equation

Standard: _cons

The results from estat sargan no longer reject the null hypothesis that the overidentifying restrictions are valid.

```
.estat sargan
Sargan test of overidentifying restrictions
H0: overidentifying restrictions are valid

    chi2(18) = 20.80081
    Prob > chi2 = 0.2896
```

Moving on to the system estimator, we note that the Sargan test rejects the null hypothesis after fitting the model with i.i.d. errors.
xtdpd L(0/1).n L(0/2).(w k) yr1980-yr1984 year, 
> div(L(0/1).(w k) yr1980-yr1984 year) dgmmiv(n) lgmmiv(n) hascons
(output omitted)

. estat sargan
Sargan test of overidentifying restrictions
H0: overidentifying restrictions are valid
  chi2(31)  =  59.22907
  Prob > chi2 =  0.0017

Now we fit the model using the additional moment conditions constructed from the second lag of n as an instrument for the level equation.

. xtdpd L(0/1).n L(0/2).(w k) yr1980-yr1984 year,
> div(L(0/1).(w k) yr1980-yr1984 year) dgmmiv(n, lag(3 .)) lgmmiv(n, lag(2))
> hascons
Dynamic panel-data estimation
Number of obs = 751
Group variable: id
Number of groups = 140
Time variable: year
  Obs per group:
    min = 5
    avg = 5.364286
    max = 7
Number of instruments = 38
Wald chi2(13) = 3680.01
Prob > chi2 = 0.0000
One-step results

|      | Coef.  | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|------|--------|-----------|-------|-------|---------------------|
| n    | L1.    | .9603675  | .095608 | 10.04 | .000    | .7729794  | 1.147756 |
| w    | --.   | -.5433987 | .068835 | -7.89 | .000    | -.6783128 | -.4084845|
|      | L1.    | .4356183  | .0881727| 4.94  | .000    | .262803   | .6084336 |
|      | L2.    | -.2785721 | .1115061| -2.50 | .012    | -.4971201 | -.0600241|
| k    | --.   | .3139331  | .0419054| 7.49  | .000    | .2317999  | .3960662 |
|      | L1.    | -.160103  | .0546915| -2.93 | .003    | -.2672963 | -.0529096|
|      | L2.    | -.1295766 | .0507752| -2.55 | .011    | -.2290943 | -.030059 |
| yr1980|    | -.0200704 | .0248954| -0.81 | .420    | -.0688644 | .0287236 |
| yr1981|    | -.0425838 | .0422155| -1.01 | .313    | -.1253246 | .040157  |
| yr1982|    | .0048723  | .0600938| 0.08  | .935    | -.1129093 | .122654  |
| yr1983|    | .0458978  | .0785687| 0.58  | .559    | -.1080941 | .198897  |
| yr1984|    | .0633219  | .1026188| 0.62  | .537    | -.1378074 | .2644511 |
| year | -0.0075599 | .019059 | -0.40 | .692    | -.0449148 | .029795 |
| _cons|    | 16.20856  | 38.00619 | 0.43  | .670    | -58.28221 | 90.69932 |

Instruments for differenced equation
  GMM-type: L(3/).n
           D.yr1984 D.year

Instruments for level equation
  GMM-type: L2D.n
  Standard: _cons

The estimate of the coefficient on L.n is now 0.96. Blundell, Bond, and Windmeijer (2000, 63–65) show that the moment conditions in the system estimator remain informative as the true coefficient on L.n approaches unity. Holtz-Eakin, Newey, and Rosen (1988) show that because the large-sample
distribution of the estimator is derived for fixed number of periods and a growing number of individuals there is no “unit-root” problem.

The results from estat sargan no longer reject the null hypothesis that the overidentifying restrictions are valid.

```
estat sargan
Sargan test of overidentifying restrictions
  H0: overidentifying restrictions are valid
  chi2(24)    =  27.22585
  Prob > chi2 =  0.2940
```

- **Technical note**

  To find the valid moment conditions for the model with MA(1) errors, we begin by writing the model

  \[
  n_{it} = \alpha n_{it-1} + \beta x_{it} + \nu_i + \epsilon_{it} + \gamma \epsilon_{it-1}
  \]

  where the \( \epsilon_{it} \) are assumed to be i.i.d.

  Because the composite error, \( \epsilon_{it} + \gamma \epsilon_{it-1} \), is MA(1), only lags two or higher are valid instruments for the level equation, assuming the initial condition that \( E[\nu_i \Delta n_{i2}] = 0 \). The key to this point is that lagging the above equation two periods shows that \( \epsilon_{it-2} \) and \( \epsilon_{it-3} \) appear in the equation for \( n_{it-2} \). Because the \( \epsilon_{it} \) are i.i.d., \( n_{it-2} \) is a valid instrument for the level equation with errors \( \nu_i + \epsilon_{it} + \gamma \epsilon_{it-1} \). \( n_{it-2} \) will be correlated with \( n_{it-1} \) but uncorrelated with the errors \( \nu_i + \epsilon_{it} + \gamma \epsilon_{it-1} \). An analogous argument works for higher lags.

  First-differencing the above equation yields

  \[
  \Delta n_{it} = \alpha \Delta n_{it-1} + \beta \Delta x_{it} + \Delta \epsilon_{it} + \gamma \Delta \epsilon_{it-1}
  \]

  Because \( \epsilon_{it-2} \) is the farthest lag of \( \epsilon_{it} \) that appears in the difference equation, lags three or higher are valid instruments for the differenced composite errors. (Lagging the level equation three periods shows that only \( \epsilon_{it-3} \) and \( \epsilon_{it-4} \) appear in the equation for \( n_{it-3} \), which implies that \( n_{it-3} \) is a valid instrument for the current difference equation. An analogous argument works for higher lags.)
Stored results

`xtdpd` stores the following in `e()`:

Scalars
- `e(N)` number of observations
- `e(N_g)` number of groups
- `e(df_m)` model degrees of freedom
- `e(g_min)` smallest group size
- `e(g_avg)` average group size
- `e(g_max)` largest group size
- `e(t_min)` minimum time in sample
- `e(t_max)` maximum time in sample
- `e(chi2)` χ²
- `e(arm#)` test for autocorrelation of order #
- `e(artests)` number of AR tests computed
- `e(sig2)` estimate of σ²
- `e(rss)` sum of squared differenced residuals
- `e(sargan)` Sargan test statistic
- `e(rank)` rank of `e(V)`
- `e(zrank)` rank of instrument matrix

Macros
- `e(cmd)` `xtdpd`
- `e(cmdline)` command as typed
- `e(depvar)` name of dependent variable
- `e(twostep)` twostep, if specified
- `e(ivar)` variable denoting groups
- `e(tvar)` variable denoting time within groups
- `e(vce)` vcetype specified in `vce()`
- `e(vcetype)` title used to label Std. Err.
- `e(system)` system, if system estimator
- `e(transform)` specified transform
- `e(diffvars)` already-differenced exogenous variables
- `e(datasignature)` checksum from `datasignature`
- `e(datasignaturevars)` variables used in calculation of checksum
- `e(properties)` `b V`
- `e(estat_cmd)` program used to implement `estat`
- `e(predict)` program used to implement `predict`
- `e(marginsok)` predictions allowed by `margins`

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

Functions
- `e(sample)` marks estimation sample

Methods and formulas

Consider dynamic panel-data models of the form

\[ y_{it} = \sum_{j=1}^{p} \alpha_j y_{i,t-j} + x_{it}\beta_1 + w_{it}\beta_2 + \nu_i + \epsilon_{it} \]

where the variables are as defined as in (1).

x and w may contain lagged independent variables and time dummies.
Let $X_{it}^{L} = (y_{i,t-1}, y_{i,t-2}, \ldots, y_{i,t-p}, x_{it}, w_{it})$ be the $1 \times K$ vector of covariates for $i$ at time $t$, where $K = p + k_1 + k_2$, $p$ is the number of included lags, $k_1$ is the number of strictly exogenous variables in $x_{it}$, and $k_2$ is the number of predetermined variables in $w_{it}$. (The superscript $L$ stands for levels.)

Now rewrite this relationship as a set of $T_i$ equations for each individual, 

$$y_i^L = X_i^L \delta + \nu_i \iota_i + \epsilon_i$$

where $T_i$ is the number of observations available for individual $i$; $y_i$, $\iota_i$, and $\epsilon_i$ are $T_i \times 1$, whereas $X_i$ is $T_i \times K$.

The estimators use both the levels and a transform of the variables in the above equation. Denote the transformed variables by an $\ast$, so that $y_i^\ast$ is the transformed $y_i^L$ and $X_i^\ast$ is the transformed $X_i^L$. The transform may be either the first difference or the forward-orthogonal deviations (FOD) transform. The $(i,t)$th observation of the FOD transform of a variable $x$ is given by

$$x_{it}^\ast = c_t \left\{ x_{it} - \frac{1}{T-t} (x_{it+1} + x_{it+2} + \cdots + x_{iT}) \right\}$$

where $c_t^2 = (T-t)/(T-t+1)$ and $T$ is the number of observations on $x$; see Arellano and Bover (1995) and Arellano (2003).

Here we present the formulas for the Arellano–Bover/Blundell–Bond system estimator. The formulas for the Arellano–Bond estimator are obtained by setting the additional level matrices in the system estimator to null matrices.

Stacking the transformed and untransformed vectors of the dependent variable for a given $i$ yields

$$y_i = \begin{pmatrix} y_i^\ast \\ y_i^L \end{pmatrix}$$

Similarly, stacking the transformed and untransformed matrices of the covariates for a given $i$ yields

$$X_i = \begin{pmatrix} X_i^\ast \\ X_i^L \end{pmatrix}$$

$Z_i$ is a matrix of instruments,

$$Z_i = \begin{pmatrix} Z_{di} & 0 & D_i & 0 & I_i^d \\ 0 & Z_{Li} & 0 & L_i & I_i^L \end{pmatrix}$$

where $Z_{di}$ is the matrix of GMM-type instruments created from the $\text{dgmmiv()}$ options, $Z_{Li}$ is the matrix of GMM-type instruments created from the $\text{lgmmiv()}$ options, $D_i$ is the matrix of standard instruments created from the $\text{div()}$ options, $L_i$ is the matrix of standard instruments created from the $\text{liv()}$ options, $I_i^d$ is the matrix of standard instruments created from the $\text{iv()}$ options for the differenced errors, and $I_i^L$ is the matrix of standard instruments created from the $\text{iv()}$ options for the level errors.

$\text{div()}$, $\text{liv()}$, and $\text{iv()}$ simply add columns to instrument matrix. The GMM-type instruments are more involved. Begin by considering a simple balanced-panel example in which our model is
\[ y_{it} = \alpha_1 y_{i,t-1} + \alpha_2 y_{i,t-2} + \nu_i + \epsilon_{it} \]

We do not need to consider covariates because strictly exogenous variables are handled using div(), iv(), or liv(), and predetermined or endogenous variables are handled analogous to the dependent variable.

Assume that the data come from a balanced panel in which there are no missing values. After first-differencing the equation, we have

\[ \Delta y_{it} = \alpha_1 \Delta y_{i,t-1} + \alpha_2 \Delta y_{i,t-2} + \Delta \epsilon_{it} \]

The first 3 observations are lost to lags and differencing. If we assume that the \( \epsilon_{it} \) are not autocorrelated, for each \( i \) at \( t = 4 \), \( y_{i1} \) and \( y_{i2} \) are valid instruments for the differenced equation. Similarly, at \( t = 5 \), \( y_{i1}, y_{i2}, \) and \( y_{i3} \) are valid instruments. We specify \( \text{dgmmiv}(y) \) to obtain an instrument matrix with one row for each period that we are instrumenting:

\[
Z_{di} = \begin{pmatrix}
y_{i1} & y_{i2} & 0 & 0 & 0 & \ldots & 0 & 0 & 0 \\
0 & 0 & y_{i1} & y_{i2} & y_{i3} & \ldots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \ldots & 0 & y_{i1} & \ldots & y_{i,T-2}
\end{pmatrix}
\]

Because \( p = 2 \), \( Z_{di} \) has \( T - p - 1 \) rows and \( \sum_{m=p}^{T-2} m \) columns.

Specifying \( \text{lgmmiv}(y) \) creates the instrument matrix

\[
Z_{Li} = \begin{pmatrix}
\Delta y_{i2} & 0 & 0 & \ldots & 0 \\
0 & \Delta y_{i3} & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & \Delta y_{i(T-1)}
\end{pmatrix}
\]

This extends to other lag structures with complete data. Unbalanced data and missing observations are handled by dropping the rows for which there are no data and filling in zeros in columns where missing data are required. Suppose that, for some \( i \), the \( t = 1 \) observation was missing but was not missing for some other panels. \( \text{dgmmiv}(y) \) would then create the instrument matrix

\[
Z_{di} = \begin{pmatrix}
0 & 0 & 0 & y_{i2} & y_{i3} & 0 & 0 & 0 & \ldots & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & y_{i2} & y_{i3} & 0 & \ldots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & y_{i2} & \ldots & y_{i,T-2}
\end{pmatrix}
\]

\( Z_{di} \) has \( T_i - p - 1 \) rows and \( \sum_{m=p}^{T-2} m \) columns, where \( T = \max_i T_i \) and \( T_i \) is the number of nonmissing observations in panel \( i \).
After defining

\[ Q_{xz} = \sum_i X'_i Z_i \]

\[ Q_{zy} = \sum_i Z'_i y_i \]

\[ W_1 = Q_{xz} A_1 Q'_{xz} \]

\[ A_1 = \left( \sum_i Z'_i H_{1i} Z_i \right)^{-1} \]

and

\[ H_{1i} = \begin{pmatrix} H_{di} & 0 \\ 0 & H_{Li} \end{pmatrix} \]

the one-step estimates are given by

\[ \hat{\beta}_1 = W_1^{-1} Q_{xz} A_1 Q_{zy} \]

When using the first-difference transform \( H_{di} \), is given by

\[ H_{di} = \begin{pmatrix} 1 & -0.5 & 0 & \ldots & 0 & 0 \\ -0.5 & 1 & -0.5 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & -0.5 \\ 0 & 0 & 0 & \ldots & -0.5 & 1 \end{pmatrix} \]

and \( H_{Li} \) is given by 0.5 times the identity matrix. When using the FOD transform, both \( H_{di} \) and \( H_{Li} \) are equal to the identity matrix.

The transformed one-step residuals are given by

\[ \hat{\epsilon}_{1i} = y_i^* - \hat{\beta}_1 X_i^* \]

which are used to compute

\[ \hat{\sigma}_1^2 = \frac{1}{(N - K)} \sum_{i=1}^{N} \hat{\epsilon}_{1i}' \hat{\epsilon}_{1i}^* \]

The GMM one-step VCE is then given by

\[ \hat{V}_{GMM}[\hat{\beta}_1] = \hat{\sigma}_1^2 W_1^{-1} \]
The one-step level residuals are given by

\[ \hat{\epsilon}^L_{1i} = y^L_i - \hat{\beta}_1 X^L_i \]

Stacking the residual vectors yields

\[ \hat{\epsilon}_{1i} = \begin{pmatrix} \hat{\epsilon}^*_{1i} \\ \hat{\epsilon}^L_{1i} \end{pmatrix} \]

which is used to compute \( H_{2i} = \hat{\epsilon}'_{1i} \hat{\epsilon}_{1i} \), which is used in

\[ A_2 = \left( \sum_i Z_i' H_{2i} Z_i \right)^{-1} \]

and the robust one-step VCE is given by

\[ \hat{V}_{\text{robust}}[\hat{\beta}_1] = W_1^{-1} Q_{xz} A_1 A_2^{-1} A_1 Q_{xz}' W_1^{-1} \]

\( \hat{V}_{\text{robust}}[\hat{\beta}_1] \) is robust to heteroskedasticity in the errors.

After defining

\[ W_2 = Q_{xz} A_2 Q_{xz}' \]

the two-step estimates are given by

\[ \hat{\beta}_2 = W_2^{-1} Q_{xz} A_2 Q_{zy} \]

The GMM two-step VCE is then given by

\[ \hat{V}_{\text{GMM}}[\hat{\beta}_2] = W_2^{-1} \]

The GMM two-step VCE is known to be severely biased. Windmeijer (2005) derived the Windmeijer bias-corrected (WC) estimator for the robust VCE of two-step GMM estimators. xtdpd implements this WC-robust estimator of the VCE. The formulas for this method are involved; see Windmeijer (2005). The WC-robust estimator of the VCE is robust to heteroskedasticity in the errors.

Acknowledgment

We thank David Roodman of the Open Philanthropy Project, who wrote xtabond2.

References


Also see

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

[XT] **xtabond** — Arellano–Bond linear dynamic panel-data estimation

[XT] **xtdpdsys** — Arellano–Bover/Blundell–Bond linear dynamic panel-data estimation

[XT] **xtivreg** — Instrumental variables and two-stage least squares for panel-data models

[XT] **xtreg** — Fixed-, between-, and random-effects and population-averaged linear models

[XT] **xtregar** — Fixed- and random-effects linear models with an AR(1) disturbance

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

[R] **gmm** — Generalized method of moments estimation

[U] **20 Estimation and postestimation commands**