Generalized method of moments estimation of linear dynamic panel data models

Sebastian Kripfganz

University of Exeter Business School, Department of Economics, Exeter, UK

Stata Conference

July 31, 2020

```
ssc install xtdpdgmm
net install xtdpdgmm, from(http://www.kripfganz.de/stata/)
```

GMM estimation of linear dynamic panel data models

- Instrumental variables (IV) / generalized method of moments (GMM) estimation is the predominant estimation technique for panel data models with unobserved unit-specific heterogeneity and endogenous variables, in particular lagged dependent variables, when the time horizon is short.
- This presentation introduces the community-contributed `xtdpdgmm` Stata command.
- For a longer version of this talk with many additional details, see my 2019 London Stata Conference presentation: `https://www.stata.com/meeting/uk19/slides/uk19_kripfganz.pdf`

GMM estimation of linear dynamic panel data models

- Official Stata commands:
  - `xtdpd` command for the Arellano and Bond (1991) *difference GMM* (diff-GMM) and the Arellano and Bover (1995) and Blundell and Bond (1998) *system GMM* (sys-GMM) estimation.
  - `xtabond` command for diff-GMM estimation; `xtdpd` wrapper.
  - `xtdpdsys` command for sys-GMM estimation; `xtdpd` wrapper.
  - `gmm` command for GMM estimation (not just of dynamic panel data models).
- Community-contributed Stata commands:
  - `xtabond2` command by Roodman (2009) for diff-GMM and sys-GMM estimation.
  - `xtdpdgmm` command for diff-GMM, sys-GMM, and GMM estimation with the Ahn and Schmidt (1995) nonlinear moment conditions.

Concerns about existing Stata commands

- Official Stata commands lack flexibility and suffer from bugs:
  - Specification of time dummies `i.`*`timevar`*: collinearity checks in `xtdpd` (and therefore also `xtabond` and `xtdpdsys`) lead to the omission of 1 time dummy too many.
  - `xtdpd` and `gmm` yield incorrect estimates in some cases of unbalanced panel data sets.
  - Option `diffvars()` of `xtabond` yields incorrect predictions.
- Community-contributed Stata command `xtabond2` suffers from bugs as well:
  - Incorrect estimates in some cases when forward-orthogonal deviations are combined with standard instruments.
  - Incorrect estimates in some cases of unbalanced panel data sets.
  - Incorrect degrees of freedom and p-values for the overidentification tests if some coefficients are shown as *omitted* (or *empty*), a typical concern with time dummies.

Linear dynamic panel data model

- Linear dynamic panel data model:

$$y_{it} = \lambda y_{i,t-1} + \mathbf{x}'_{it}\beta + \underbrace{\alpha_i + u_{it}}_{=e_{it}}$$

with many cross-sectional units $i = 1, 2, \ldots, N$ and few time periods $t = 1, 2, \ldots, T$.

  - Further lags of y_{it} and $\mathbf{x}_{it}$ can be added as regressors.
  - The regressors $\mathbf{x}_{it}$ can be strictly exogenous, weakly exogenous (predetermined), or endogenous.
  - The idiosyncratic error term u_{it} shall be serially uncorrelated.
  - The unobserved unit-specific heterogeneity α_i can be correlated with the regressors $\mathbf{x}_{it}$. It is correlated by construction with the lagged dependent variable $y_{i,t-1}$.

Model transformations supported by xtdpdgmm

- First-difference transformation (Anderson and Hsiao, 1981; Arellano and Bond, 1991), option `model(difference)`:

$$\Delta y_{it} = \lambda \Delta y_{i,t-1} + \Delta \mathbf{x}_{it}' \beta + \Delta e_{it}$$

- Forward-orthogonal deviations (Arellano and Bover, 1995), option `model(fodev)`:

$$\tilde{\Delta}_t y_{it} = \lambda \tilde{\Delta}_t y_{i,t-1} + \tilde{\Delta}_t \mathbf{x}_{it}' \beta + \tilde{\Delta}_t e_{it}$$

where $\tilde{\Delta}_t e_{it} = \sqrt{\frac{T-t+1}{T-t}} \left(e_{it} - \frac{1}{T-t+1} \sum_{s=0}^{T-t} e_{i,t+s} \right)$.

- Deviations from within-group means, option `model(mdev)`:

$$\ddot{\Delta} y_{it} = \lambda \ddot{\Delta} y_{i,t-1} + \ddot{\Delta} \mathbf{x}_{it}' \beta + \ddot{\Delta} e_{it}$$

where $\ddot{\Delta} e_{it} = \sqrt{\frac{T}{T-1}} (e_{it} - \bar{e}_i)$.

GMM-type instruments

- Stacked moment conditions (for the first-differenced model):

$$E\left[\mathbf{Z}_i^{D\prime}\Delta\mathbf{e}_i\right] = \mathbf{0}$$

where $\Delta\mathbf{e}_i = (\Delta e_{i2}, \Delta e_{i3}, \ldots, \Delta e_{iT})'$, and $\mathbf{Z}_i^D = (\mathbf{Z}_{yi}^D, \mathbf{Z}_{xi}^D)$, with *GMM-type* instruments

$$\mathbf{Z}_{yi}^D = \begin{pmatrix} y_{i0} & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & y_{i0} & y_{i1} & \cdots & 0 & 0 & \cdots & 0 \\ & & & \ddots & & & & \\ 0 & 0 & 0 & \cdots & y_{i0} & y_{i1} & \cdots & y_{i,T-2} \end{pmatrix} \begin{matrix} \leftarrow t = 2 \\ \leftarrow t = 3 \\ \vdots \\ \leftarrow t = T \end{matrix}$$

and similarly for $\mathbf{Z}_{xi}^D$.

- Moment conditions for other model transformations are stacked likewise.

One-step diff-GMM estimation

- *GMM-type* instruments specified with the `gmmiv()` option, exemplarily for predetermined `w` and strictly exogenous `k`:

```
. webuse abdata

. xtdpdgmm L(0/1).n w k, model(diff) gmm(n, lag(2 .)) gmm(w, lag(1 .)) gmm(k, lag(. .)) nocons
note: standard errors may not be valid

Generalized method of moments estimation

Fitting full model:
Step 1         f(b) =  .01960406

Group variable: id                           Number of obs         =       891
Time variable: year                          Number of groups      =       140

Moment conditions:     linear =      126     Obs per group:    min =         6
                    nonlinear =        0                       avg =  6.364286
                        total =      126                       max =         8

------------------------------------------------------------------------------
           n |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   .4144164   .0341502    12.14   0.000     .3474833    .4813495
             |
           w |  -.8292293   .0588914   -14.08   0.000    -.9446543   -.7138042
           k |   .3929936   .0223829    17.56   0.000     .3491239    .4368634
------------------------------------------------------------------------------
```

(Continued on next page)

One-step diff-GMM estimation

```
Instruments corresponding to the linear moment conditions:
 1, model(diff):
   1978:L2.n 1979:L2.n 1980:L2.n 1981:L2.n 1982:L2.n 1983:L2.n 1984:L2.n
   1979:L3.n 1980:L3.n 1981:L3.n 1982:L3.n 1983:L3.n 1984:L3.n 1980:L4.n
   1981:L4.n 1982:L4.n 1983:L4.n 1984:L4.n 1981:L5.n 1982:L5.n 1983:L5.n
   1984:L5.n 1982:L6.n 1983:L6.n 1984:L6.n 1983:L7.n 1984:L7.n 1984:L8.n
 2, model(diff):
   1978:L1.w 1979:L1.w 1980:L1.w 1981:L1.w 1982:L1.w 1983:L1.w 1984:L1.w
   1978:L2.w 1979:L2.w 1980:L2.w 1981:L2.w 1982:L2.w 1983:L2.w 1984:L2.w
   1979:L3.w 1980:L3.w 1981:L3.w 1982:L3.w 1983:L3.w 1984:L3.w 1980:L4.w
   1981:L4.w 1982:L4.w 1983:L4.w 1984:L4.w 1981:L5.w 1982:L5.w 1983:L5.w
   1984:L5.w 1982:L6.w 1983:L6.w 1984:L6.w 1983:L7.w 1984:L7.w 1984:L8.w
 3, model(diff):
   1978:F6.k 1978:F5.k 1979:F5.k 1978:F4.k 1979:F4.k 1980:F4.k 1978:F3.k
   1979:F3.k 1980:F3.k 1981:F3.k 1978:F2.k 1979:F2.k 1980:F2.k 1981:F2.k
   1982:F2.k 1978:F1.k 1979:F1.k 1980:F1.k 1981:F1.k 1982:F1.k 1983:F1.k
   1978:k 1979:k 1980:k 1981:k 1982:k 1983:k 1984:k 1978:L1.k 1979:L1.k
   1980:L1.k 1981:L1.k 1982:L1.k 1983:L1.k 1984:L1.k 1978:L2.k 1979:L2.k
   1980:L2.k 1981:L2.k 1982:L2.k 1983:L2.k 1984:L2.k 1979:L3.k 1980:L3.k
   1981:L3.k 1982:L3.k 1983:L3.k 1984:L3.k 1980:L4.k 1981:L4.k 1982:L4.k
   1983:L4.k 1984:L4.k 1981:L5.k 1982:L5.k 1983:L5.k 1984:L5.k 1982:L6.k
   1983:L6.k 1984:L6.k 1983:L7.k 1984:L7.k 1984:L8.k
```

- `xtdpdgmm` has the options `nolog`, `noheader`, `notable`, and `nofootnote` to suppress undesired output.

Too-many-instruments problem

- Too many instruments relative to the cross-sectional sample size can aggravate finite-sample biases in the coefficient and standard error estimates and potentially weakens specification tests (Roodman, 2009a).
- To reduce the number of instruments, two main approaches are typically used (Roodman, 2009a, 2009b; Kiviet, 2020):
  - Curtailing: Limit the number of lags used as instruments, suboption `lagrange()`, e.g. $y_{i,t-2}, y_{i,t-3}, \ldots, y_{i,t-l}$.
  - Collapsing: Use *standard* instruments instead of *GMM-type* instruments, suboption `collapse` or option `iv()`, e.g.

$$\mathbf{Z}_{yi}^{D} = \begin{pmatrix} y_{i0} & 0 & \cdots & 0 \\ y_{i1} & y_{i0} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ y_{i,T-2} & y_{i,T-3} & \cdots & y_{i0} \end{pmatrix} \begin{matrix} \leftarrow & t = 2 \\ \leftarrow & t = 3 \\ & \vdots \\ \leftarrow & t = T \end{matrix}$$

Sys-GMM estimation

- Instruments for different model transformations can be combined with each other and with instruments for the untransformed model, option `model(level)`.
  - Instruments for the level model might require an additional initial-conditions / mean stationarity assumption to ensure that they are uncorrelated with the unobserved unit-specific heterogeneity α_i (Blundell and Bond, 1998; Blundell, Bond, and Windmeijer; 2001).
- Stacked moment conditions:

$$E\left[\begin{pmatrix} \mathbf{Z}_i^{D'}\Delta\mathbf{e}_i \\ \mathbf{Z}_i^{L'}\mathbf{e}_i \end{pmatrix}\right] = \mathbf{0}$$

where $\mathbf{e}_i = (e_{i2}, e_{i3}, \ldots, e_{iT})'$.

Sys-GMM as level GMM

- Alternative formulation of the stacked moment conditions, noting that $\Delta\mathbf{e}_i = \mathbf{D}_i\mathbf{e}_i$ (where $\mathbf{D}_i$ is the first-difference transformation matrix):

$$E\left[\begin{pmatrix} \mathbf{Z}_i^{D\prime}\mathbf{D}_i\mathbf{e}_i \\ \mathbf{Z}_i^{L\prime}\mathbf{e}_i \end{pmatrix}\right] = E\left[\begin{pmatrix} \mathbf{Z}_i^{D\prime}\mathbf{D}_i \\ \mathbf{Z}_i^{L\prime} \end{pmatrix}\mathbf{e}_i\right] = E[\mathbf{Z}_i'\mathbf{e}_i] = \mathbf{0}$$

  where $\mathbf{Z}_i = (\tilde{\mathbf{Z}}_i^D, \mathbf{Z}_i^L)$ is a set of instruments for the level model with transformed instruments $\tilde{\mathbf{Z}}_i^D = \mathbf{D}_i'\mathbf{Z}_i^D$, and analogously for other model transformations.
  - The sys-GMM estimator can be written as a *level GMM* estimator (Arellano and Bover, 1995).
  - Internally, this is how `xtdpdgmm` is implemented.

Two-step estimation with optimal weighting matrix

- One-step diff-GMM is efficient only under a strong homoskedasticity assumption.
- One-step sys-GMM is inefficient even under homoskedasticity.
- For efficient two-step estimation with an optimal weighting matrix, option `twostep`, the Windmeijer (2005) finite-sample correction is applied for panel-robust or cluster-robust standard errors, options `vce(robust)` or `vce(cluster` *`clustvar`*`)`, respectively.

Two-step sys-GMM estimation

- Combination of curtailed and collapsed instruments:

```
. xtdpdgmm L(0/1).n w k, model(diff) collapse gmm(n, lag(2 4)) gmm(w k, lag(1 3)) ///
> gmm(n, lag(1 1) diff model(level)) gmm(w k, lag(0 0) diff model(level)) two vce(r) nofootnote

Generalized method of moments estimation

Fitting full model:
Step 1         f(b) =  .00285146
Step 2         f(b) =  .11568719

Group variable: id                          Number of obs         =        891
Time variable: year                         Number of groups      =        140

Moment conditions:     linear =        13   Obs per group:    min =          6
                    nonlinear =         0                     avg =   6.364286
                        total =        13                     max =          8

                                   (Std. Err. adjusted for 140 clusters in id)
------------------------------------------------------------------------------
             |              WC-Robust
           n |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   .5117523   .1208484     4.23   0.000     .2748937    .7486109
             |
           w |  -1.323125   .2383451    -5.55   0.000    -1.790273    -.855977
           k |   .1931365   .0941343     2.05   0.040     .0086367    .3776363
       _cons |   4.698425   .7943584     5.91   0.000     3.141511    6.255339
------------------------------------------------------------------------------
```

Postestimation specification tests

- Arellano and Bond (1991) tests for absence of higher-order serial correlation: `estat serial`.
- Sargan (1958) / Hansen (1982) tests for the validity of the overidentifying restrictions: `estat overid`.

```
. quietly xtdpdgmm L(0/1).n w k, model(diff) collapse gmm(n, lag(2 4)) gmm(w k, lag(1 3)) ///
> gmm(n, lag(1 1) diff model(level)) gmm(w k, lag(0 0) diff model(level)) two vce(r)

. estat serial, ar(1/3)

Arellano-Bond test for autocorrelation of the first-differenced residuals
H0: no autocorrelation of order 1:     z =   -3.3341   Prob > |z|  =    0.0009
H0: no autocorrelation of order 2:     z =   -1.2436   Prob > |z|  =    0.2136
H0: no autocorrelation of order 3:     z =   -0.1939   Prob > |z|  =    0.8462

. estat overid

Sargan-Hansen test of the overidentifying restrictions
H0: overidentifying restrictions are valid

2-step moment functions, 2-step weighting matrix       chi2(9)      =   16.1962
                                                       Prob > chi2 =    0.0629

2-step moment functions, 3-step weighting matrix       chi2(9)      =   13.8077
                                                       Prob > chi2 =    0.1293
```

Incremental overidentification tests

- Under the assumption that the diff-GMM estimator is correctly specified, we can test the validity of the additional moment conditions for the level model with incremental overidentification tests / *difference Sargan-Hansen* tests
  - `xtdpdgmm` specified with option `overid` computes incremental overidentification tests for each set of `gmmiv()` or `iv()` instruments, and jointly for all moment conditions referring to the same model transformation. The incremental tests are displayed by the postestimation command `estat overid` when called with option `difference`.
- A generalized Hausman (1978) test can be performed as an alternative to incremental Sargan-Hansen tests: `estat hausman`.

Incremental overidentification tests

```
. xtdpdgmm L(0/1).n w k, model(diff) collapse gmm(n, lag(2 4)) gmm(w k, lag(1 3)) ///
> gmm(n, lag(1 1) diff model(level)) gmm(w k, lag(0 0) diff model(level)) two vce(r) overid

Generalized method of moments estimation

Fitting full model:
Step 1         f(b) =  .00285146
Step 2         f(b) =  .11568719

Fitting reduced model 1:
Step 1         f(b) =  .10476123

Fitting reduced model 2:
Step 1         f(b) =  .02873833

Fitting reduced model 3:
Step 1         f(b) =   .1131458

Fitting reduced model 4:
Step 1         f(b) =  .08632894

Fitting no-diff model:
Step 1         f(b) =  8.476e-19

Fitting no-level model:
Step 1         f(b) =  .05779984
(Some output omitted)
(Continued on next page)
```

Incremental overidentification tests

```
Instruments corresponding to the linear moment conditions:
 1, model(diff):
   L2.n L3.n L4.n
 2, model(diff):
   L1.w L2.w L3.w L1.k L2.k L3.k
 3, model(level):
   L1.D.n
 4, model(level):
   D.w D.k
 5, model(level):
   _cons

. estat overid, difference

Sargan-Hansen (difference) test of the overidentifying restrictions
H0: (additional) overidentifying restrictions are valid

2-step weighting matrix from full model

                  | Excluding                   | Difference
Moment conditions |       chi2     df         p |        chi2     df         p
------------------+-----------------------------+-----------------------------
   1, model(diff) |    14.6666      6    0.0230 |      1.5296      3    0.6754
   2, model(diff) |     4.0234      3    0.2590 |     12.1728      6    0.0582
  3, model(level) |    15.8404      8    0.0447 |      0.3558      1    0.5509
  4, model(level) |    12.0861      7    0.0978 |      4.1102      2    0.1281
      model(diff) |     0.0000      0         . |     16.1962      9    0.0629
     model(level) |     8.0920      6    0.2314 |      8.1042      3    0.0439
```

Model and moment selection criteria

- The Andrews and Lu (2001) model and moment selection criteria (MMSC) can support the specification search.
  - The `xtdpdgmm` postestimation command `estat mmsc` computes the Akaike (AIC), Bayesian (BIC), and Hannan-Quinn (HQIC) versions of the Andrews-Lu MMSC.
  - Models with lower values of the criteria are preferred.

```
. estimates store noxlags

. quietly xtdpdgmm L(0/1).n L(0/1).(w k), model(diff) collapse gmm(n, lag(2 4)) ///
> gmm(w k, lag(1 3)) gmm(n, lag(1 1) diff model(level)) gmm(w k, lag(0 0) diff model(level)) two vce(r)

. estimates store xlags

. quielty xtdpdgmm L(0/1).n L(0/1).(w k) c.w#c.k, model(diff) collapse gmm(n, lag(2 4)) ///
> gmm(w k, lag(1 3)) gmm(n, lag(1 1) diff model(level)) gmm(w k, lag(0 0) diff model(level)) two vce(r)

. estat mmsc xlags noxlags

Andrews-Lu model and moment selection criteria

       Model | ngroups          J  nmom  npar   MMSC-AIC   MMSC-BIC  MMSC-HQIC
-------------+----------------------------------------------------------------
           . |     140     1.5797    13     7   -10.4203   -28.0702   -17.7844
       xlags |     140    12.9784    13     6    -1.0216   -21.6131    -9.6130
     noxlags |     140    16.1962    13     4    -1.8038   -28.2786   -12.8499
```

Sys-GMM estimation: transformed instruments

- The postestimation command `predict` with option `iv` generates the transformed instruments for the level model, $\mathbf{Z}_i = (\tilde{\mathbf{Z}}_i^D, \mathbf{Z}_i^L)$ (excluding the intercept), as new variables, e.g. for subsequent use with the official `ivregress` command, the community-contributed `ivreg2` command (Baum, Schaffer, and Stillman, 2003, 2007), or any other tool.

```
. quietly predict iv*, iv

. describe iv*

              storage   display    value
variable name   type    format     label      variable label
-------------------------------------------------------------------------------
iv1             float   %9.0g                 1, model(diff): L2.n
iv2             float   %9.0g                 1, model(diff): L3.n
iv3             float   %9.0g                 1, model(diff): L4.n
iv4             float   %9.0g                 2, model(diff): L1.w
iv5             float   %9.0g                 2, model(diff): L2.w
iv6             float   %9.0g                 2, model(diff): L3.w
iv7             float   %9.0g                 2, model(diff): L1.k
iv8             float   %9.0g                 2, model(diff): L2.k
iv9             float   %9.0g                 2, model(diff): L3.k
iv10            float   %9.0g                 3, model(level): L1.D.n
iv11            float   %9.0g                 4, model(level): D.w
iv12            float   %9.0g                 4, model(level): D.k
```

Two-step sys-GMM estimation

```
. ivregress gmm n (L.n w k = iv*), wmat(cluster id)

Instrumental variables (GMM) regression          Number of obs   =        891
                                                 Wald chi2(3)    =     485.45
                                                 Prob > chi2     =     0.0000
                                                 R-squared       =     0.8545
GMM weight matrix: Cluster (id)                  Root MSE        =     .51125

                                   (Std. Err. adjusted for 140 clusters in id)
------------------------------------------------------------------------------
             |               Robust
           n |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   .5117523    .098918     5.17   0.000     .3178765    .7056281
             |
           w |  -1.323125   .2031404    -6.51   0.000    -1.721273    -.924977
           k |   .1931365   .0873607     2.21   0.027     .0219126    .3643604
       _cons |   4.698425   .6369462     7.38   0.000     3.450034    5.946817
------------------------------------------------------------------------------
Instrumented:  L.n w k
Instruments:   iv1 iv2 iv3 iv4 iv5 iv6 iv7 iv8 iv9 iv10 iv11 iv12

. estat overid

  Test of overidentifying restriction:

  Hansen's J chi2(9) = 16.1962 (p = 0.0629)
```

Two-step sys-GMM estimation

```
. ivreg2 n (L.n w k = iv*), gmm2s cluster(id)

2-Step GMM estimation
---------------------

Estimates efficient for arbitrary heteroskedasticity and clustering on id
Statistics robust to heteroskedasticity and clustering on id

Number of clusters (id) =          140              Number of obs =      891
                                                    F(  3,    139) =   230.77
                                                    Prob > F      =   0.0000
Total (centered) SS     =  1601.042507              Centered R2   =   0.8545
Total (uncentered) SS   =  2564.249196              Uncentered R2 =   0.9092
Residual SS             =  232.8868955              Root MSE      =    .5113

------------------------------------------------------------------------------
             |               Robust
           n |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   .5117523   .0822341     6.22   0.000     .3505763    .6729282
             |
           w |  -1.323125   .1621898    -8.16   0.000    -1.641011   -1.005239
           k |   .1931365   .0660458     2.92   0.003     .0636892    .3225838
       _cons |   4.698425   .5321653     8.83   0.000     3.655401     5.74145
------------------------------------------------------------------------------
```

(Continued on next page)

Two-step sys-GMM estimation

```
Underidentification test (Kleibergen-Paap rk LM statistic):          30.312
                                           Chi-sq(10) P-val =    0.0008
------------------------------------------------------------------------------
Weak identification test (Cragg-Donald Wald F statistic):              0.376
                         (Kleibergen-Paap rk Wald F statistic):        5.128
Stock-Yogo weak ID test critical values:  5% maximal IV relative bias    17.80
                                         10% maximal IV relative bias    10.01
                                         20% maximal IV relative bias     5.90
                                         30% maximal IV relative bias     4.42
Source: Stock-Yogo (2005).  Reproduced by permission.
NB: Critical values are for Cragg-Donald F statistic and i.i.d. errors.
------------------------------------------------------------------------------
Hansen J statistic (overidentification test of all instruments):     16.196
                                           Chi-sq(9) P-val =     0.0629
------------------------------------------------------------------------------
Instrumented:         L.n w k
Excluded instruments: iv1 iv2 iv3 iv4 iv5 iv6 iv7 iv8 iv9 iv10 iv11 iv12
------------------------------------------------------------------------------
```

Underidentification tests

- While it is standard practice to test for overidentification, the potential problem of underidentification is largely ignored in the empirical practice.
- The new `underid` command (now on SSC) by Mark Schaffer and Frank Windmeijer presents underidentification statistics (Windmeijer, 2018). From the users' perspective, `underid` works as a postestimation command for `xtdpdgmm`.
  - The null hypothesis of the underidentification tests is that the model is underidentfied. (The aim is to reject the null hypothesis, as opposed to overidentification tests.)

Underidentification tests

```
. quietly xtdpdgmm L(0/1).n w k, model(diff) collapse gmm(n, lag(2 4)) gmm(w k, lag(1 3)) ///
> gmm(n, lag(1 1) diff model(level)) gmm(w k, lag(0 0) diff model(level)) two vce(r)

. underid
Number of obs:      891
Number of panels:   140
Dep var:            n
Endog Xs (3):       L.n w k
Exog Xs (1):        _cons
Excl IVs (12):      __alliv_1 __alliv_2 __alliv_3 __alliv_4 __alliv_5 __alliv_6
                    __alliv_7 __alliv_8 __alliv_9 __alliv_10 __alliv_11
                    __alliv_12

Underidentification test: Cragg-Donald robust CUE-based (LM version)
  Test statistic robust to heteroskedasticity and clustering on id
j=   26.92  Chi-sq( 10) p-value=0.0027

. underid, kp sw noreport

Underidentification test: Kleibergen-Paap robust LIML-based (LM version)
  Test statistic robust to heteroskedasticity and clustering on id
j=   30.31  Chi-sq( 10) p-value=0.0008

2-step GMM J underidentification stats by regressor:
j=   30.00  Chi-sq( 10) p-value=0.0009 L.n
j=   29.07  Chi-sq( 10) p-value=0.0012 w
j=   26.01  Chi-sq( 10) p-value=0.0037 k
```

Nonlinear moment conditions

- Absence of serial correlation in u_{it} is a necessary condition for the validity of $y_{i,t-2}, y_{i,t-3}, \dots$ as instruments for the first-differenced model.
- The nonlinear (quadratic) moment conditions suggested by Ahn and Schmidt (1995) can help to improve the efficiency and to achieve identification.
  - Absence of serial correlation: option `nl(noserial)`.
  - Absence of serial correlation plus homoskedasticity: option `nl(iid)`.
- While GMM estimators with only linear moment conditions have a closed-form solution, this is no longer the case with nonlinear moment conditions.
  - `xtdpdgmm` minimizes the GMM criterion function numerically with Stata's Gauss-Newton algorithm.

Estimation with nonlinear moment conditions

- The nonlinear moment conditions can be optionally collapsed into a single moment condition.

```
. xtdpdgmm L(0/1).n w k, model(diff) collapse gmm(n, lag(2 4)) gmm(w k, lag(1 3)) nl(noserial) igmm
> vce(r) nolog nofootnote

Generalized method of moments estimation

Group variable: id                              Number of obs         =       891
Time variable: year                             Number of groups      =       140

Moment conditions:      linear =        10      Obs per group:    min =         6
                     nonlinear =         1                        avg =  6.364286
                         total =        11                        max =         8

                                   (Std. Err. adjusted for 140 clusters in id)
------------------------------------------------------------------------------
             |              WC-Robust
           n |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   .5048501   .1229569     4.11   0.000     .2638591    .7458411
             |
           w |  -1.712339   .2553838    -6.70   0.000    -2.212882   -1.211796
           k |   .0645476   .1152549     0.56   0.575    -.1613478    .2904429
       _cons |   5.884724   .7948763     7.40   0.000     4.326795    7.442653
------------------------------------------------------------------------------
```

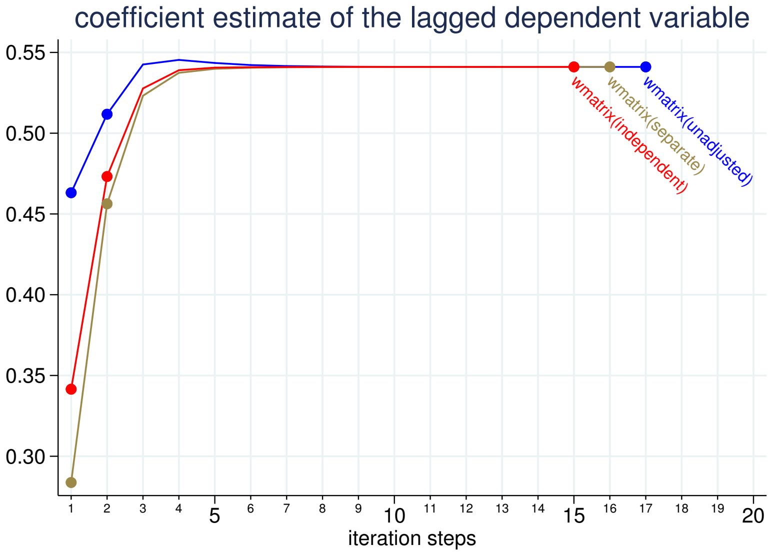
Iterated GMM estimation

- While the two-step estimator is asymptotically efficient (for a given set of instruments), in finite samples the estimation of the optimal weighting matrix might be sensitive to the (arbitrarily) chosen initial weighting matrix.
- Hansen, Heaton, and Yaron (1996) suggest to use an iterated GMM estimator that updates the weighting matrix and coefficient estimates until convergence.
  - Similar to Stata's `gmm` or `ivregress` command, `xtdpdgmm` provides the option `igmm` as alternatives to `onestep` and `twostep`.

Iterated sys-GMM estimation

```
. xtdpdgmm L(0/1).n w k, model(diff) collapse gmm(n, lag(2 4)) gmm(w k, lag(1 3)) ///
> gmm(n, lag(1 1) diff model(level)) gmm(w k, lag(0 0) diff model(level)) igmm vce(r) nofootnote

Generalized method of moments estimation

Fitting full model:
Steps
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
.................                                                    17

Group variable: id                           Number of obs         =        891
Time variable: year                          Number of groups      =        140

Moment conditions:      linear =        13      Obs per group:     min =          6
                     nonlinear =         0                         avg =   6.364286
                         total =        13                         max =          8

                                   (Std. Err. adjusted for 140 clusters in id)
------------------------------------------------------------------------------
             |              WC-Robust
           n |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |    .541044   .1265822     4.27   0.000     .2929474    .7891406
             |
           w |  -1.527984    .304707    -5.01   0.000    -2.125199   -.9307697
           k |   .1075032   .1115814     0.96   0.335    -.1111923    .3261986
       _cons |   5.275027   .9736502     5.42   0.000     3.366707    7.183346
------------------------------------------------------------------------------
```

Iterated sys-GMM estimation: initial weighting matrices



Continuously updated GMM estimation

- As an alternative to the iterated GMM estimator, Hansen, Heaton, and Yaron (1996) also suggest a continuously updated GMM estimator, where the optimal weighting matrix is obtained directly as part of the minimization process.
    - This estimator is not currently implemented in `xtdpdgmm` but the `ivreg2` command can be used with the instruments previously generated from `xtdpdgmm`.

Continuously updated sys-GMM estimation

```
. ivreg2 n (L.n w k = iv*), cue cluster(id)
Iteration 0:   f(p) =  24.858945  (not concave)
(Some output omitted)
Iteration 21:  f(p) =  8.2335574

CUE estimation
--------------

Estimates efficient for arbitrary heteroskedasticity and clustering on id
Statistics robust to heteroskedasticity and clustering on id
(Some output omitted)
------------------------------------------------------------------------------
             |               Robust
           n |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   .5239428   .1138624     4.60   0.000     .3007766    .7471089
             |
           w |  -2.025771   .2810169    -7.21   0.000    -2.576555   -1.474988
           k |  -.0193789   .1221278    -0.16   0.874    -.2587449    .2199872
       _cons |   6.781101   .8346986     8.12   0.000     5.145122     8.41708
------------------------------------------------------------------------------
(Some output omitted)
Hansen J statistic (overidentification test of all instruments):         8.234
                                                   Chi-sq(9) P-val =    0.5108
------------------------------------------------------------------------------
Instrumented:         L.n w k
Excluded instruments: iv1 iv2 iv3 iv4 iv5 iv6 iv7 iv8 iv9 iv10 iv11 iv12
------------------------------------------------------------------------------
```

Time effects

- To account for global shocks, it is common practice to include a set of time dummies in the regression model:

$$y_{it} = \lambda y_{i,t-1} + \mathbf{x}_{it}'\boldsymbol{\beta} + \delta_t + \underbrace{\alpha_i + u_{it}}_{=e_{it}}$$

  - Without loss of generality, time dummies δ_t can be treated as strictly exogenous and uncorrelated with the unit-specific effects α_i. Hence, time dummies can be instrumented by themselves.

GMM estimation with time effects

- xtdpdgmm has the option teffects that automatically adds the correct number of time dummies and corresponding instruments:

```
. xtdpdgmm L(0/1).n w k, model(diff) collapse gmm(n, lag(2 4)) gmm(w k, lag(1 3)) nl(noserial) ///
> teffects igmm vce(r)

Generalized method of moments estimation

Fitting full model:
Steps
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
...................................                  35

Group variable: id                       Number of obs         =       891
Time variable: year                      Number of groups      =       140

Moment conditions:     linear =      17     Obs per group:    min =         6
                    nonlinear =       1                       avg =  6.364286
                        total =      18                       max =         8

                                   (Std. Err. adjusted for 140 clusters in id)
```

(Continued on next page)

GMM estimation with time effects

```
------------------------------------------------------------------------------
             |              WC-Robust
           n |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |    .715963    .2630756     2.72   0.006     .2003442    1.231582
             |
           w |  -.7645527    .6235711    -1.23   0.220     -1.98673    .4576242
           k |   .4043948     .270444     1.50   0.135    -.1256657    .9344553
             |
        year |
       1978  |  -.0656579    .0317356    -2.07   0.039    -.1278586   -.0034572
       1979  |  -.0825628    .0346171    -2.39   0.017    -.1504111   -.0147145
       1980  |  -.1035026    .0263053    -3.93   0.000      -.15506   -.0519452
       1981  |  -.1335986    .0313492    -4.26   0.000    -.1950419   -.0721553
       1982  |  -.0661445    .0574973    -1.15   0.250    -.1788372    .0465482
       1983  |   .0033487    .0685548     0.05   0.961    -.1310163    .1377137
       1984  |   .0538893    .1010754     0.53   0.594    -.1442148    .2519933
             |
       _cons |   2.932618    2.345137     1.25   0.211    -1.663767    7.529002
------------------------------------------------------------------------------
Instruments corresponding to the linear moment conditions:
 1, model(diff):
   L2.n L3.n L4.n
 2, model(diff):
   L1.w L2.w L3.w L1.k L2.k L3.k
 3, model(level):
   1978bn.year 1979.year 1980.year 1981.year 1982.year 1983.year 1984.year
 4, model(level):
   _cons
```

Summary: the xtdpdgmm package for Stata

- The xtdpdgmm package enables generalized method of moments estimation of linear (dynamic) panel data models.
  - Besides the conventional *difference GMM*, *system GMM*, and GMM with forward-orthogonal deviations, additional nonlinear moment conditions can be incorporated.
  - Besides one-step and feasible efficient two-step estimation, iterated GMM estimation is possible as well.
  - Combining the command with other packages in the Stata universe opens up further possibilities.

```
ssc install xtdpdgmm
net install xtdpdgmm, from(http://www.kripfganz.de/stata/)

help xtdpdgmm
help xtdpdgmm postestimation
```

References

- Ahn, S. C., and P. Schmidt (1995). Efficient estimation of models for dynamic panel data. *Journal of Econometrics 68* (1): 5–27.
- Anderson, T. W., and C. Hsiao (1981). Estimation of dynamic models with error components. *Journal of the American Statistical Association 76* (375): 598–606.
- 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 S. R. Bond (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies 58* (2): 277–297.
- Arellano, M., and O. Bover (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics 68* (1): 29–51.
- Baum, C. F., M. E. Schaffer, and S. Stillman (2003). Instrumental variables and GMM: Estimation and testing. *Stata Journal 3* (1): 1–31.
- Baum, C. F., M. E. Schaffer, and S. Stillman (2007). Enhanced routines for instrumental variables/generalized method of moments estimation and testing. *Stata Journal 7* (4): 465–506.
- Blundell, R., and S. R. Bond (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics 87* (1): 115–143.
- Blundell, R., S. R. Bond, and F. Windmeijer (2001). Estimation in dynamic panel data models: Improving on the performance of the standard GMM estimator. *Advances in Econometrics 15* (1): 53–91.

References

- Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50* (4): 1029–1054.
- Hansen, L. P., J. Heaton, and A. Yaron (1996). Finite-sample properties of some alternative GMM estimators. *Journal of Business & Economic Statistics 14* (3): 262–280.
- Hausman, J. A. (1978). Specification tests in Econometrics. *Econometrica 46* (6): 1251–1271.
- Kiviet, J. F. (2020). Microeconometric dynamic panel data methods: Model specification and selection issues. *Econometrics and Statistics 13*: 16–45.
- Roodman, D. (2009a). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics 71* (1): 135–158.
- Roodman, D. (2009b). How to do xtabond2? An introduction to difference and system GMM in Stata. *Stata Journal 9* (1): 86–136.
- Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica 26* (3): 393–415.
- Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics 126* (1): 25–51.
- Windmeijer, F. (2018). Testing over- and underidentification in linear models, with applications to dynamic panel data and asset-pricing models. *Economics Discussion Paper* 18/696, University of Bristol.