☺A

DYNAMIC MODELS FOR PANEL DATA: APPLICATION IN STATA

1. Introduction

The aim of this note is to show some alternative estimates of a dynamic panel model: OLS, within, first-differences, IV, GMM. We use a dynamic panel model as an example of the bias deriving from endogeneity and of the most appropriate estimation methods.[1] Estimation methods like GMM can be used each time we tackle simultaneity or measurement errors, also in static model.
Moreover, we try to show how the same estimate may be obtained in different ways in Stata; this in order to better understand the syntax of two new commands, `ivreg2.ado` and `xtabond2.ado`, that represent a very useful extension of the Stata environment in dealing with endogeneity.

`ivreg2` (see C. F. Baum, M. E. Schaffer and S. Stillman (2003) Instrumental variables and GMM: Estimation and testing, The Stata Journal, 3, 1, 1-31) is a general estimating comand. In fact, it implements a range of single-equation estimation methods for the linear regression model: OLS, instrumental variables (IV, also known as two-stage least squares, 2SLS), the two-step feasible generalized method of moments (GMM), limited-information maximum likelihood (LIML), and k-class estimators. It also implement a number of tests. Below some examples.

- The test of overidentifying restrictions. The joint null hypothesis is that the instruments are valid instruments, i.e., uncorrelated with the error term, and that the excluded instruments are correctly excluded from the estimated equation. Under the null, the test statistic is distributed as chi-squared in the number of overidentifying restrictions. A rejection casts doubt on the validity of the instruments. For the efficient GMM estimator, the test statistic is Hansen's J statistic[2], the minimised value of the GMM criterion function and saved in the output as `e(j), e(jp), e(jdf)` (see below). For the 2SLS estimator, the test statistic is Sargan's statistic[3], typically calculated as NR^2 from a regression of the IV residuals on the full set of instruments and saved in the output as `e(sargan), e(sarganp), e(sargandf)`. The J statistic is consistent in the presence of heteroskedasticity and autocorrelation; Sargan's statistic is consistent if the disturbance is homoskedastic. Hence, with

[1] A good survey of these estimation methods applied to dynamic panel is Bond S.R. (2002), Dynamic panel data models: a guide to micro data methods and practice, WP Nuffield College, Oxford and Institute for Fiscal Studies.
[2] Hansen L. P. (1982) Large Sample Properties of Generalised Method of Moments Estimators, Econometrica, 50, 1029-54.
[3] Sargan, J.D. (1958) The Estimation of Economic Relationships Using Instrumental Variables, Econometrica, 26, 393-415; (1998) Testing for misspecification after estimating using instrumental variables, in Maasoumi E. (ed.) Contributions to econometrics: John Denis Sargan, vol. 1, Cambridge University Press.

robust or cluster(.) options, Hansen's J statistic is reported; under the assumption of conditional homoskedasticity, Hansen's J statistic becomes Sargan's statistic[4].

- Anderson canonical correlations likelihood-ratio test of whether the equation is identified, i.e., that the excluded instruments are relevant.[5] This test assumes the regressors are distributed as multivariate normal and is saved in the output as e(idstat), e(idp), e(iddf). The null hypothesis of the test is that the matrix of reduced form coefficients has rank=K-1, where K is the number of regressors, i.e, that the equation is underidentified. Under the null of underidentification, the statistic is distributed as chi-squared with degrees of freedom=(L-K+1), where L is the number of instruments (included+excluded). The statistic provides a measure of instrument relevance, and rejection of the null indicates that the model is identified. Note: a result of rejection of the null should be treated with caution, because weak instrument problems may still be present.
- Various identification statistics, if the option ffirst is used, saved in the matrix e(first). For example: the Shea's (1997) "partial R-squared"[6] that measures the instrument relevance taking into accout intercorrelations among instruments; the more common form of "partial R-squared", given by the squared partial correlation between the excluded instruments and the endogenous regressor; the F-test of the excluded instruments that corresponds to the latter R-squared measure. The two measures of "partial R-squared" coincide when the model has only one endogenous regressor. If heteroskedastic-robust standard errors are used, the F-test is also heteroskedastic-robust.

xtabond2 (see D. Roodman (2005) xtabond2: stata module to extend xtabond dynamic panel data estimator, Center for Global Development, Washington, DC, http://econpapers.repec.org/software/bocbocode/s435901.htm) offers the following opportunities.

- "Difference GMM" (GMM-dif) is the Arellano-Bond (1991)[7] GMM estimator that treats the model as a system of equations, one for each time period; the equation differ only in their instrument/moment condition sets. The predetermined [correlated with past errors, but not with current and future errors: $E(\varepsilon_{it}|x_{is})=0\ \forall\ t \geq s$ and $E(\varepsilon_{it}|x_{is})\neq 0\ \forall\ t < s$, $\forall$ i=1,...,N, $\forall$ t, s=1,...,T] and endogenous [correlated with past and present errors, but not with current errors: $E(\varepsilon_{it}|x_{is})=0\ \forall\ t > s$ and $E(\varepsilon_{it}|x_{is})\neq 0\ \forall\ t \leq s$, $\forall$ i=1,...,N, $\forall$ t, s=1,...,T] variables in first differences are instrumented with suitable lags of their own levels. Strictly exogenous regressors [not correlated with errors in all temporal periods: $E(\varepsilon_{it}|x_{is})=0$, $\forall$ i=1,...,N, $\forall$ t, s=1,...,T;], as well as any other instruments, can enter the instrument matrix in the conventional instrumental variables fashion: in first differences, with one column per instrument.
- "System GMM" (GMM-sys) is the augmented version of GMM outlined in Arellano and Bover (1995) and fully developed in Blundell and Bond (1998) who more precisely articulated the necessary assumptions for this augmented estimator and tested it with Monte Carlo simulations.[8] Lagged levels are often poor instruments for first differences, especially for variables that are close to a random walk. Thus, the original equations in levels can be added to the system, and the additional moment conditions could increase efficiency. In these equations, predetermined and endogenous variables in levels are instrumented with suitable lags of their own first differences.
- The updated version of xtabond2 also allows for "Level GMM" (GMM-lev) in which IVs in first-differences or IVs in levels can be used. Pay attention that IVs in levels are not valid under the standard assumptions of dynamic panel data model.

---

4 See Hayashi F. (2000) Econometrics, Princeton University Press, pp. 227-228, 407, and 417.

5 Hall, A. R., G. D. Rudebusch, and D. W. Wilcox (1996) Judging instrument relevance in instrumental variables estimation, International Economic Review, 37, 2, 283-298.

6 Shea, J (1997) Instrument relevance in multivariate linear models: a simple measure, Review of Economics and Statistics, 49, 2, 348-352.

7 Arellano, M. and S. Bond (1991), Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, The Review of Economic Studies, 58, 277-97.

8 Arellano, M. and O. Bover (1995), Another look at the instrumental variable estimation of error-components models, Journal of Econometrics, 68, 29-51. Blundell, R. and S. Bond (1998), Initial conditions and moment restrictions in dynamic panel data models, Journal of Econometrics, 87, 115-43.

- Of course, many tests are included. The Arellano-Bond test for autocorrelation, saved in the output as `e(ari), e(arip), i=1,2,...`, and applied to the first-difference equation residuals: AR(1) is expected in first differences, while higher-order autocorrelation indicates that some lags of the dependent variable, which might be used as instruments, are in fact endogenous, thus bad instruments. The Sargan or Hansen J statistics, saved in the output as `e(sargan), e(sarganp), e(sar_df)`, to perform tests of over-identifying restrictions -whether the instruments, as a group, appear exogenous. For one-step non-robust estimation, the Sargan statistic is reported, which is the minimized value of the one-step GMM criterion function. For one-step robust estimation and for all two-step estimation, the Hansen J statistic is reported, which is the minimized value of the two-step GMM criterion function, and it is robust.

If compared to the original `xtabond` command available since stata 7, `xtabond2` presents many advantages.

- Also the GMM-sys and GMM-lev, and not only the GMM-dif, estimators are available.
- As GMM estimators, we have one- and two-step variants, with two-step estimates asymptotically more efficient.[9]. However, the dependence of the two-step weight matrix on estimated parameters makes asymptotic distribution approximations less reliable in small samples: in the `xtabond` command, the two-step estimates of the standard errors tend to be severely downward biased, or the t-ratios tend to be upward biased. In `xtabond2`, the finite-sample correction for the asymptotic variance of the two-step estimator suggested by Windmeijer (2005)[10] is available.
- There is full control over the instrument matrix, thanks to a syntax that almost completely decouples specification of regressors from specification of instruments (the cost is a loss of parsimony, because most variables used will appear twice).
- The constant term is not included in the first-difference equations in GMM-dif, since this would be equivalent to including time (a trend) as a variable in the levels model[11]. Rather, the constant is differenced out. By default, the constant term is included as a regressor and IV-style instrument in the levels equation, but it can be suppressed with the `noconstant` option.
- It is written to take advantage of the Mata (matrix) programming language first included in Stata 9.0. It works in Stata 7 and 8, but runs much faster in Stata 9 or later. The `xtabond2 nomata` option can be used to prevent the use of Mata even when it is available (this is convenient in large sample because it could exist a difficult trade-off between speed and memory).

2. The data-set

The data-set abdata.dta contains the data used in Arellano-Bond (1991) and Blundell-Bond (1998). The aim is to estimate a model for employment in a panel of companies in UK:
id and *year* are the individual and temporal indicators;
n is the log of employees (*emp*);
k is the log of gross capital stock (*cap*);
ys is the log of output of each industry (*indoutpt*, included in order to capture demand shocks);
w is the log of per-employee real wage (*wage*, deflated by output prices).
Originally, the data have been employed by Layard R.-Nickell S.J. (1986) "Unemployment in Britain, Econometrica", 53, Supplement, 5121-69.

```
use abdata, clear
```

[9] Many applied works use one-step estimates because simulation studies suggest very modest efficiecy gains from two-step, even in presence of heteroskedasticity. See, for example, Blundell R.W., Bond S.R. and Windmeijer F. (2000), Estimation in dynamic panel data models: improving on the perfomance of the standard GMM estimator, in B. Baltagi (ed) *Advances in Econometrics, Volume 15: Non Stationary panels, panel cointegration, and dynamic panels*, JAI Elsevier science.

[10] Windmeijer, F. 2005, A Finite Sample correction for the variance of linear efficient two-step GMM estimators, Journal of Econometrics, 126, 25-51.

[11] This is not possible in `xtabond` and in DPD, the original software that implemented dynamic panel estimators (Doornik, J.A., M. Arellano, and S. Bond (2002) Panel data estimation using DPD for Ox, http://www.nuff.ox.ac.uk/Users/Doornik).

```
descr
Contains data from C:\banchedati\GMM\abdata.dta
  obs:         1,031
 vars:            21                          12 Jan 2005 19:54
 size:        93,821 (99.1% of memory free)
-------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
c1              str9   %9s
year            float  %9.0g
rec             float  %9.0g
yearm1          float  %9.0g
id              float  %9.0g
nL1             float  %9.0g
nL2             float  %9.0g
wL1             float  %9.0g
kL1             float  %9.0g
kL2             float  %9.0g
ysL1            float  %9.0g
ysL2            float  %9.0g
ind             int    %8.0g                  industry
emp             float  %9.0g                  employment
wage            float  %9.0g                  real wage
cap             float  %9.0g                  gross capital stock
indoutpt        float  %9.0g                  industry output
n               float  %9.0g                  log(employment)
w               float  %9.0g                  log(real wage)
k               float  %9.0g                  log(gross capital stock)
ys              float  %9.0g                  log(industry output)
-------------------------------------------------------------------------------
Sorted by:  id  year
```

Variables like *nL1* are just lags, generated, for example, as:

```
g nL1=l.n
(140 missing values generated)
g nL2=l2.n
(280 missing values generated)
```

This in case some commands do not understand time-series operator.

The data-set is unbalanced.

```
. tsset id year
       panel variable:  id, 1 to 140
        time variable:  year, 1976 to 1984

. xtdes, pattern(1000)

      id:  1, 2, ..., 140                                  n =        140
    year:  1976, 1977, ..., 1984                           T =          9
           Delta(year) = 1; (1984-1976)+1 = 9
           (id*year uniquely identifies each observation)

Distribution of T_i:   min      5%     25%       50%       75%     95%     max
                         7       7       7         7         8       9       9

     Freq.  Percent    Cum. |  Pattern
 ---------------------------+-----------
       62     44.29   44.29 |  1111111..
       39     27.86   72.14 |  .1111111.
       19     13.57   85.71 |  .11111111
       14     10.00   95.71 |  111111111
        4      2.86   98.57 |  11111111.
        2      1.43  100.00 |  ..1111111
 ---------------------------+-----------
      140    100.00         |  XXXXXXXXX
```

3. Model estimates

The model presented in Arellano-Bond (1991) is:

$$n_{it} = \alpha_1 n_{it-1} + \alpha_2 n_{it-2} + \beta_1 w_{it} + \beta_2 w_{it-1} + \beta_3 k_{it} + \beta_4 k_{it-1} + \beta_5 k_{it-2} + \beta_6 ys_{it} + \beta_7 ys_{it-1} + \beta_8 ys_{it-2} + \lambda_t + \eta_i + \upsilon_{it}$$

The model estimated in Blundell-Bond (1998) is:

$$n_{it} = a_1 n_{it-1} + b_1 w_{it} + b_2 w_{it-1} + b_3 k_{it} + b_4 k_{it-1} + \lambda_t + \eta_i + \upsilon_{it}$$

The procedure dtime.ado allows for generating the temporal dummies used to estimate λ_t, a time effect that is common to all the companies.

```
. dtime year 1976 1984
```

The following options influence the behaviour of Stata in running complex models and the behaviour of the Mata code.
`set matsize` sets the maximum number of variables that can be included in any of Stata's estimation commands. For Stata/SE the default value is 400, but it may be changed upward or downward, with upper limit equal to 11,000.
`mata query` shows the values of Mata's system parameters.
`mata set` sets the value of the system parameters. For example, `mata set matacache` specifies the maximum amount of memory, in kilobytes, that may be consumed before Mata starts looking to drop autoloaded functions that are not currently being used. The default value is 400, meaning 400 kilobytes. This parameter affects the efficiency with which Stata runs. Larger values cannot hurt, but once matacache is large enough, larger values will not improve performance. As another example, `mata set matafavor` specifies whether, when executing code, Mata should favor conserving memory (`space`) or running quickly (`speed`). The default setting is `space`. Switching to `speed` will make Mata, in a few instances, run a little quicker but consume more memory.

```
. *set matsize 7000
. mata: mata query
. *mata: mata set matacache 900
. mata: mata set matafavor speed
. *mata: mata set matafavor space
```

3.1. OLS estimates: Arellano-Bond (1991), Table 5, col. (g), p. 292

```
. reg n l(1/2).n l(0/1).w l(0/2).(k ys) tau1979-tau1984, cluster(id)

Linear regression                                      Number of obs =      751
                                                       F( 16,    139) =13990.88
                                                       Prob > F      =   0.0000
                                                       R-squared     =   0.9944
Number of clusters (id) = 140                          Root MSE      =   .10158

------------------------------------------------------------------------------
             |               Robust
           n |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   1.044643   .0517969    20.17   0.000     .9422313    1.147055
         L2. |  -.0765426   .0488082    -1.57   0.119    -.1730451    .0199598
           w |
         --. |  -.5236727   .1740911    -3.01   0.003    -.8678817   -.1794637
         L1. |   .4767538   .1717904     2.78   0.006     .1370937    .8164139
           k |
         --. |   .3433951    .048649     7.06   0.000     .2472074    .4395829
         L1. |  -.2018991   .0650327    -3.10   0.002    -.3304803    -.073318
```

```
         L2. |  -.1156467   .0358966    -3.22   0.002    -.1866206   -.0446727
          ys |
         --. |   .4328752     .17894     2.42   0.017      .079079    .7866715
         L1. |  -.7679125   .2514336    -3.05   0.003    -1.265041   -.2707836
         L2. |   .3124721   .1322678     2.36   0.020     .0509551    .5739891
     tau1979 |   .0158888   .0090408     1.76   0.081    -.0019865    .0337641
     tau1980 |   .0219933   .0149899     1.47   0.145    -.0076444    .0516309
     tau1981 |  -.0221532   .0242324    -0.91   0.362    -.0700648    .0257585
     tau1982 |  -.0150344   .0214242    -0.70   0.484    -.0573938    .0273251
     tau1983 |   .0073931     .01963     0.38   0.707    -.0314189    .0462052
     tau1984 |   .0153956   .0204269     0.75   0.452     -.024992    .0557832
       _cons |   .2747256   .3194854     0.86   0.391    -.3569538     .906405
------------------------------------------------------------------------------

. abar, lags(2)
Arellano-Bond test for AR(1): z =   1.29  Pr > z = 0.1978
Arellano-Bond test for AR(2): z =  -1.03  Pr > z = 0.3039

.
. ivreg2 n l(1/2).n l(0/1).w l(0/2).(k ys) tau1979-tau1984, cluster(id) small

OLS regression with robust standard errors
------------------------------------------

Number of clusters (id) = 140                         Number of obs =      751
                                                      F( 16,   139) = 13990.88
                                                      Prob > F      =   0.0000
Total (centered) SS     =  1350.891752                Centered R2   =   0.9944
Total (uncentered) SS   =  2122.555626                Uncentered R2 =   0.9964
Residual SS             =  7.573781642                Root MSE      =    .1016

------------------------------------------------------------------------------
             |               Robust
           n |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   1.044643   .0517969    20.17   0.000     .9422313    1.147055
         L2. |  -.0765426   .0488082    -1.57   0.119    -.1730451    .0199598
           w |
         --. |  -.5236727   .1740911    -3.01   0.003    -.8678817   -.1794637
         L1. |   .4767538   .1717904     2.78   0.006     .1370937    .8164139
           k |
         --. |   .3433951    .048649     7.06   0.000     .2472074    .4395829
         L1. |  -.2018991   .0650327    -3.10   0.002    -.3304803    -.073318
         L2. |  -.1156467   .0358966    -3.22   0.002    -.1866206   -.0446727
          ys |
         --. |   .4328752     .17894     2.42   0.017      .079079    .7866715
         L1. |  -.7679125   .2514336    -3.05   0.003    -1.265041   -.2707836
         L2. |   .3124721   .1322678     2.36   0.020     .0509551    .5739892
     tau1979 |   .0158888   .0090408     1.76   0.081    -.0019865    .0337641
     tau1980 |   .0219933   .0149899     1.47   0.145    -.0076444    .0516309
     tau1981 |  -.0221532   .0242324    -0.91   0.362    -.0700648    .0257585
     tau1982 |  -.0150344   .0214242    -0.70   0.484    -.0573938    .0273251
     tau1983 |   .0073931     .01963     0.38   0.707    -.0314189    .0462052
     tau1984 |   .0153956   .0204269     0.75   0.452     -.024992    .0557832
       _cons |   .2747256   .3194854     0.86   0.391    -.3569538     .906405
------------------------------------------------------------------------------

. abar, lags(2)
Arellano-Bond test for AR(1): z =   1.29  Pr > z = 0.1978
Arellano-Bond test for AR(2): z =  -1.03  Pr > z = 0.3039

.
. ivreg2 n l(1/2).n l(0/1).w l(0/2).(k ys) tau1979-tau1984, cluster(id)

OLS regression with robust standard errors
------------------------------------------

Number of clusters (id) = 140                         Number of obs =      751
                                                      F( 16,   139) = 13990.88
```

```
                                                      Prob > F      =   0.0000
Total (centered) SS     =  1350.891752                Centered R2   =   0.9944
Total (uncentered) SS   =  2122.555626                Uncentered R2 =   0.9964
Residual SS             =  7.573781642                Root MSE      =    .1004

------------------------------------------------------------------------------
             |               Robust
           n |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   1.044643   .0510581    20.46   0.000     .9445709    1.144715
         L2. |  -.0765426    .048112    -1.59   0.112    -.1708404    .0177551
           w |
         --. |  -.5236727   .1716079    -3.05   0.002     -.860018   -.1873274
         L1. |   .4767538     .16934     2.82   0.005     .1448535    .8086542
           k |
         --. |   .3433951   .0479551     7.16   0.000     .2494048    .4373854
         L1. |  -.2018991   .0641051    -3.15   0.002    -.3275428   -.0762555
         L2. |  -.1156467   .0353846    -3.27   0.001    -.1849992   -.0462942
          ys |
         --. |   .4328752   .1763877     2.45   0.014     .0871617    .7785888
         L1. |  -.7679125   .2478472    -3.10   0.002    -1.253684   -.2821409
         L2. |   .3124721   .1303812     2.40   0.017     .0569297    .5680146
     tau1979 |   .0158888   .0089119     1.78   0.075    -.0015781    .0333557
     tau1980 |   .0219933   .0147761     1.49   0.137    -.0069673    .0509538
     tau1981 |  -.0221532   .0238867    -0.93   0.354    -.0689703    .0246639
     tau1982 |  -.0150344   .0211186    -0.71   0.477    -.0564261    .0263574
     tau1983 |   .0073931     .01935     0.38   0.702    -.0305322    .0453185
     tau1984 |   .0153956   .0201355     0.76   0.445    -.0240693    .0548605
       _cons |   .2747256   .3149284     0.87   0.383    -.3425227    .8919739
------------------------------------------------------------------------------

. abar, lags(2)
Arellano-Bond test for AR(1): z =   1.47  Pr > z = 0.1421
Arellano-Bond test for AR(2): z =  -1.03  Pr > z = 0.3036

.
. xtabond2 n l(1/2).n l(0/1).w l(0/2).(k ys) tau1979-tau1984, iv(l(1/2).n l(0/1).w l(0/2).(k
ys) tau1979-tau1984, eq(level)) arle robust
Favoring speed over space. To switch, type or click on mata: mata set matafavor space.

Arellano-Bond dynamic panel-data estimation, one-step system GMM results
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =       751
Time variable : year                            Number of groups   =       140
Number of instruments = 17                      Obs per group: min =         5
Wald chi2(16) = 230379.33                                      avg =      5.36
Prob > chi2   =     0.000                                      max =         7
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   1.044643   .0510581    20.46   0.000     .9445709    1.144715
         L2. |  -.0765426    .048112    -1.59   0.112    -.1708404    .0177551
           w |
         --. |  -.5236727   .1716079    -3.05   0.002     -.860018   -.1873274
         L1. |   .4767538     .16934     2.82   0.005     .1448535    .8086542
           k |
         --. |   .3433951   .0479551     7.16   0.000     .2494048    .4373854
         L1. |  -.2018991   .0641051    -3.15   0.002    -.3275428   -.0762555
         L2. |  -.1156467   .0353846    -3.27   0.001    -.1849992   -.0462942
          ys |
         --. |   .4328754   .1763877     2.45   0.014     .0871618     .778589
         L1. |  -.7679126   .2478473    -3.10   0.002    -1.253684   -.2821409
         L2. |   .3124721   .1303812     2.40   0.017     .0569297    .5680146
     tau1979 |   .0158888   .0089119     1.78   0.075    -.0015781    .0333557
     tau1980 |   .0219933   .0147761     1.49   0.137    -.0069673    .0509538
     tau1981 |  -.0221532   .0238867    -0.93   0.354    -.0689703    .0246639
     tau1982 |  -.0150344   .0211186    -0.71   0.477    -.0564261    .0263574
```

```
     tau1983 |   .0073931     .01935     0.38   0.702    -.0305322    .0453185
     tau1984 |   .0153956   .0201355     0.76   0.445    -.0240693    .0548605
       _cons |   .2747254   .3149284     0.87   0.383    -.3425228    .8919737
------------------------------------------------------------------------------
Hansen test of overid. restrictions: chi2(0) =   0.00     Prob > chi2 =      .

Arellano-Bond test for AR(1) in levels:           z =   1.47  Pr > z =  0.142
Arellano-Bond test for AR(2) in levels:           z =  -1.03  Pr > z =  0.304
------------------------------------------------------------------------------

. predict yhat
(280 missing values generated)

. qui corr yhat n if e(sample)

. scalar XTAB_r2=r(rho)^2

. drop yhat

. di XTAB_r2
.99439349
```

Considerations.

- Years 1976 and 1977 are lost due to the use of 2 lags as explanatory variables. the dummy for 1978 is not included because we leave the constant in the model. The inclusion of right dummies (1979-1984, especially 1979) is fundamental in order to obtain the same results as Arellano-Bond paper.
- By default, both `ivreg2` and `xtabond2` use asymptotic standard errors, the same reported by Arellano-Bond. In order to reproduce the same standard errors as `reg`, we need the `small` option. This allows for coincidence between `reg` and `ivreg2` also when the options `robust` or `cluster(.)` are used[12]. On the contrary, `xtabond2`, `robust` ignores `small`. NOTE: `xtabond2` does not have option `cluster(.)`: `robust` in `xtabond2` is equal to `cluster(.)` in `ivreg2`.[13] `Cluster` is the option used by Arellano-Bond.
- The test for second-order (or higher order) serial correlation of the residuals is automatically done by `xtabond2`; option `arlevels` specifies that the autocorrelation tests should be applied to the residuals from the levels, not first-difference, equations. In the `reg` and `ivreg2` cases, the procedure `abar.ado`, written by D. Roodman, Center for Global Development, Washington, DC, is tyo be used. This procedure, originally proposed for a particular linear GMM dynamic panel data estimator, can be applied to linear GMM regressions in general, and thus to OLS and 2SLS regressions, which can be seen as special cases of linear GMM. It is appropriate for both time-series and cross-section time-series (panel) regressions. It is consistent in the presence of various patterns of error covariance when the options `robust` or `cluster` are used.
- In `xtabond2` the option `ivstyle()` specifies a set of variables to serve as standard instruments, with one column in the instrument matrix per variable. Normally, strictly exogenous regressors are included in `ivstyle` options in order to enter the instrument matrix, as well as being listed before the main comma of the command line. The suboption `eq(.)` specifies which set of equation(s) should use the instruments: first-difference only (`equation(diff)`), levels only

---

[12] In `ivreg2` and `xtabond2` the MSE is, by default, defined as the squared root of RSS/NT; if `small` is specified, it is the squared root of RSS/(NT-K). With `robust` and `cluster` there is, by default, no degrees of freedom adjustment. If `small robust` is specified, the finite sample adjustement to the robust variance-covariance matrix is NT/(NT-K). If `small cluster` is specified, the finite sample adjustement to the robust variance-covariance matrix is (NT-1)/(NT-K)× N_c/(N_c-1), where N_c is the total number of clusters.

[13] Summarising, we have the following combinations:

`reg = ivreg2, small = xtabond2, small`
`reg, robust = ivreg2, small robust` and `reg, cluster = ivreg2, small cluster`
`ivreg2, robust`
`ivreg2, cluster = xtabond2, robust`

(`equation(level)`), or both (`equation(both)`, the default). By default, the instruments are first-differenced for use in the first-difference equations and taken as is for instrumenting the levels equations. The `eq(.)` option can be useful when x is predetermined: since it is assumed to be uncorrelated with the contemporaneous and future error term, it is a valid instrument for the levels equations; however, x becomes endogenous in first differences, so Δx is not a valid instrument for the first-difference equations.

- In `xtabond2` the R-squared is not available. Thus, we compute it as the squared correlation coefficient between actual and fitted values.

3.2. Within estimates: Arellano-Bond (1991), Table 5, col. (h), p. 292

```
. xtreg n l(1/2).n l(0/1).w l(0/2).(k ys) tau1979-tau1984, fe cluster(id)

Fixed-effects (within) regression               Number of obs      =       751
Group variable (i): id                          Number of groups   =       140

R-sq:  within  = 0.7973                         Obs per group: min =         5
       between = 0.9809                                        avg =       5.4
       overall = 0.9758                                        max =         7

                                                F(16,139)          =    152.18
corr(u_i, Xb)  = 0.5459                         Prob > F           =    0.0000

                                 (Std. Err. adjusted for 140 clusters in id)
------------------------------------------------------------------------------
             |               Robust
           n |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   .7329476   .0596831    12.28   0.000     .6149436    .8509516
         L2. |  -.1394773   .0781564    -1.78   0.077    -.2940065    .0150519
           w |
         --. |  -.5597445   .1596195    -3.51   0.001    -.8753406   -.2441484
         L1. |   .3149987   .1430587     2.20   0.029     .0321463    .5978511
           k |
         --. |   .3884188    .056928     6.82   0.000      .275862    .5009756
         L1. |  -.0805185   .0538774    -1.49   0.137    -.1870436    .0260066
         L2. |  -.0278013   .0426222    -0.65   0.515    -.1120728    .0564703
          ys |
         --. |    .468666   .1712492     2.74   0.007     .1300759    .8072561
         L1. |  -.6285587   .2066106    -3.04   0.003    -1.037065   -.2200527
         L2. |   .0579764   .1326758     0.44   0.663    -.2043473    .3203001
     tau1979 |   .0046562   .0092572     0.50   0.616    -.0136469    .0229593
     tau1980 |   .0112327   .0158528     0.71   0.480    -.0201111    .0425765
     tau1981 |  -.0253693   .0249902    -1.02   0.312    -.0747794    .0240408
     tau1982 |  -.0343973    .022882    -1.50   0.135    -.0796391    .0108444
     tau1983 |  -.0280344   .0262074    -1.07   0.287     -.079851    .0237822
     tau1984 |  -.0119152   .0281652    -0.42   0.673    -.0676028    .0437723
       _cons |    1.79212    .633652     2.83   0.005     .5392775    3.044963
-------------+----------------------------------------------------------------
     sigma_u |  .22568151
     sigma_e |  .09395847
         rho |  .85227336   (fraction of variance due to u_i)
------------------------------------------------------------------------------

. *qui within id n nL1 nL2 w wL1 k kL1 kL2 ys ysL1 ysL2 tau1979 tau1980 tau1981 tau1982
tau1983 tau1984 if nL2!=.
. qui within id n nL1 nL2 w wL1 k kL1 kL2 ys ysL1 ysL2 tau1979 tau1980 tau1981 tau1982 tau1983
tau1984 if e(sample)

.
. ivreg2 n_WD nL1_WD nL2_WD w_WD wL1_WD k_WD kL1_WD kL2_WD ys_WD ysL1_WD ysL2_WD tau1979_WD-
tau1984_WD, small noconstant cluster(id)

OLS regression with robust standard errors
-----------------------------------------
```

```
Number of clusters (id) = 140                          Number of obs =      751
                                                       F( 16,   139) =   152.38
                                                       Prob > F      =   0.0000
Total (centered) SS     =  25.91406328                 Centered R2   =   0.7973
Total (uncentered) SS   =  25.91406328                 Uncentered R2 =   0.7973
Residual SS             =  5.252775373                 Root MSE      =   .08454

------------------------------------------------------------------------------
             |               Robust
        n_WD |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      nL1_WD |   .7329477   .0596424    12.29   0.000      .615024    .8508714
      nL2_WD |  -.1394773   .0781033    -1.79   0.076    -.2939013    .0149468
        w_WD |  -.5597445   .1595108    -3.51   0.001    -.8751258   -.2443632
       wL1_WD |   .3149987   .1429613     2.20   0.029     .0323388    .5976586
        k_WD |   .3884188   .0568893     6.83   0.000     .2759386     .500899
      kL1_WD |  -.0805185   .0538407    -1.50   0.137    -.1869711    .0259341
      kL2_WD |  -.0278013   .0425932    -0.65   0.515    -.1120155     .056413
       ys_WD |    .468666   .1711327     2.74   0.007     .1303063    .8070257
     ysL1_WD |  -.6285587     .20647    -3.04   0.003    -1.036787   -.2203307
     ysL2_WD |   .0579764   .1325855     0.44   0.663    -.2041688    .3201216
  tau1979_WD |   .0046562   .0092509     0.50   0.616    -.0136344    .0229468
  tau1980_WD |   .0112327    .015842     0.71   0.479    -.0200898    .0425552
  tau1981_WD |  -.0253693   .0249732    -1.02   0.311    -.0747458    .0240072
  tau1982_WD |  -.0343973   .0228664    -1.50   0.135    -.0796083    .0108136
  tau1983_WD |  -.0280344   .0261895    -1.07   0.286    -.0798157    .0237469
  tau1984_WD |  -.0119152    .028146    -0.42   0.673    -.0675649    .0437344
------------------------------------------------------------------------------

.
. ivreg2 n_WD nL1_WD nL2_WD w_WD wL1_WD k_WD kL1_WD kL2_WD ys_WD ysL1_WD ysL2_WD tau1979_WD-
tau1984_WD, noconstant cluster(id)

OLS regression with robust standard errors
------------------------------------------

Number of clusters (id) = 140                          Number of obs =      751
                                                       F( 16,   139) =   152.38
                                                       Prob > F      =   0.0000
Total (centered) SS     =  25.91406328                 Centered R2   =   0.7973
Total (uncentered) SS   =  25.91406328                 Uncentered R2 =   0.7973
Residual SS             =  5.252775373                 Root MSE      =   .08363

------------------------------------------------------------------------------
             |               Robust
        n_WD |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      nL1_WD |   .7329477   .0588318    12.46   0.000     .6176395    .8482558
      nL2_WD |  -.1394773   .0770417    -1.81   0.070    -.2904761    .0115216
        w_WD |  -.5597445   .1573427    -3.56   0.000    -.8681306   -.2513584
      wL1_WD |   .3149987   .1410181     2.23   0.025     .0386083    .5913891
        k_WD |   .3884188    .056116     6.92   0.000     .2784334    .4984042
      kL1_WD |  -.0805185   .0531089    -1.52   0.129    -.1846099     .023573
      kL2_WD |  -.0278013   .0420142    -0.66   0.508    -.1101476    .0545451
       ys_WD |    .468666   .1688066     2.78   0.005     .1378112    .7995208
     ysL1_WD |  -.6285587   .2036636    -3.09   0.002    -1.027732   -.2293853
     ysL2_WD |   .0579764   .1307834     0.44   0.658    -.1983543    .3143072
  tau1979_WD |   .0046562   .0091251     0.51   0.610    -.0132287    .0225411
  tau1980_WD |   .0112327   .0156267     0.72   0.472    -.0193951    .0418605
  tau1981_WD |  -.0253693   .0246338    -1.03   0.303    -.0736506     .022912
  tau1982_WD |  -.0343973   .0225556    -1.53   0.127    -.0786055    .0098108
  tau1983_WD |  -.0280344   .0258335    -1.09   0.278    -.0786672    .0225984
  tau1984_WD |  -.0119152   .0277634    -0.43   0.668    -.0663305    .0425001
------------------------------------------------------------------------------

.
. xtabond2 n_WD nL1_WD nL2_WD w_WD wL1_WD k_WD kL1_WD kL2_WD ys_WD ysL1_WD ysL2_WD tau1979_WD-
tau1984_WD, iv(nL1_WD nL2_WD w_WD wL1_WD k_WD kL1_WD kL2_WD ys_WD ysL1_WD ysL2_WD tau1979_WD-
tau1984_WD, eq(level)) noconstant arle robust
```

```
Favoring speed over space. To switch, type or click on mata: mata set matafavor space.

Arellano-Bond dynamic panel-data estimation, one-step system GMM results
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =       751
Time variable : year                            Number of groups   =       140
Number of instruments = 16                      Obs per group: min =         5
Wald chi2(16) =   2505.81                                      avg =      5.36
Prob > chi2   =     0.000                                      max =         7
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      nL1_WD |   .7329477   .0588318    12.46   0.000     .6176395    .8482558
      nL2_WD |  -.1394773   .0770417    -1.81   0.070    -.2904761    .0115216
        w_WD |  -.5597445   .1573427    -3.56   0.000    -.8681306   -.2513584
       wL1_WD |   .3149987   .1410181     2.23   0.025     .0386083    .5913891
        k_WD |   .3884188    .056116     6.92   0.000     .2784334    .4984042
      kL1_WD |  -.0805185   .0531089    -1.52   0.129    -.1846099     .023573
      kL2_WD |  -.0278013   .0420142    -0.66   0.508    -.1101476    .0545451
       ys_WD |    .468666   .1688066     2.78   0.005     .1378112    .7995208
     ysL1_WD |  -.6285587   .2036636    -3.09   0.002    -1.027732   -.2293853
     ysL2_WD |   .0579764   .1307834     0.44   0.658    -.1983543    .3143072
  tau1979_WD |   .0046562   .0091251     0.51   0.610    -.0132287    .0225411
  tau1980_WD |   .0112327   .0156267     0.72   0.472    -.0193951    .0418605
  tau1981_WD |  -.0253693   .0246338    -1.03   0.303    -.0736506     .022912
  tau1982_WD |  -.0343973   .0225556    -1.53   0.127    -.0786055    .0098108
  tau1983_WD |  -.0280344   .0258335    -1.09   0.278    -.0786672    .0225984
  tau1984_WD |  -.0119152   .0277634    -0.43   0.668    -.0663305    .0425001
------------------------------------------------------------------------------
Hansen test of overid. restrictions: chi2(0) =   0.00     Prob > chi2 =      .

Arellano-Bond test for AR(1) in levels:           z =  -3.87  Pr > z =  0.000
Arellano-Bond test for AR(2) in levels:           z =  -4.68  Pr > z =  0.000
------------------------------------------------------------------------------

. predict yhat
(280 missing values generated)

. qui corr yhat n_WD if e(sample)

. scalar XTAB_r2=r(rho)^2

. drop yhat

. di XTAB_r2
.7973002
```

Considerations.

- In Stata 9 `xtreg` allows for robust/cluster standard errors and it understands time-series operators.
- `ivreg2` and `xtabond2` require within data transformation (`within.ado` procedure). Note that the "right" sample must be used, either by imposing a condition on the data availability or by imposing to use the same estimation sample of `xtreg`. Note also that all the variables, temporal dummies included, must be transformed. When using `cluster`, is the downwards bias in the standard errors relevant as seen in lecture_panel_static_application?[14] Hint: try to run the same estimation commands with the option `robust`.

[14] With `small` option, the MSE is defined as the squared root of RSS/(NT-K). If `cluster` is specified, the MSE is obtained by a finite sample adjustement to the variance-covariance matrix: RSS/NT × (NT-1)/(NT-K) × N_c/(N_c-1), where N_c is the total number of clusters. In this case, we need a correction different from the ones used in lecture_panel_static_application. In `ivreg2` without the option `small` and in `xtabond2` the standard errors are asymptotic; thus, the MSE is defined as the squared root of RSS/NT, both without and with `cluster` option, because cluster does not creates any degrees of freedom adjustment.

- The abar procedure does not work in this case, and the output of autocorrelation test reported by xtabond2 must to be not considered. This test is not appropriate for fixed-effects regressions for dynamic models, assuming those are done via a mean-deviation transformation. This is because the Arellano-Bond AR(p) test assumes that right-hand-side variables are not "post-determined" i.e. not correlated with future errors. In a dynamic setting, future values of regressors can depend on future errors. And after the mean-deviations transformation, future values of the original regressors affect current values of the transformed versions.

3.3. First-differences estimates

```
. xtivreg n (l(1/2).n l(0/1).w l(0/2).(k ys) tau1979-tau1984 = l(1/2).n l(0/1).w l(0/2).(k ys)
tau1979-tau1984), fd noconst

First-differenced IV regression                 Number of obs      =       611
Group variable: id                              Number of groups   =       140

R-sq:  within  = 0.1603                         Obs per group: min =         4
       between = 0.5014                                        avg =       4.4
       overall = 0.8637                                        max =         6

                                                chi2(16)           =    699.56
corr(u_i, Xb)  = 0.7056                         Prob > chi2        =    0.0000

------------------------------------------------------------------------------
         d.n |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         LD. |   .1218396   .0434764     2.80   0.005     .0366275    .2070517
        L2D. |  -.0556518   .0419881    -1.33   0.185     -.137947    .0266434
           w |
         D1. |  -.5551606    .055246   -10.05   0.000    -.6634407   -.4468804
         LD. |   .1238423   .0637157     1.94   0.052    -.0010381    .2487227
           k |
         D1. |   .3938555   .0307857    12.79   0.000     .3335167    .4541943
         LD. |   .1252152   .0328176     3.82   0.000     .0608938    .1895365
        L2D. |    .040182   .0329258     1.22   0.222    -.0243515    .1047154
          ys |
         D1. |   .5692903    .117116     4.86   0.000     .3397473    .7988334
         LD. |  -.3754761   .1282028    -2.93   0.003    -.6267491   -.1242032
        L2D. |   .0272238   .1313197     0.21   0.836     -.230158    .2846056
     tau1979 |
         D1. |  -.0000954   .0125153    -0.01   0.994     -.024625    .0244342
     tau1980 |
         D1. |   .0068434   .0181881     0.38   0.707    -.0288047    .0424914
     tau1981 |
         D1. |  -.0229064   .0265084    -0.86   0.388    -.0748619    .0290491
     tau1982 |
         D1. |  -.0495964   .0317921    -1.56   0.119    -.1119078     .012715
     tau1983 |
         D1. |  -.0591959   .0373131    -1.59   0.113    -.1323283    .0139366
     tau1984 |
         D1. |  -.0529525   .0419136    -1.26   0.206    -.1351017    .0291967
-------------+----------------------------------------------------------------
     sigma_u |  .71755048
     sigma_e |  .10926003
         rho |  .97733982   (fraction of variance due to u_i)
------------------------------------------------------------------------------
Instrumented:    L.n  L2.n  w  L.w  k  L.k  L2.k  ys  L.ys  L2.ys  tau1979  tau1980  tau1981  tau1982
tau1983 tau1984
Instruments:      L.n  L2.n  w  L.w  k  L.k  L2.k  ys  L.ys  L2.ys  tau1979  tau1980  tau1981  tau1982
tau1983 tau1984

. ivreg2 d.n dl.n dl2.n d.w dl.w d.k dl.k dl2.k d.ys dl.ys dl2.ys d.(tau1979 tau1980 tau1981
tau1982 tau1983 tau1984), noconst small

Ordinary Least Squares (OLS) regression
```

```
---------------------------------------

                                                      Number of obs =      611
                                                      F( 16,   595) =    43.72
                                                      Prob > F      =   0.0000
Total (centered) SS     =  12.59997848                Centered R2   =   0.4363
Total (uncentered) SS   =  15.45414725                Uncentered R2 =   0.5404
Residual SS             =  7.102963775                Root MSE      =    .1093

------------------------------------------------------------------------------
         D.n |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         LD. |   .1218396   .0434764     2.80   0.005     .0364538    .2072254
        L2D. |  -.0556518   .0419881    -1.33   0.186    -.1381148    .0268111
           w |
         D1. |  -.5551606    .055246   -10.05   0.000    -.6636614   -.4466597
         LD. |   .1238423   .0637157     1.94   0.052    -.0012926    .2489773
           k |
         D1. |   .3938555   .0307857    12.79   0.000     .3333937    .4543173
         LD. |   .1252152   .0328176     3.82   0.000     .0607627    .1896676
        L2D. |    .040182   .0329258     1.22   0.223     -.024483     .104847
          ys |
         D1. |   .5692903    .117116     4.86   0.000     .3392794    .7993013
         LD. |  -.3754761   .1282028    -2.93   0.004    -.6272613    -.123691
        L2D. |   .0272238   .1313197     0.21   0.836    -.2306826    .2851302
     tau1979 |
         D1. |  -.0000954   .0125153    -0.01   0.994     -.024675    .0244842
     tau1980 |
         D1. |   .0068434   .0181881     0.38   0.707    -.0288773     .042564
     tau1981 |
         D1. |  -.0229064   .0265084    -0.86   0.388    -.0749678     .029155
     tau1982 |
         D1. |  -.0495964   .0317921    -1.56   0.119    -.1120348     .012842
     tau1983 |
         D1. |  -.0591959   .0373131    -1.59   0.113    -.1324773    .0140856
     tau1984 |
         D1. |  -.0529525   .0419136    -1.26   0.207    -.1352691    .0293641
------------------------------------------------------------------------------

. abar, lags(2)
Arellano-Bond test for AR(1): z =  -1.56  Pr > z = 0.1176
Arellano-Bond test for AR(2): z =  -0.16  Pr > z = 0.8767

. ivreg2 d.n dl.n dl2.n d.w dl.w d.k dl.k dl2.k d.ys dl.ys dl2.ys d.(tau1979 tau1980 tau1981
tau1982 tau1983 tau1984), noconst cluster(id)

OLS regression with robust standard errors
------------------------------------------

Number of clusters (id) = 140                         Number of obs =      611
                                                      F( 16,   139) =    42.36
                                                      Prob > F      =   0.0000
Total (centered) SS     =  12.59997848                Centered R2   =   0.4363
Total (uncentered) SS   =  15.45414725                Uncentered R2 =   0.5404
Residual SS             =  7.102963775                Root MSE      =    .1078

------------------------------------------------------------------------------
             |               Robust
         D.n |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         LD. |   .1218396   .0470986     2.59   0.010     .0295282    .2141511
        L2D. |  -.0556518   .0393147    -1.42   0.157    -.1327073    .0214036
           w |
         D1. |  -.5551606    .160396    -3.46   0.001     -.869531   -.2407901
         LD. |   .1238423   .0690606     1.79   0.073    -.0115139    .2591985
           k |
         D1. |   .3938555   .0577114     6.82   0.000     .2807432    .5069678
         LD. |   .1252152   .0400739     3.12   0.002     .0466719    .2037585
```

```
        L2D. |    .040182   .0294136     1.37   0.172    -.0174675    .0978315
          ys |
         D1. |   .5692903   .1675732     3.40   0.001     .2408529    .8977277
         LD. |  -.3754761   .1494496    -2.51   0.012    -.6683919   -.0825604
        L2D. |   .0272238   .1210112     0.22   0.822    -.2099538    .2644014
     tau1979 |
         D1. |  -.0000954   .0089401    -0.01   0.991    -.0176177    .0174268
     tau1980 |
         D1. |   .0068434   .0167252     0.41   0.682    -.0259374    .0396241
     tau1981 |
         D1. |  -.0229064   .0299926    -0.76   0.445    -.0816907    .0358779
     tau1982 |
         D1. |  -.0495964   .0360871    -1.37   0.169    -.1203258     .021133
     tau1983 |
         D1. |  -.0591959   .0418965    -1.41   0.158    -.1413115    .0229198
     tau1984 |
         D1. |  -.0529525   .0489099    -1.08   0.279    -.1488142    .0429092
------------------------------------------------------------------------------

. abar, lags(2)
Arellano-Bond test for AR(1): z =  -1.51  Pr > z = 0.1305
Arellano-Bond test for AR(2): z =  -0.14  Pr > z = 0.8879

.
. xtabond2 n l(1/2).n l(0/1).w l(0/2).(k ys) tau1979-tau1984, iv(l(1/2).n l(0/1).w l(0/2).(k
ys) tau1979-tau1984, eq(diff)) noleveleq robust
Favoring speed over space. To switch, type or click on mata: mata set matafavor space.

Arellano-Bond dynamic panel-data estimation, one-step difference GMM results
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =       611
Time variable : year                            Number of groups   =       140
Number of instruments = 16                      Obs per group: min =         4
Wald chi2(16) =    699.92                                      avg =      4.36
Prob > chi2   =     0.000                                      max =         6
------------------------------------------------------------------------------
             |              Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   .1218396   .0470986     2.59   0.010     .0295282    .2141511
         L2. |  -.0556518   .0393147    -1.42   0.157    -.1327073    .0214036
           w |
         --. |  -.5551606    .160396    -3.46   0.001     -.869531   -.2407901
         L1. |   .1238423   .0690606     1.79   0.073    -.0115139    .2591985
           k |
         --. |   .3938555   .0577114     6.82   0.000     .2807432    .5069678
         L1. |   .1252152   .0400739     3.12   0.002     .0466719    .2037585
         L2. |    .040182   .0294136     1.37   0.172    -.0174675    .0978315
          ys |
         --. |   .5692903   .1675732     3.40   0.001     .2408529    .8977277
         L1. |  -.3754761   .1494496    -2.51   0.012    -.6683919   -.0825604
         L2. |   .0272238   .1210112     0.22   0.822    -.2099538    .2644014
     tau1979 |  -.0000954   .0089401    -0.01   0.991    -.0176177    .0174268
     tau1980 |   .0068434   .0167252     0.41   0.682    -.0259374    .0396241
     tau1981 |  -.0229064   .0299926    -0.76   0.445    -.0816907    .0358779
     tau1982 |  -.0495964   .0360871    -1.37   0.169    -.1203258     .021133
     tau1983 |  -.0591959   .0418965    -1.41   0.158    -.1413115    .0229198
     tau1984 |  -.0529525   .0489099    -1.08   0.279    -.1488142    .0429092
------------------------------------------------------------------------------
Hansen test of overid. restrictions: chi2(0) =   0.00     Prob > chi2 =      .

Arellano-Bond test for AR(1) in first differences: z =  -1.51  Pr > z =  0.131
Arellano-Bond test for AR(2) in first differences: z =  -0.14  Pr > z =  0.888
------------------------------------------------------------------------------

. predict yhat
(280 missing values generated)

. qui corr yhat d.n if e(sample)
```

```
. scalar XTAB_r2_D=r(rho)^2

. qui corr yhat n if e(sample)

. scalar XTAB_r2_L=r(rho)^2

. drop yhat

. di XTAB_r2_D
.01403893

. di XTAB_r2_L
.86391605
```

Considerations

- Estimating in first-differences implies the lost of another year (1978). Temporal dummy 1979 is included because, by definition, the constant is differenced out. In fact, the inclusion of a constant would be equivalent to including a trend as a variable in the levels model:
  $y_{it} = a + b\ y_{it-1} + c\ x_{it} + \varepsilon_{it} \quad \Rightarrow \Delta y_{it} = b\ \Delta y_{it-1} + c\ \Delta x_{it} + \Delta\varepsilon_{it}$
  $y_{it} = \alpha t + \beta\ y_{it-1} + \gamma\ x_{it} + \eta_{it} \quad \Rightarrow \Delta y_{it} = \alpha + \beta\ \Delta y_{it-1} + \gamma\ \Delta x_{it} + \Delta\eta_{it}$.
  In `xtabond2` with the suboption `eq(diff)` we use equations in first-differences instrumented by first-differences only. Note that it is better to add the option `noleveleq`, otherwise equation in levels are implicitly added (GMM-sys) and the number of observations is 751 rather than 611. With these option (and unlike the old `xtabond`, or DPD, or `xtabond2, eq(level)`, or `xtivreg`, or `ivreg2`, in which the option `noconstant` is to be used) the constant is automatically differenced out.
- Despite it is appropriate to run the autocorrelation tests, `abar.ado` does not work after `xtivreg, fd`. This command understands time-series operators, but does not allows for robust/cluster standard errors even in Stata 9 (but bootstrapping is allowed). It is to be noted that it transforms in first-differences all the variables, exogenous, endogenous, dummies, and the instruments.
- The `ivreg2, small` command perfectly reproduces `xtivreg, fd`. Thus, it can be used to obtain first-differenced estimates with robust and asymptotic standard errors. The same it is true for `xtabond2`. Note that the first-differences transformation has to be esplicitly written in the `ivreg2` command, while it is automatically produced by `xtabond2, eq(diff)` that also keeps the original names of the variables. The command `xtabond2` could also works as the command `ivreg2`: by using the options `eq(lev) arle` and by writing all the variables in first-differences.

General consideration about OLS, FE and FD estimation of a dynamic panel model.

- OLS pooled estimate of lagged dependent variable is upwards biased because the same constant is imposed to all the individuals.
- The F test in the fixed effects estimate without `cluster` option shows that significant firm-specific effects exist.
- However, also fixed effects estimate is inconsistent as $N \rightarrow \infty$ if T is kept fixed.[15] This bias is due to the correlation between lagged dependent variable and error term.
- The bias of estimates is even more evident when first difference transformation are used.

3.4. Anderson-Hsiao (1981)[16] estimates with IV in differences: Arellano-Bond (1991), Table 5, col. (e), p. 292

---

[15] Nickell S. (1981) Biases in dynamic models with fixed effects, Econometrica, 49, 6, 1417-1426.
[16] Anderson T.W., Hsiao C. (1981), Estimation of dynamic models with error com ponents, *Journal of the American Statistical Association*, 76, 598-606.

```
. xtivreg n (l.n l2.n l(0/1).w l(0/2).(k ys) tau1980-tau1984 = l3.n l2.n l(0/1).w l(0/2).(k
ys) tau1980-tau1984), fd noconst

First-differenced IV regression                  Number of obs      =       471
Group variable: id                               Number of groups   =       140

R-sq:  within  = 0.0121                          Obs per group: min =         3
       between = 0.9178                                         avg =       3.4
       overall = 0.9895                                         max =         5

                                                 chi2(15)           =    223.84
corr(u_i, Xb)  = 0.9270                          Prob > chi2        =    0.0000

------------------------------------------------------------------------------
         d.n |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         LD. |   1.422765   1.583053     0.90   0.369    -1.679962    4.525493
        L2D. |  -.1645517   .1647179    -1.00   0.318    -.4873928    .1582894
           w |
         D1. |  -.7524675   .1765733    -4.26   0.000    -1.098545   -.4063902
         LD. |   .9627611   1.086506     0.89   0.376    -1.166752    3.092275
           k |
         D1. |   .3221686   .1466086     2.20   0.028     .0348211    .6095161
         LD. |  -.3248778   .5800599    -0.56   0.575    -1.461774    .8120187
        L2D. |  -.0953947   .1960883    -0.49   0.627    -.4797207    .2889314
          ys |
         D1. |   .7660906    .369694     2.07   0.038     .0415037    1.490678
         LD. |  -1.361881   1.156835    -1.18   0.239    -3.629237    .9054744
        L2D. |   .3212993   .5440403     0.59   0.555        -.745    1.387599
     tau1980 |
         D1. |   .0161204   .0336264     0.48   0.632    -.0497861     .082027
     tau1981 |
         D1. |  -.0251788   .0480382    -0.52   0.600    -.1193319    .0689742
     tau1982 |
         D1. |  -.0399339   .0614387    -0.65   0.516    -.1603515    .0804838
     tau1983 |
         D1. |  -.0418336   .0730074    -0.57   0.567    -.1849255    .1012584
     tau1984 |
         D1. |  -.0366085   .0774638    -0.47   0.637    -.1884349    .1152178
-------------+----------------------------------------------------------------
     sigma_u |  .29195156
     sigma_e |  .18855982
         rho |  .70564904   (fraction of variance due to u_i)
------------------------------------------------------------------------------
Instrumented:    L.n L2.n w L.w k L.k L2.k ys L.ys L2.ys tau1980 tau1981 tau1982 tau1983
tau1984
Instruments:      L3.n L2.n w L.w k L.k L2.k ys L.ys L2.ys tau1980 tau1981 tau1982 tau1983
tau1984

. xtabond2 n l(1/2).n l(0/1).w l(0/2).(k ys) tau1980-tau1984, iv(l(2/3).n l(0/1).w l(0/2).(k
ys) tau1980-tau1984, eq(diff)) robust noleveleq
Favoring speed over space. To switch, type or click on mata: mata set matafavor space.

Arellano-Bond dynamic panel-data estimation, one-step difference GMM results
------------------------------------------------------------------------------
Group variable: id                               Number of obs      =       471
Time variable : year                             Number of groups   =       140
Number of instruments = 15                       Obs per group: min =         3
Wald chi2(15) =    836.36                                       avg =      3.36
Prob > chi2   =     0.000                                       max =         5
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   1.422765   1.001091     1.42   0.155    -.5393375    3.384868
         L2. |  -.1645517   .1276498    -1.29   0.197    -.4147407    .0856372
           w |
         --. |  -.7524675    .229792    -3.27   0.001    -1.202852   -.3020834
```

```
         L1. |   .9627611   .7683295     1.25   0.210    -.5431371    2.468659
           k |
         --. |   .3221686    .104688     3.08   0.002     .1169839    .5273533
         L1. |  -.3248779   .3860559    -0.84   0.400    -1.081534    .4317779
         L2. |  -.0953947   .1234367    -0.77   0.440    -.3373261    .1465368
          ys |
         --. |   .7660906   .3113873     2.46   0.014     .1557827    1.376399
         L1. |  -1.361881   .8814085    -1.55   0.122     -3.08941    .3656478
         L2. |   .3212993   .4156362     0.77   0.440    -.4933326    1.135931
     tau1980 |   .0161204   .0249057     0.65   0.517    -.0326939    .0649348
     tau1981 |  -.0251788    .037007    -0.68   0.496    -.0977112    .0473535
     tau1982 |  -.0399339   .0396221    -1.01   0.314    -.1175918     .037724
     tau1983 |  -.0418336   .0434498    -0.96   0.336    -.1269936    .0433265
     tau1984 |  -.0366085   .0475629    -0.77   0.441    -.1298301     .056613
------------------------------------------------------------------------------
Hansen test of overid. restrictions: chi2(0) =   0.00     Prob > chi2 =      .

Arellano-Bond test for AR(1) in first differences: z =  -1.24  Pr > z =  0.216
Arellano-Bond test for AR(2) in first differences: z =  -0.78  Pr > z =  0.435
------------------------------------------------------------------------------
```

Considerations.

- Instrumenting first-differences of y_{it-1} with Δy_{it-3} [17] implies the lost of another year (1979). The correct specification is to drop the constant and to include the 1980 temporal dummy. Again, the only estimation command in which we do not need to use the option `noconstant` is `xtabond2 eq(diff)`.
- From now on it is important to test H_0: errors not correlated at the second order i.e. dynamics correctly specified. Of course, H_0: errors not correlated at the first order is always rejected because in the first difference equations errors ~ MA(1).

3.5. Anderson-Hsiao (1982)[18] estimates with IV in levels: Arellano-Bond (1991), Table 5, col. (f), p. 292

```
. facum id n
(140 missing values generated)
(140 real changes made)

.  xtivreg  n  (l.n  l2.n  l(0/1).w  l(0/2).(k  ys)  tau1979-tau1984  =  l2.cumnl  l2.n  l(0/1).w
l(0/2).(k ys) tau1979-tau1984), noconst fd

First-differenced IV regression               Number of obs      =       611
Group variable: id                            Number of groups   =       140

R-sq:  within  = 0.0561                       Obs per group: min =         4
       between = 0.8740                                      avg =       4.4
       overall = 0.9615                                      max =         6

                                              chi2(16)           =    133.13
corr(u_i, Xb)  = -0.7735                      Prob > chi2        =    0.0000

------------------------------------------------------------------------------
         d.n |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         LD. |   2.307626   1.999548     1.15   0.248    -1.611416    6.226668
        L2D. |  -.2240271   .1814343    -1.23   0.217    -.5796318    .1315777
           w |
         D1. |  -.8103626   .2653018    -3.05   0.002    -1.330345   -.2903807
```

[17] The IV-set is composed by all the exogenous variables of the system (both included and excluded from the equation of interest): hence, Δy_{it-2}, Δy_{it-3}, plus contemporaneous and lagged first-differences of w, k, ys. Given that Δy_{it-2} is an exogenous included in the equation, the actual IV for y_{it-1} is Δy_{it-3}.

[18] Anderson T.W., Hsiao C. (1982), Formulation and Estimation of dynamic models using panel data, *Journal of Econometrics*, 18, 47-82.

```
         LD. |   1.422246   1.195245     1.19   0.234    -.9203922    3.764884
           k |
         D1. |   .2530975   .1466736     1.73   0.084    -.0343774    .5405725
         LD. |  -.5524614   .6237136    -0.89   0.376    -1.774918    .6699949
        L2D. |  -.2126364   .2429936    -0.88   0.382    -.6888952    .2636224
          ys |
         D1. |   .9905803   .4691946     2.11   0.035     .0709758    1.910185
         LD. |  -1.937912   1.457434    -1.33   0.184    -4.794431    .9186063
        L2D. |   .4870838   .5167524     0.94   0.346    -.5257324      1.4999
     tau1979 |
         D1. |   .0626485   .0640965     0.98   0.328    -.0629784    .1882754
     tau1980 |
         D1. |   .1080019   .1013824     1.07   0.287     -.090704    .3067078
     tau1981 |
         D1. |   .0704241   .1046861     0.67   0.501    -.1347568    .2756051
     tau1982 |
         D1. |    .076515   .1363112     0.56   0.575    -.1906502    .3436801
     tau1983 |
         D1. |   .0895758   .1605703     0.56   0.577    -.2251361    .4042878
     tau1984 |
         D1. |   .0956024   .1662583     0.58   0.565    -.2302579    .4214626
-------------+----------------------------------------------------------------
     sigma_u |  .33809698
     sigma_e |  .25030024
         rho |  .64596356   (fraction of variance due to u_i)
------------------------------------------------------------------------------
Instrumented:    L.n  L2.n  w  L.w  k  L.k  L2.k  ys  L.ys  L2.ys  tau1979  tau1980  tau1981  tau1982
tau1983 tau1984
Instruments:     L2.cumnl L2.n w L.w k L.k L2.k ys L.ys L2.ys tau1979 tau1980 tau1981 tau1982
tau1983
                tau1984

. xtabond2 n l(1/2).n l(0/1).w l(0/2).(k ys) tau1979-tau1984, iv(l2.n l(0/1).w l(0/2).(k ys)
tau1979-tau1984, eq(diff)) iv(l3.n , eq(diff) passthru) robust noleveleq
Favoring speed over space. To switch, type or click on mata: mata set matafavor space.

Arellano-Bond dynamic panel-data estimation, one-step difference GMM results
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =       611
Time variable : year                            Number of groups   =       140
Number of instruments = 16                      Obs per group: min =         4
Wald chi2(16) =    302.91                                      avg =      4.36
Prob > chi2   =     0.000                                      max =         6
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   2.307626   1.054549     2.19   0.029     .2407485    4.374503
         L2. |  -.2240271   .1172406    -1.91   0.056    -.4538144    .0057602
           w |
         --. |  -.8103626    .283096    -2.86   0.004    -1.365221   -.2555047
         L1. |   1.422246   .8507098     1.67   0.095    -.2451148    3.089606
           k |
         --. |   .2530975   .1103882     2.29   0.022     .0367407    .4694543
         L1. |  -.5524613   .3572037    -1.55   0.122    -1.252568     .147645
         L2. |  -.2126364   .1453826    -1.46   0.144    -.4975811    .0723083
          ys |
         --. |   .9905803   .3376145     2.93   0.003     .3288681    1.652293
         L1. |  -1.937912   .9923162    -1.95   0.051    -3.882816     .006992
         L2. |   .4870838   .4247408     1.15   0.251    -.3453929     1.31956
     tau1979 |   .0626485   .0333423     1.88   0.060    -.0027012    .1279982
     tau1980 |   .1080019   .0536941     2.01   0.044     .0027633    .2132405
     tau1981 |   .0704241   .0632658     1.11   0.266    -.0535746    .1944228
     tau1982 |    .076515   .0697354     1.10   0.273    -.0601639    .2131938
     tau1983 |   .0895758   .0821999     1.09   0.276     -.071533    .2506847
     tau1984 |   .0956023   .0780306     1.23   0.221    -.0573348    .2485395
------------------------------------------------------------------------------
Hansen test of overid. restrictions: chi2(0) =   0.00     Prob > chi2 =      .
```

```
Arellano-Bond test for AR(1) in first differences: z =  -1.98  Pr > z =  0.048
Arellano-Bond test for AR(2) in first differences: z =  -0.92  Pr > z =  0.358
------------------------------------------------------------------------------
```

Considerations:

- Estimating in first-differences and instrumentint Δy_{it-1} with y_{it-3} [19] implies that the year 1979 is saved. Thus, the correct specification is to drop the constant (the model is in first differences so that the constant is differenced out) and to include 1979 temporal dummy.
- The command `xtivreg, fd` needs the `facum.ado`, a procedure that we wrote in order to preserve the levels of the instruments despite the first-differencing: the trick is to create the cumulated of the involved variable; it is lagged, eplaining why, in the estimation command, we need to put lag t-2, instead of lag t-3. This procedure also replace initial missing with zero.
- The option `passthru` with `eq(diff)` in `xtabond2` leaves IVs in levels for the first-differences equations. There is also an option `mz` that, if added to the `ivstyle`, convertes missing values in the IVs into zeroes. In this context, however, this option is irrelevant, because results do not change.
- Until now we are estimating a model exactly identified, with IV (2SLS). Hence, the overidentification tests do not work; onestep and twostep coincide; no estimation of the variance-covariance matrix of the error is done, i.e. the MA(1) structure in the error term induced by the first-differencing is not taken into account; the choice of `h(.)`, an option in `xtabond2` that selects the a priori estimate of the variance-covariance matrix of the error terms, is irrelevant.
- The AndersonHsiao estimates are less precise if compared to the GMM ones because they do not use all the relevant information about the structure of the error term, nor all the available moment conditions.

3.6. GMM-dif estimates: Arellano-Bond (1991), Table 4, col. (a1) (a2), p. 290

```
. xtabond2 n l(1/2).n l(0/1).w l(0/2).(k ys) tau1979-tau1984, iv(l(0/1).w l(0/2).(k ys)
tau1979-tau1984, eq(diff)) gmm(n, laglimits(2 .) eq(diff)) noleveleq h(2)
Favoring speed over space. To switch, type or click on mata: mata set matafavor space.

Arellano-Bond dynamic panel-data estimation, one-step difference GMM results
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =       611
Time variable : year                            Number of groups   =       140
Number of instruments = 41                      Obs per group: min =         4
Wald chi2(16) =   1804.32                                      avg =      4.36
Prob > chi2   =     0.000                                      max =         6
------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   .6862261   .1466575     4.68   0.000     .3987827    .9736696
         L2. |  -.0853582   .0438509    -1.95   0.052    -.1713043    .0005879
           w |
         --. |  -.6078208   .0649026    -9.37   0.000    -.7350275   -.4806141
         L1. |   .3926237   .1077977     3.64   0.000     .1813441    .6039032
           k |
         --. |   .3568456   .0365434     9.76   0.000     .2852219    .4284693
         L1. |  -.0580012   .0575366    -1.01   0.313    -.1707708    .0547685
         L2. |  -.0199475   .0410788    -0.49   0.627    -.1004604    .0605654
          ys |
         --. |   .6085073   .1327679     4.58   0.000      .348287    .8687276
         L1. |  -.7111651   .1820286    -3.91   0.000    -1.067935   -.3543955
         L2. |   .1057969    .140974     0.75   0.453     -.170507    .3821008
     tau1979 |   .0095545     .01402     0.68   0.496    -.0179242    .0370332
     tau1980 |   .0220152   .0183075     1.20   0.229    -.0138669    .0578973
     tau1981 |  -.0117743   .0242915    -0.48   0.628    -.0593848    .0358362
```

[19] The IV-set is composed by all the exogenous variables of the system (both included and excluded from the equation of interest): hence, y_{it-2}, y_{it-3}, plus contemporaneous and lagged levels of w, k, ys. Given that y_{it-2} is an exogenous included in the equation, the actual IV for y_{it-1} is y_{it-3}.

```
     tau1982 |  -.0270588   .0256089    -1.06   0.291    -.0772514    .0231338
     tau1983 |  -.0213204   .0279544    -0.76   0.446      -.07611    .0334691
     tau1984 |  -.0077033   .0304006    -0.25   0.800    -.0672875    .0518808
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(26) =  67.59    Prob > chi2 =  0.000

Arellano-Bond test for AR(1) in first differences: z =  -3.99  Pr > z =  0.000
Arellano-Bond test for AR(2) in first differences: z =  -0.55  Pr > z =  0.583
------------------------------------------------------------------------------

.  xtabond2  n  l(1/2).n  l(0/1).w  l(0/2).(k  ys)  tau1979-tau1984,  iv(l(0/1).w  l(0/2).(k  ys)
tau1979-tau1984, eq(diff)) gmm(n, laglimits(2 .) eq(diff)) robust noleveleq h(2)
Favoring speed over space. To switch, type or click on mata: mata set matafavor space.

Arellano-Bond dynamic panel-data estimation, one-step difference GMM results
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =       611
Time variable : year                            Number of groups   =       140
Number of instruments = 41                      Obs per group: min =         4
Wald chi2(16) =   1727.45                                      avg =      4.36
Prob > chi2   =     0.000                                      max =         6
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   .6862261   .1445943     4.75   0.000     .4028266    .9696257
         L2. |  -.0853582   .0560155    -1.52   0.128    -.1951467    .0244302
           w |
         --. |  -.6078208   .1782055    -3.41   0.001    -.9570972   -.2585445
         L1. |   .3926237   .1679931     2.34   0.019     .0633632    .7218842
           k |
         --. |   .3568456   .0590203     6.05   0.000      .241168    .4725233
         L1. |  -.0580012   .0731797    -0.79   0.428    -.2014308    .0854284
         L2. |  -.0199475   .0327126    -0.61   0.542    -.0840631    .0441681
          ys |
         --. |   .6085073   .1725313     3.53   0.000     .2703522    .9466624
         L1. |  -.7111651   .2317163    -3.07   0.002    -1.165321   -.2570095
         L2. |   .1057969   .1412021     0.75   0.454    -.1709542     .382548
     tau1979 |   .0095545   .0102896     0.93   0.353    -.0106127    .0297217
     tau1980 |   .0220152   .0177104     1.24   0.214    -.0126966     .056727
     tau1981 |  -.0117743   .0295079    -0.40   0.690    -.0696086      .04606
     tau1982 |  -.0270588   .0292751    -0.92   0.355    -.0844369    .0303193
     tau1983 |  -.0213204   .0304599    -0.70   0.484    -.0810207    .0383798
     tau1984 |  -.0077033   .0314106    -0.25   0.806     -.069267    .0538604
------------------------------------------------------------------------------
Hansen test of overid. restrictions: chi2(26) =  31.38    Prob > chi2 =  0.214

Arellano-Bond test for AR(1) in first differences: z =  -3.60  Pr > z =  0.000
Arellano-Bond test for AR(2) in first differences: z =  -0.52  Pr > z =  0.606
------------------------------------------------------------------------------

.  xtabond2  n  l(1/2).n  l(0/1).w  l(0/2).(k  ys)  tau1979-tau1984,  iv(l(0/1).w  l(0/2).(k  ys)
tau1979-tau1984, eq(diff)) gmm(n, laglimits(2 .) eq(diff)) robust noleveleq h(2)  twostep
Favoring speed over space. To switch, type or click on mata: mata set matafavor space.

Arellano-Bond dynamic panel-data estimation, two-step difference GMM results
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =       611
Time variable : year                            Number of groups   =       140
Number of instruments = 41                      Obs per group: min =         4
Wald chi2(16) =   1104.72                                      avg =      4.36
Prob > chi2   =     0.000                                      max =         6
------------------------------------------------------------------------------
             |              Corrected
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   .6287089   .1934138     3.25   0.001     .2496248    1.007793
         L2. |  -.0651882   .0450501    -1.45   0.148    -.1534847    .0231084
```

```
           w |
         --. |  -.5257597   .1546107    -3.40   0.001     -.828791   -.2227284
         L1. |   .3112899   .2030006     1.53   0.125     -.086584    .7091638
           k |
         --. |   .2783619   .0728019     3.82   0.000     .1356728    .4210511
         L1. |   .0140994   .0924575     0.15   0.879     -.167114    .1953129
         L2. |  -.0402484   .0432745    -0.93   0.352    -.1250649    .0445681
          ys |
         --. |   .5919243   .1730916     3.42   0.001      .252671    .9311776
         L1. |  -.5659863   .2611008    -2.17   0.030    -1.077734   -.0542381
         L2. |   .1005433   .1610987     0.62   0.533    -.2152043    .4162908
     tau1979 |   .0112155   .0116783     0.96   0.337    -.0116735    .0341045
     tau1980 |   .0230688    .020056     1.15   0.250    -.0162402    .0623778
     tau1981 |  -.0213579   .0332439    -0.64   0.521    -.0865147    .0437989
     tau1982 |  -.0311159   .0339723    -0.92   0.360    -.0977005    .0354686
     tau1983 |  -.0179932   .0369328    -0.49   0.626    -.0903802    .0543938
     tau1984 |  -.0233675   .0366145    -0.64   0.523    -.0951306    .0483956
------------------------------------------------------------------------------
Hansen test of overid. restrictions: chi2(26) =  31.38    Prob > chi2 =  0.214

Arellano-Bond test for AR(1) in first differences: z =  -2.13  Pr > z =  0.034
Arellano-Bond test for AR(2) in first differences: z =  -0.35  Pr > z =  0.725
------------------------------------------------------------------------------

. xtabond2 n l(1/2).n l(0/1).w l(0/2).(k ys) tau1979-tau1984, iv(l(0/1).w l(0/2).(k ys)
tau1979-tau1984, eq(diff)) gmm(n, laglimits(2 .) eq(diff)) noleveleq h(2)  twostep
Favoring speed over space. To switch, type or click on mata: mata set matafavor space.

Arellano-Bond dynamic panel-data estimation, two-step difference GMM results
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =       611
Time variable : year                            Number of groups   =       140
Number of instruments = 41                      Obs per group: min =         4
Wald chi2(16) =   2216.93                                      avg =      4.36
Prob > chi2   =     0.000                                      max =         6
------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   .6287089   .0904543     6.95   0.000     .4514216    .8059961
         L2. |  -.0651882   .0265009    -2.46   0.014     -.117129   -.0132474
           w |
         --. |  -.5257597   .0537692    -9.78   0.000    -.6311453    -.420374
         L1. |   .3112899   .0940116     3.31   0.001     .1270305    .4955492
           k |
         --. |   .2783619   .0449083     6.20   0.000     .1903432    .3663807
         L1. |   .0140994   .0528046     0.27   0.789    -.0893957    .1175946
         L2. |  -.0402484   .0258038    -1.56   0.119    -.0908229     .010326
          ys |
         --. |   .5919243   .1162114     5.09   0.000     .3641542    .8196943
         L1. |  -.5659863   .1396738    -4.05   0.000    -.8397419   -.2922306
         L2. |   .1005433   .1126749     0.89   0.372    -.1202955     .321382
     tau1979 |   .0112155   .0077507     1.45   0.148    -.0039756    .0264066
     tau1980 |   .0230688   .0136626     1.69   0.091    -.0037094    .0498471
     tau1981 |  -.0213579   .0224104    -0.95   0.341    -.0652815    .0225657
     tau1982 |  -.0311159   .0231606    -1.34   0.179    -.0765099     .014278
     tau1983 |  -.0179932   .0232122    -0.78   0.438    -.0634883     .027502
     tau1984 |  -.0233675   .0235452    -0.99   0.321    -.0695153    .0227803
------------------------------------------------------------------------------
Warning: Uncorrected two-step standard errors are unreliable.

Hansen test of overid. restrictions: chi2(26) =  31.38    Prob > chi2 =  0.214

Arellano-Bond test for AR(1) in first differences: z =  -3.00  Pr > z =  0.003
Arellano-Bond test for AR(2) in first differences: z =  -0.42  Pr > z =  0.678
------------------------------------------------------------------------------
```

Considerations.

- All the explanatory variables different from the lagged dependent variable are assumed to be strictly exogenous i.e. $E(\varepsilon_{it} \mid x_{is})=0 \ \forall \ t,s=1,\ldots,T, \ \forall \ i=1,\ldots,N$.
- Column (a) reports robust standard error and homoskedastic Sargan test; hence, it is a mixture of the two first GMM-dif one-step estimates previously reported. In the one-step, the heteroskedastic Hansen test is reported when the option `robust` is specified.
- Column (b) is the fourth GMM-dif estimate, without `robust` option. Note that, in `twostep robust` (third GMM-dif estimate), the Windmeijer (2005) small sample correction for the variance of linear two-step GMM estimators is computed. Warning: do not use it in panel of not negligible size! Note also that, in the two-step, the heteroskedastic Hansen test is always reported, independently from the use or not of the option `robust`.
- The `gmm` option specifies a set of variables to be used as bases for "GMM-style" instrument sets described in Holtz-Eakin,Newey, and Rosen (1988)[20] and Arellano-Bond (1991). By default `xtabond2` uses, for each time period, all available lags of the specified variables in levels <u>dated t-1 or earlier</u> as instruments for the first-difference equations, and the <u>contemporaneous</u> first differences as instruments in the levels equations.[21] The suboption `laglimits(a b)` can override these defaults: for the first-difference equations, lagged levels dated t-a to t-b are used as instruments, while for the levels equations, the first-difference dated t-a+1 is normally used. Note that a and b can each be missing (".") intending infinity; they can even be negative, implying "forward" lags. Hence, there are different ways of writing gmm for `eq(diff)`. For example, the use of y_{it-2} and as an IV for y_{it-1} in equation in first-differences can be written as: `gmm(y, laglimits(2 .))`, or `gmm(l2.y, laglimits (0 .))`, or `gmm(l.y, laglimits (1 .))` or, the default, `gmm(l.y)`.
- In `gmm` the moment conditions are imposed for each time period, variable and lag distance. For example, when we use lag(2 2), we have 5 over-identifying restrictions:

Equations	*IVs*
$y_{i1979} = \beta_1 y_{i1978} + \beta_2 y_{i1977}$	y_{i1977}
$y_{i1980} = \beta_1 y_{i1979} + \beta_2 y_{i1978}$	y_{i1978}
$y_{i1981} = \beta_1 y_{i1980} + \beta_2 y_{i1979}$	y_{i1979}
$y_{i1982} = \beta_1 y_{i1981} + \beta_2 y_{i1980}$	y_{i1980}
$y_{i1983} = \beta_1 y_{i1982} + \beta_2 y_{i1981}$	y_{i1981}
$y_{i1984} = \beta_1 y_{i1983} + \beta_2 y_{i1982}$	y_{i1982}

  For 1 endogenous variable, y_{it-1}, 6 instruments, y_{it-2}, are available (one in each equation, from 1979 to 1984).[22]

  If we use lag(2 3), we have 11 over-identifying restrictions:

Equations	*IVs*
$y_{i1979} = \beta_1 y_{i1978} + \beta_2 y_{i1977}$	$y_{i1976} y_{1977}$
$y_{i1980} = \beta_1 y_{i1979} + \beta_2 y_{i1978}$	$y_{i1977} y_{i1978}$
$y_{i1981} = \beta_1 y_{i1980} + \beta_2 y_{i1979}$	$y_{i1978} y_{i1979}$
$y_{i1982} = \beta_1 y_{i1981} + \beta_2 y_{i1980}$	$y_{i1979} y_{i1980}$
$y_{i1983} = \beta_1 y_{i1982} + \beta_2 y_{i1981}$	$y_{1980} y_{i1981}$
$y_{i1984} = \beta_1 y_{i1983} + \beta_2 y_{i1982}$	$y_{i1981} y_{i1982}$

  For 1 endogenous variable, y_{it-1}, 12 instruments, y_{it-2} and y_{it-3}, are available (two in each equation from 1979 to 1984).[23]

---

[20] Holtz-Eakin, D., W. Newey, and H.S. Rosen (1988) Estimating vector autoregressions with panel data, Econometrica, 56, 1371-95.

[21] These defaults are appropriate for predetermined variables (Bond 2000).

[22] Note that, in the Anderson-Hsiao IV estimation method, where one instrument is imposed to each variable and lag distance, y_{it-2} alone is not able to identify the equation of interest because it is also the included exogenous.

Finally, if we use lag(2 .), we have 26 over-identifying restrictions:

Equations	*IVs*
$y_{i1979} = \beta_1 y_{i1978} + \beta_2 y_{i1977}$	$y_{i1976} y_{1977}$
$y_{i1980} = \beta_1 y_{i1979} + \beta_2 y_{i1978}$	$y_{i1976} y_{i1977} y_{1978}$
$y_{i1981} = \beta_1 y_{i1980} + \beta_2 y_{i1979}$	$y_{i1976} y_{i1977} y_{i1978} y_{1979}$
$y_{i1982} = \beta_1 y_{i1981} + \beta_2 y_{i1980}$	$y_{i1976} y_{i1977} y_{i1978} y_{i1979} y_{1980}$
$y_{i1983} = \beta_1 y_{i1982} + \beta_2 y_{i1981}$	$y_{i1976} y_{i1977} y_{i1978} y_{1979} y_{i1980} y_{1981}$
$y_{i1984} = \beta_1 y_{i1983} + \beta_2 y_{i1982}$	$y_{i1976} y_{i1977} y_{i1978} y_{1979} y_{i1980} y_{i1981} y_{1982}$

For 1 endogenous variable, y_{it-1}, we have 27 moment conditions available, y_{it-2}, y_{it-3}, y_{it-4}, …, y_{i0} (from two in the first equation of 1979, to three in the second equation of 1980, to all the seven availble lags in the last equation of 1984).[24]

- In GMM, by default, missing values are always replaced by zeros.
- When we use `xtabond2` with `ivstyle` only (or with `gmm collapse`) we need only onestep: twostep is irrelevant because no estimation of the covariance matrix of the error is done; hence, also the choice of `h(.)`, the a priori estimate of the covariance matrix of the error terms that serves as the first-step inner product matrix to weight the sample moments whose magnitudes are minimised, is irrelevant. Matrix `h(.)` always has a block diagonal form, with all the T×T blocks the same. When we use `xtabond2 gmm` the choice of `h(.)` is relevant: the default, `h(3)`, will be discussed in section 3.8 on GMM-sys. Option `h(1)` selects the identity matrix used to produce IV/2SLS. Option `h(2)` accounts for the MA(1) structure of the error term in the first-differenced model: the blocks of H are like

```
 2 -1  0 ...
-1  2 -1 ...
 0 -1  2 ...
```

Consider the error term $v_{it}=\mu_i + \varepsilon_{it}$. If the VCOV matrix of the error term, E(εε'), is the identity matrix, i.e. ε is homoskedastic, in the first-differenced equations we have the following VCOV matrix: $E(\Delta v \Delta v') = E(\Delta\varepsilon\Delta\varepsilon') = Var(\varepsilon_{it}-\varepsilon_{it-1}) = Var(\varepsilon_{it}) - 2Cov(\varepsilon_{it}, \varepsilon_{it-1}) + Var(\varepsilon_{it-1}) = 1-2*0+1$. Hence, $Var(\Delta\varepsilon_{it})$, a diagonal element of the error covariance matrix, is 2; similarly, $Cov(\Delta\varepsilon_{it}, \Delta\varepsilon_{it-1})$, on the first subdiagonal of the error covariance matrix, works out to equal -1.

3.7. GMM-dif estimates: Arellano-Bond (1991), Table 4, col. (b) (c), p. 290

```
. xtabond2 n l(1/2).n l(0/1).w k l(0/2).ys tau1979-tau1984, iv(l(0/1).w k l(0/2).ys tau1979-
tau1984, eq(diff)) gmm(n, laglimits(2 .) eq(diff)) noleveleq h(2)  twostep
Favoring speed over space. To switch, type or click on mata: mata set matafavor space.

Arellano-Bond dynamic panel-data estimation, two-step difference GMM results
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =       611
Time variable : year                            Number of groups   =       140
Number of instruments = 39                      Obs per group: min =         4
Wald chi2(14) =   1582.82                                      avg =      4.36
Prob > chi2   =     0.000                                      max =         6
------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
```

[23] Note that, in the Anderson-Hsiao IV estimation method, y_{it-2} and y_{it-3} produce an exactly identified model, given that y_{it-2} is also the exogenous variable included in the equation of interest.

[24] Note that, in the Anderson-Hsiao IV estimation method, y_{it-2}, y_{it-3}, y_{it-4}, y_{it-5}, y_{it-6}, y_{it-7} and y_{it-8} produce 5 over-identifying restrictions; the moment condition linked to y_{it-2} is lost because it is the also the exogenous variable included in the equation of interest. The Anderson-Hsiao IV estimation method that uses all the available lags can be easily obtained by the command `xtabond2 n l(1/2).n l(0/1).w l(0/2).(k ys) tau1979-tau1984, iv(l(0/1).w l(0/2).(k ys) tau1979-tau1984, eq(diff)) gmm(n, laglimits(2 .) eq(diff) collapse) noleveleq h(2) robust`

```
           n |
         L1. |   .4642432   .0862874     5.38   0.000     .2951229    .6333634
         L2. |  -.0566277   .0268605    -2.11   0.035    -.1092734   -.0039821
           w |
         --. |  -.5085975   .0513303    -9.91   0.000     -.609203   -.4079919
         L1. |   .2254592   .0795659     2.83   0.005      .069513    .3814054
           k |   .2899209   .0390319     7.43   0.000     .2134198    .3664221
          ys |
         --. |   .5889862   .1105944     5.33   0.000     .3722253    .8057472
         L1. |   -.445183   .1274288    -3.49   0.000    -.6949389   -.1954271
         L2. |   .0874253    .107306     0.81   0.415    -.1228907    .2977412
     tau1979 |   .0092642   .0072076     1.29   0.199    -.0048624    .0233909
     tau1980 |   .0214197   .0122459     1.75   0.080    -.0025817    .0454212
     tau1981 |  -.0217047   .0206961    -1.05   0.294    -.0622683     .018859
     tau1982 |  -.0402667   .0229537    -1.75   0.079    -.0852552    .0047218
     tau1983 |  -.0357929   .0244132    -1.47   0.143     -.083642    .0120561
     tau1984 |  -.0443511   .0256751    -1.73   0.084    -.0946734    .0059712
------------------------------------------------------------------------------
Warning: Uncorrected two-step standard errors are unreliable.

Hansen test of overid. restrictions: chi2(26) =  29.58    Prob > chi2 =  0.285

Arellano-Bond test for AR(1) in first differences: z =  -2.39  Pr > z =  0.017
Arellano-Bond test for AR(2) in first differences: z =  -0.24  Pr > z =  0.808
------------------------------------------------------------------------------

. xtabond2 n l(1/2).n l(0/1).w k l(0/1).ys tau1979-tau1984, iv(l(0/1).ys tau1979-tau1984,
eq(diff)) gmm(n, lag(2 .) eq(diff)) gmm(w k, lag(2 3) eq(diff)) noleveleq h(2) twostep
Favoring speed over space. To switch, type or click on mata: mata set matafavor space.

Arellano-Bond dynamic panel-data estimation, two-step difference GMM results
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =       611
Time variable : year                            Number of groups   =       140
Number of instruments = 59                      Obs per group: min =         4
Wald chi2(13) =   2668.33                                      avg =      4.36
Prob > chi2   =     0.000                                      max =         6
------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   .8074308   .0491071    16.44   0.000     .7111827    .9036789
         L2. |  -.1134945   .0228508    -4.97   0.000    -.1582813   -.0687076
           w |
         --. |  -.5686242   .0598841    -9.50   0.000    -.6859949   -.4512536
         L1. |    .640707   .0672002     9.53   0.000      .508997     .772417
           k |   .1833987   .0567488     3.23   0.001     .0721731    .2946243
          ys |
         --. |   .8599923   .1097787     7.83   0.000     .6448299    1.075155
         L1. |  -.8632569   .1101426    -7.84   0.000    -1.079132   -.6473814
     tau1979 |   .0162529    .006453     2.52   0.012     .0036053    .0289005
     tau1980 |   .0420327   .0105855     3.97   0.000     .0212854    .0627799
     tau1981 |   .0025669   .0162173     0.16   0.874    -.0292183    .0343522
     tau1982 |  -.0370388   .0186426    -1.99   0.047    -.0735776   -.0004999
     tau1983 |  -.0511972   .0238629    -2.15   0.032    -.0979676   -.0044267
     tau1984 |  -.0587939    .026186    -2.25   0.025    -.1101176   -.0074701
------------------------------------------------------------------------------
Warning: Uncorrected two-step standard errors are unreliable.

Hansen test of overid. restrictions: chi2(49) =  58.71    Prob > chi2 =  0.161

Arellano-Bond test for AR(1) in first differences: z =  -4.07  Pr > z =  0.000
Arellano-Bond test for AR(2) in first differences: z =  -0.68  Pr > z =  0.498
------------------------------------------------------------------------------
```

Considerations.

- Column (b) just omits insignificant dynamics; only the twostep estimates are reported. Column (c) relaxes the assumption that real wage (w) and capital stock (k) are strictly exogenous and, instead, treats them as being endogenous. $E(\varepsilon_{it}|x_{is})=0 \ \forall \ t > s$ and $E(\varepsilon_{it}|x_{is})\neq 0 \ \forall \ t \leq s, \ \forall \ i=1,...,N, \ \forall \ t, s=1,...,T$. Hence, lags of w and k dated (t-2) and earlier can be used as IVs for w_{it}, w_{it-1} and k_{it}. However, estimates of Column (c) can just be approximated because the available data-set lacks company's real sales and real stocks that Arellano-Bond use as additional IVs.

3.8. GMM-dif and GMM-sys estimates: Blundell-Bond (1998), Table 4, col. III and IV, p. 137

```
. xtabond2 n l.n l(0/1).(w k) tau1978-tau1984, iv(tau1978-tau1984, eq(diff)) gmm(n w k, lag(2
.) eq(diff)) noleveleq h(2) robust
Favoring speed over space. To switch, type or click on mata: mata set matafavor space.

Arellano-Bond dynamic panel-data estimation, one-step difference GMM results
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =       751
Time variable : year                            Number of groups   =       140
Number of instruments = 91                      Obs per group: min =         5
Wald chi2(12) =   1163.33                                      avg =      5.36
Prob > chi2   =     0.000                                      max =         7
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   .7074701   .0841788     8.40   0.000     .5424827    .8724576
           w |
         --. |  -.7087965    .117102    -6.05   0.000    -.9383122   -.4792809
         L1. |   .5000149   .1113282     4.49   0.000     .2818157    .7182141
           k |
         --. |   .4659776    .101044     4.61   0.000      .267935    .6640203
         L1. |  -.2151309   .0858525    -2.51   0.012    -.3833987   -.0468631
     tau1978 |   .0057636   .0166077     0.35   0.729    -.0267868     .038314
     tau1979 |   .0136366   .0193748     0.70   0.482    -.0243374    .0516106
     tau1980 |  -.0071557   .0213479    -0.34   0.737    -.0489969    .0346855
     tau1981 |  -.0340692   .0264327    -1.29   0.197    -.0858763    .0177379
     tau1982 |  -.0059175   .0272325    -0.22   0.828    -.0592922    .0474573
     tau1983 |   .0187213   .0288529     0.65   0.516    -.0378294     .075272
     tau1984 |   .0352279   .0331578     1.06   0.288    -.0297603    .1002161
------------------------------------------------------------------------------
Hansen test of overid. restrictions: chi2(79) =  88.80    Prob > chi2 =  0.211

Arellano-Bond test for AR(1) in first differences: z =  -5.60  Pr > z =  0.000
Arellano-Bond test for AR(2) in first differences: z =  -0.14  Pr > z =  0.891
------------------------------------------------------------------------------

. xtabond2 n l.n l(0/1).(w k) tau1978-tau1984, iv(tau1978-tau1984, eq(both)) gmm(n w k, lag(2
.) eq(both)) h(1) robust
Favoring speed over space. To switch, type or click on mata: mata set matafavor space.

Arellano-Bond dynamic panel-data estimation, one-step system GMM results
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =       891
Time variable : year                            Number of groups   =       140
Number of instruments = 113                     Obs per group: min =         6
Wald chi2(12) =   4147.85                                      avg =      6.36
Prob > chi2   =     0.000                                      max =         8
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   .8108394   .0579982    13.98   0.000     .6971649    .9245138
           w |
         --. |  -.7945394   .0971517    -8.18   0.000    -.9849532   -.6041257
         L1. |     .55012    .151645     3.63   0.000     .2529012    .8473388
           k |
         --. |   .4285055   .0763361     5.61   0.000     .2788895    .5781215
         L1. |  -.2802184   .0776689    -3.61   0.000    -.4324466   -.1279903
```

```
    tau1978 |   .0077488   .0200664     0.39   0.699    -.0315806    .0470781
    tau1979 |    .020829   .0236973     0.88   0.379     -.025617    .0672749
    tau1980 |  -.0002589   .0252166    -0.01   0.992    -.0496826    .0491648
    tau1981 |  -.0271456     .02961    -0.92   0.359    -.0851801     .030889
    tau1982 |   .0012306    .026954     0.05   0.964    -.0515983    .0540596
    tau1983 |    .014436   .0254967     0.57   0.571    -.0355367    .0644087
    tau1984 |   .0003278   .0307739     0.01   0.992     -.059988    .0606436
      _cons |   1.006162    .430149     2.34   0.019     .1630853    1.849238
------------------------------------------------------------------------------
Hansen test of overid. restrictions: chi2(106) = 115.73   Prob > chi2 =  0.244

Arellano-Bond test for AR(1) in first differences: z =  -6.49  Pr > z =  0.000
Arellano-Bond test for AR(2) in first differences: z =  -0.08  Pr > z =  0.934
------------------------------------------------------------------------------

. xtabond2 n l.n l(0/1).(w k) tau1978-tau1984, iv(tau1978-tau1984, eq(both)) gmm(n w k, lag(2
.) eq(both)) h(2) robust
Favoring speed over space. To switch, type or click on mata: mata set matafavor space.

Arellano-Bond dynamic panel-data estimation, one-step system GMM results
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =       891
Time variable : year                            Number of groups   =       140
Number of instruments = 113                     Obs per group: min =         6
Wald chi2(12) =   4883.35                                      avg =      6.36
Prob > chi2   =     0.000                                      max =         8
------------------------------------------------------------------------------
             |              Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   .8709811   .0441784    19.72   0.000      .784393    .9575692
           w |
         --. |    -.78236   .1166924    -6.70   0.000    -1.011073   -.5536472
         L1. |    .512276   .1685898     3.04   0.002     .1818462    .8427059
           k |
         --. |    .472313   .0703289     6.72   0.000     .3344709    .6101551
         L1. |  -.3595169   .0716555    -5.02   0.000    -.4999591   -.2190748
     tau1978 |   .0052094   .0207006     0.25   0.801     -.035363    .0457818
     tau1979 |   .0181131   .0242935     0.75   0.456    -.0295012    .0657275
     tau1980 |   .0027193   .0247796     0.11   0.913    -.0458478    .0512864
     tau1981 |  -.0208888   .0296276    -0.71   0.481    -.0789579    .0371803
     tau1982 |   .0155336   .0272278     0.57   0.568    -.0378319    .0688991
     tau1983 |   .0314346   .0253854     1.24   0.216    -.0183198     .081189
     tau1984 |   .0203841   .0314453     0.65   0.517    -.0412474    .0820157
       _cons |    1.00307   .3916881     2.56   0.010     .2353753    1.770765
------------------------------------------------------------------------------
Hansen test of overid. restrictions: chi2(106) = 111.96   Prob > chi2 =  0.327

Arellano-Bond test for AR(1) in first differences: z =  -5.99  Pr > z =  0.000
Arellano-Bond test for AR(2) in first differences: z =  -0.16  Pr > z =  0.872
------------------------------------------------------------------------------

. xtabond2 n l.n l(0/1).(w k) tau1978-tau1984, iv(tau1978-tau1984, eq(both)) gmm(n w k, lag(2
.) eq(both)) h(3) robust
Favoring speed over space. To switch, type or click on mata: mata set matafavor space.

Arellano-Bond dynamic panel-data estimation, one-step system GMM results
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =       891
Time variable : year                            Number of groups   =       140
Number of instruments = 113                     Obs per group: min =         6
Wald chi2(12) =  13267.69                                      avg =      6.36
Prob > chi2   =     0.000                                      max =         8
------------------------------------------------------------------------------
             |              Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           n |
         L1. |   .9326197   .0265224    35.16   0.000     .8806368    .9846025
```

```
           w |
         --. |  -.6305321   .1195813    -5.27   0.000    -.8649071    -.396157
         L1. |   .4597504   .1458203     3.15   0.002     .1739479     .745553
           k |
         --. |   .4820807   .0538588     8.95   0.000     .3765195    .5876419
         L1. |  -.4203043   .0587666    -7.15   0.000    -.5354847   -.3051239
     tau1978 |    .004515   .0194143     0.23   0.816    -.0335364    .0425664
     tau1979 |   .0186472   .0225926     0.83   0.409    -.0256335    .0629279
     tau1980 |   .0038579   .0223114     0.17   0.863    -.0398716    .0475874
     tau1981 |  -.0224858   .0277068    -0.81   0.417    -.0767901    .0318185
     tau1982 |   .0117896    .027819     0.42   0.672    -.0427347    .0663139
     tau1983 |   .0258224    .025958     0.99   0.320    -.0250543    .0766991
     tau1984 |    .020294   .0300096     0.68   0.499    -.0385237    .0791117
       _cons |   .6048596   .2317178     2.61   0.009      .150701    1.059018
------------------------------------------------------------------------------
Hansen test of overid. restrictions: chi2(106) = 109.87   Prob > chi2 =  0.379

Arellano-Bond test for AR(1) in first differences: z =  -5.37  Pr > z =  0.000
Arellano-Bond test for AR(2) in first differences: z =  -0.27  Pr > z =  0.786
------------------------------------------------------------------------------
```

Considerations.

- Ther model is treated as a system of equations, one for each time period, where now: the predetermined and endogenous variables in first-differences are instrumented with suitable lags of their own levels; the predetermined and endogenous variables in levels are instrumented with suitable lags of their own first differences; strictly exogenous regressors, as well as any other instruments, enter the instrument matrix in the conventional instrumental variables fashion: with one column per instrument.
- The default of `xtabond2` in estimating GMM-sys is `h(3)`. First-difference equations are treated as being for periods 1 to T and levels equations as being for periods T+1 to 2T. Observations for the levels equations hold the original values of the variables, while those for the first-differenced equations are of course first-differenced. The blocks of H are then 2Tx2T, and are a priori estimates of the VCOV matrix of the vector (Δv' v')'; however, in the system GMM case, there is not as natural a choice for it. As a result, several variants have been used. In general, still assuming that E(εε') is the identity matrix, i.e. ε is homoskedastic, the VCOV matrix of the vector (Δv' v')' will have blocks like:

```
 2 -1  0  ...  1    0    0  ...
-1  2 -1 ...  -1    1    0  ...
 0 -1  2 ...   0   -1    1  ...
 .  .  . ...   .    .    .  ...
 1 -1  0 ...  1+s1 0    0  ...
 0  1 -1 ...   0  1+s2 0  ...
 0  0  1 ...   0    0  1+s3...
 .  .  . ...   .    .    .  ...
```

where si=Var(μ_i). For example, a diagonal element in the bottom right quadrant is: $Var(\mu_i+\varepsilon_{it}) = Var(\mu_i) + 2Cov(\mu_i, \varepsilon_{it}) + Var(\varepsilon_{it}) = Var(\mu_i) + 2*0+1 = 1+si$. Similarly, on the diagonals of the bottom left and upper right quadrants, we have: $Cov(v_{it}, \Delta v_{it}) = Cov(v_{it}, \Delta\varepsilon_{it}) = Cov(\mu_i, \Delta\varepsilon_{it}) + Var(\varepsilon_{it}) - Cov(\varepsilon_{it}, \varepsilon_{it-1}) = 0+1-0=1$.
- Blundell-Bond use, in their 1998 paper, `h(1)`. The matrix `h(2)` is the one used in the 2002 version of DPD for Ox. Note how results change according to the choice of the a priori estimate of the covariance matrix of the error terms.