2022 Essex Summer School
3K: Dynamics and Heterogeneity

Robert W. Walker, Ph. D.
Associate Professor of Quantitative Methods
Atkinson Graduate School of Management
Willamette University
Salem, Oregon USA
rwalker@willamette.edu

August 18, 2022

Outline for Day 9

- A Brief Bit on IV
- DPD

Instrumental Variables

Three key conditions for instruments in general:

Orthogonality with y

Relevance to endogenous x

Variance components of instruments are equivalent

Literature on weak instruments is relevant.

There are two classes of instrumental variables estimators in Stata.

- Hausman-Taylor: Subset of RHS variables are correlated with random effects. The idea is that we can use time-varying covariates to achieve identification for time-invariant things correlated with random effects. Not all that useful in applied setting because convincing instruments are hard to find.

- General IV: Endogenous covariates
  `xtivreg` implements this for the usual models (FE, RE, BE, FD)

xtfevd

Plümper and Troeger have designed a procedure to solve one of the principle problems that arises in fixed effects regressions: it is either impossible or suboptimal to estimate the effects of time-invariant or nearly time-invariant regressors. Their approach plays off of the generic consistency of the fixed effects estimator. In general, they begin by estimating an LSDV model.

$$y_{it} = \alpha_i + X_{it}\beta + \epsilon_{it}$$

They then proceed to model the unit effects as a function of (largely) time-invariant regressors that they denote as Z

$$\alpha_i = Z_i\gamma + \psi_i$$

In a third stage, they then construct the regression with an offset. In effect, they take the offset and add it to the regression such as,

$$y_{it} = \psi_i + X_{it}\beta + Z_i\gamma + \nu_{it}$$

and adjust the variance/covariance matrix of the errors accordingly.

Dynamic Panel Data

We have encountered Hurwicz/Nickell bias. Dynamic panel data estimators are an effort to avoid this problem.

GMM

Generalized method of moments estimators are a class of estimators created by analogs of the population moment conditions for sample moments. For example, linear regression is a GMM estimator and the moment restriction that must hold for OLS is that $\mathbb{E}[\mathbf{X}'\epsilon] = 0$. With endogenous $\mathbf{x} \in \mathbf{X}$, we instrument using $\mathbf{z}$. If there is one $\mathbf{z}$ for each endogenous x, we have a standard IV. Without exact identification, we need iteration and GMM estimators will typically involve testing these overidentifying restrictions using a Sargan test, as we will see.

GMM for Panels

The trick here is that the panel structure gives us numerous instruments for "free". Comes in two forms. Single-equation and systems estimators. With systems estimators, assumptions give us leverage on moment conditions in both level and difference forms, we use these jointly to estimate the parameters of interest.

Introducing DPD

- We are interested in estimating the parameters of models of the form

$$y_{it} = y_{i,t-1}\gamma + X_{it}\beta + \alpha_i + \epsilon_{it}$$

  for $i = \{1, \ldots, N\}$ and $t = \{1, \ldots, T\}$ using datasets with large N and fixed T

- By construction, $y_{i,t-1}$ is correlated with the unobserved individual-level effect α_i.

- Removing α_i by the within transform produces an inconsistent estimator with T fixed.

- First difference both sides and look for instrumental-variables (IV) and generalized method-of-moments (GMM) estimators

Arrelano-Bond

- First differencing the model equation yields

$$\Delta y_{it} = \Delta y_{i,t-1}\gamma + \Delta x_{it}\beta + \Delta\epsilon_{it}$$

- The α_i are gone, but the $y_{i,t-1}$ in $\Delta y_{i,t-1}$ is a function of the $\epsilon_{i,t-1}$ which is also in $\Delta\epsilon_{it}$.

- $\Delta y_{i,t-1}$ is correlated with $\Delta\epsilon_{it}$ by construction

- Anderson and Hsiao (1981) give a 2SLS estimator based on (further) lags of Δy_{it} as instruments for $\Delta y_{i,t-1}$. E.g. if ϵ_{it} is IID over i and t, $\Delta y_{i,t-2}$ is valid for $\Delta y_{i,t-1}$.

- Anderson and Hsiao (1981) also suggest a 2SLS estimator based on lagged levels of y_{it} as instruments for $\Delta y_{i,t-1}$. E.g. if ϵ_{it} is IID over i and t, $y_{i,t-2}$ can instrument for $\Delta y_{i,t-1}$.

- Holtz-Eakin, and co-authors (1988) and Arellano and Bond (1991) showed how to construct estimators based on moment equations constructed from further lagged levels of y_{it} and the first-differenced errors.

- We are creating moment conditions using lagged levels of the dependent variable with first differences, $\Delta\epsilon_{it}$. First-differences of strictly exogenous covariates also create moment conditions.

- Assume that ϵ_{it} are IID over i and t (no serial correlation)

- GMM is needed because there are more instruments than parameters.

Strict Exogeneity vs. Predetermined

- If regressors are strictly exogenous: $\mathbb{E}[x_{it}\epsilon_{is}] = 0 \; \forall s, t$.

- If predetermined, $\mathbb{E}[x_{it}\epsilon_{is}] \neq 0 \text{ if} s < t$ but $\mathbb{E}[x_{it}\epsilon_{is}] = 0 \; \forall s \geq t$

- Dynamic panel data models allow predetermined regressors. [backward feedback, no forward feedback]

A bit more on this and GMM

- The moment conditions formed by assuming that particular lagged levels of the dependent variable are orthogonal to the differenced disturbances are known as GMM-type moment conditions

- Sometimes they are called sequential moment conditions

- The moment conditions formed using the strictly exogenous covariates are just standard IV moment conditions, so they are called standard moment conditions

- The dynamic panel-data estimators in Stata report which transforms of which variables were used as instruments

- In GMM estimators, we weight the vector of sample-average moment conditions by the inverse of a positive definite matrix

- When that matrix is the covariance matrix of the moment conditions, we have an efficient GMM estimator
- In the case of nonidentically distributed disturbances, we can use a two-step GMM estimator that estimates the covariance matrix of the moment conditions using the first-step residuals
- Although the large-sample robust variance-covariance matrix of the two-step estimator does not depend on the fact that estimated residuals were used, simulation studies have found that that Windmejier's bias-corrected estimator performs much better
- Specifying vce(robust) produces an estimated VCE that is robust to heteroskedasticity
- There is a result in the large-sample theory for GMM which states that the

VCE of the two-step estimator does not depend on the fact that it uses the residuals from the first step. Windmeijer 2005 bias-corrects the VCE of the two-step GMM.

- No robust Sargan test but Arrelano-Bond test exists.

- When the variables are predetermined, it means that we cannot include the whole vector of differences of observed xit into the instrument matrix

- We just include the levels of x_{it} for those time periods that are assumed to be unrelated to $\Delta\epsilon_{it}$.

- The Arellano-Bond estimator formed moment conditions using lagged-levels of the dependent variable and the predetermined variables with first-differences of the disturbances

- Arellano and Bover(1995) and Blundell and Bond (1998) found that if the autoregressive process is too persistent, then the lagged-levels are weak instruments

- These authors proposed using additional moment conditions in which lagged differences of the dependent variable are orthogonal to levels of the disturbances

- To get these additional moment conditions, they assumed that panel-level effect is unrelated to the first observable first-difference of the dependent variable

- `xtdpdsys` is syntactically similar to `xtabond`

The Data for Implementation

```
Contains data from abdata.dta
  obs:        1,031                          Layard & Nickell, Unemployment
                                               in Britain, Economica 53, 1986
 -------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
ind             int    %8.0g                  industry
year            int    %8.0g
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)
yr1980          float  %9.0g
yr1981          float  %9.0g
yr1982          float  %9.0g
yr1983          float  %9.0g
yr1984          float  %9.0g
id              float  %9.0g                  firm ID
-------------------------------------------------------------------------------
Sorted by:  id  year
```

Implementation

- `xtregar`: , `re` and `fe` options

  Fit a first order autoregressive structure to TSCS data.
  Defaults to an iterative estimator but `twostep` is available.
  `lbi` gives a test of the hypothesis that ρ is zero. (not a default)

- `xtabond`

  `estat abond` gives a test for autocorrelation
  `estat sargan` gives the overidentifying restrictions test

- `xtlsdvc y x, initial(ah or ab or bb) vcov(1000 bs iter)` will handle unbalanced
  Bias-corrected least-squares dummy variable (LSDV) estimators for the standard autoregressive panel-data model using the bias approximations in Bruno (2005a) for unbalanced panels

- `xtivreg`

- `xtdpd` fits Arellano-Bond and Arellano-Bover/Blundell-Bond

  `estat abond` gives a test for autocorrelation
  `estat sargan` gives the overidentifying restrictions test (Rejection implies failure of assumptions)

More on DPD

- David Roodman's excellent and well documented `xtabond2` extends the Stata command and incorporates orthogonal deviations transformation that assist in gapped panels. I personally think it is the best software for this.

- Systems DPD is complicated but perhaps very useful.

- As an aside, I laughed pretty hard at a post on econ job rumours where someone claimed that no one actually understands these models! [Not true, I am positive that Hansen does.....]

	firm	year	sector	emp	wage	capital
Grand mean	73.20	1979.65	5.12	7.89	23.92	2.51
S.D.	41.23	2.22	2.68	15.93	5.65	6.25
TSS	1751193.23	5058.30	7387.36	261539.39	32861.76	40217.79
Between S.D.	40.56	0.60	2.68	16.17	5.18	6.10
BSS	1751193.23	368.30	7387.36	256508.78	28458.33	39065.07
Within S.D.	0.00	2.13	0.00	2.21	2.07	1.06
WSS	0.00	4690.00	0.00	5030.61	4403.43	1152.72
% Within	0.00	0.93	0.00	0.02	0.13	0.03

```
> # To make it match the Stata data.
> EmplUK$n <- log(EmplUK$emp)
> EmplUK$w <- log(EmplUK$wage)
> EmplUK$k <- log(EmplUK$capital)
> EmplUK$ys <- log(EmplUK$output)
```

```
> # Can just use log syntax to solve it.
> # Arellano and Bond (1991), table 4(a1)
> Table4.a1 <- pgmm(log(emp) ~ lag(log(emp), 1:2) + lag(log(wage),
> summary(Table4.a1)

Twoways effects One step model

Call:
pgmm(formula = log(emp) ~ lag(log(emp), 1:2) + lag(log(wage),
    0:1) + lag(log(capital), 0:2) + lag(log(output), 0:2) | lag(log(
    2:99), data = EmplUK, effect = "twoways", model = "onestep")

Unbalanced Panel: n=140, T=7-9, N=1031

Number of Observations Used:  611
```

```
Residuals
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-0.6007000 -0.0299500  0.0000000 -0.0001193  0.0311500  0.5693000

Coefficients
                         Estimate Std. Error z-value  Pr(>|z|)
lag(log(emp), 1:2)1      0.686226   0.144594  4.7459 2.076e-06 ***
lag(log(emp), 1:2)2     -0.085358   0.056016 -1.5238 0.1275510
lag(log(wage), 0:1)0    -0.607821   0.178205 -3.4108 0.0006478 ***
lag(log(wage), 0:1)1     0.392623   0.167993  2.3371 0.0194319 *
lag(log(capital), 0:2)0  0.356846   0.059020  6.0462 1.483e-09 ***
lag(log(capital), 0:2)1 -0.058001   0.073180 -0.7926 0.4280206
lag(log(capital), 0:2)2 -0.019948   0.032713 -0.6098 0.5420065
lag(log(output), 0:2)0   0.608506   0.172531  3.5269 0.0004204 ***
lag(log(output), 0:2)1  -0.711164   0.231716 -3.0691 0.0021469 **
lag(log(output), 0:2)2   0.105798   0.141202  0.7493 0.4536974
```

```
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Sargan Test: chisq(25) = 48.74983 (p.value=0.0030295)
Autocorrelation test (1): normal = -3.599593 (p.value=0.00031872)
Autocorrelation test (2): normal = -0.5160282 (p.value=0.60583)
Wald test for coefficients: chisq(10) = 408.2859 (p.value=< 2.22e-16
Wald test for time dummies: chisq(6) = 11.57904 (p.value=0.072046)
```

```
> ## Arellano and Bond (1991), table 4b
> Table4.b <- pgmm(log(emp) ~ lag(log(emp), 1:2) + lag(log(wage), (
+            + log(capital) + lag(log(output), 0:1) | lag(log(emp),
+             data = EmplUK, effect = "twoways", model = "twosteps"
> # To make it match Stata
> summary(Table4.b, robust=FALSE)
```

```
Twoways effects Two steps model

Call:
pgmm(formula = log(emp) ~ lag(log(emp), 1:2) + lag(log(wage),
    0:1) + log(capital) + lag(log(output), 0:1) | lag(log(emp),
    2:99), data = EmplUK, effect = "twoways", model = "twosteps")

Unbalanced Panel: n=140, T=7-9, N=1031

Number of Observations Used:  611

Residuals
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-0.6191000 -0.0255700  0.0000000 -0.0001339  0.0332000  0.6410000

Coefficients
```

```
                         Estimate Std. Error  z-value  Pr(>|z|)
lag(log(emp), 1:2)1      0.474151   0.085303   5.5584 2.722e-08 ***
lag(log(emp), 1:2)2     -0.052967   0.027284  -1.9413 0.0522200 .
lag(log(wage), 0:1)0    -0.513205   0.049345 -10.4003 < 2.2e-16 ***
lag(log(wage), 0:1)1     0.224640   0.080063   2.8058 0.0050192 **
log(capital)             0.292723   0.039463   7.4177 1.191e-13 ***
lag(log(output), 0:1)0  0.609775   0.108524   5.6188 1.923e-08 ***
lag(log(output), 0:1)1 -0.446373   0.124815  -3.5763 0.0003485 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Sargan Test: chisq(25) = 30.11247 (p.value=0.22011)
Autocorrelation test (1): normal = -2.427829 (p.value=0.01519)
Autocorrelation test (2): normal = -0.3325401 (p.value=0.73948)
Wald test for coefficients: chisq(7) = 371.9877 (p.value=< 2.22e-16)
Wald test for time dummies: chisq(6) = 26.9045 (p.value=0.0001509)
```

```
> # Or with Robust [Notice it is default]
> summary(Table4.b)

Twoways effects Two steps model

Call:
pgmm(formula = log(emp) ~ lag(log(emp), 1:2) + lag(log(wage),
    0:1) + log(capital) + lag(log(output), 0:1) | lag(log(emp),
    2:99), data = EmplUK, effect = "twoways", model = "twosteps")

Unbalanced Panel: n=140, T=7-9, N=1031

Number of Observations Used:  611

Residuals
      Min.    1st Qu.      Median        Mean    3rd Qu.        Max.
```

```
-0.6191000 -0.0255700  0.0000000 -0.0001339  0.0332000  0.6410000

Coefficients
                        Estimate Std. Error z-value  Pr(>|z|)
lag(log(emp), 1:2)1     0.474151   0.185398  2.5575 0.0105437 *
lag(log(emp), 1:2)2    -0.052967   0.051749 -1.0235 0.3060506
lag(log(wage), 0:1)0   -0.513205   0.145565 -3.5256 0.0004225 ***
lag(log(wage), 0:1)1    0.224640   0.141950  1.5825 0.1135279
log(capital)            0.292723   0.062627  4.6741 2.953e-06 ***
lag(log(output), 0:1)0  0.609775   0.156263  3.9022 9.530e-05 ***
lag(log(output), 0:1)1 -0.446373   0.217302 -2.0542 0.0399605 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Sargan Test: chisq(25) = 30.11247 (p.value=0.22011)
Autocorrelation test (1): normal = -1.53845 (p.value=0.12394)
```

```
Autocorrelation test (2): normal = -0.2796829 (p.value=0.77972)
Wald test for coefficients: chisq(7) = 142.0353 (p.value=< 2.22e-16)
Wald test for time dummies: chisq(6) = 16.97046 (p.value=0.0093924)

> ## Blundell and Bond (1998) table 4
> Table4.BB <- pgmm(log(emp) ~ lag(log(emp), 1)+ lag(log(wage), 0:1
+             lag(log(capital), 0:1) | lag(log(emp), 2:99) +
+             lag(log(wage), 2:99) + lag(log(capital), 2:99),
+             data = EmplUK, effect = "twoways", model = "onestep",
+             transformation = "ld")
> summary(Table4.BB, robust = TRUE)

Twoways effects One step model

Call:
pgmm(formula = log(emp) ~ lag(log(emp), 1) + lag(log(wage), 0:1) +
```

```
    lag(log(capital), 0:1) | lag(log(emp), 2:99) + lag(log(wage),
    2:99) + lag(log(capital), 2:99), data = EmplUK, effect = "twoway
    model = "onestep", transformation = "ld")

Unbalanced Panel: n=140, T=7-9, N=1031

Number of Observations Used:  1642

Residuals
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-0.7530000 -0.0369000  0.0000000  0.0002882  0.0466100  0.6002000

Coefficients
                         Estimate Std. Error z-value  Pr(>|z|)
lag(log(emp), 1)         0.935605   0.026295 35.5810 < 2.2e-16 ***
lag(log(wage), 0:1)0    -0.630976   0.118054 -5.3448 9.050e-08 ***
```

```
lag(log(wage), 0:1)1      0.482620   0.136887  3.5257 0.0004224 ***
lag(log(capital), 0:1)0  0.483930   0.053867  8.9838 < 2.2e-16 ***
lag(log(capital), 0:1)1 -0.424393   0.058479 -7.2572 3.952e-13 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Sargan Test: chisq(100) = 118.763 (p.value=0.097096)
Autocorrelation test (1): normal = -4.808434 (p.value=1.5212e-06)
Autocorrelation test (2): normal = -0.2800133 (p.value=0.77947)
Wald test for coefficients: chisq(5) = 11174.82 (p.value=< 2.22e-16)
Wald test for time dummies: chisq(7) = 14.71138 (p.value=0.039882)
```

References

The manual for R package `plm` was published in the Journal of Statistical Software. It is nice and extensive excepting the application of dpd models. Kit Baum has a very nice discussion of this in Stata in a set of course slides on the web at Boston College [search google for Baum Dynamic Panel Data Estimators].