

Dynamic Panel Data models: estimation and forecasting

Christopher F Baum

Boston College and DIW Berlin

An empirical exercise

To illustrate the performance of the several estimators, we make use of the original AB dataset, available within Stata with `webuse abdata`. This is an unbalanced panel of annual data from 140 UK firms for 1976–1984. In their original paper, they modeled firms' employment n using a partial adjustment model to reflect the costs of hiring and firing, with two lags of employment.

Other variables included were the current and lagged wage level w , the current, once- and twice-lagged capital stock (k) and the current, once- and twice-lagged output in the firm's sector (y_s). All variables are expressed as logarithms. A set of time dummies is also included to capture business cycle effects.

An empirical exercise

To illustrate the performance of the several estimators, we make use of the original AB dataset, available within Stata with `webuse abdata`. This is an unbalanced panel of annual data from 140 UK firms for 1976–1984. In their original paper, they modeled firms' employment n using a partial adjustment model to reflect the costs of hiring and firing, with two lags of employment.

Other variables included were the current and lagged wage level w , the current, once- and twice-lagged capital stock (k) and the current, once- and twice-lagged output in the firm's sector (y_s). All variables are expressed as logarithms. A set of time dummies is also included to capture business cycle effects.

If we were to estimate this model ignoring its dynamic panel nature, we could merely apply `regress` with panel-clustered standard errors:

Try it out:

```
regress n nL1 nL2 w wL1 k kL1 kL2 ys ysL1 ysL2 yr*, cluster(id)
```

One obvious difficulty with this approach is the likely importance of firm-level unobserved heterogeneity. We have accounted for potential correlation between firms' errors over time with the cluster-robust VCE, but this does not address the potential impact of unobserved heterogeneity on the conditional mean.

We can apply the within transformation to take account of this aspect of the data: **Try it out:**

```
xtreg n nL1 nL2 w wL1 k kL1 kL2 ys ysL1 ysL2 yr*, fe cluster(id)
```

If we were to estimate this model ignoring its dynamic panel nature, we could merely apply `regress` with panel-clustered standard errors:

Try it out:

```
regress n nL1 nL2 w wL1 k kL1 kL2 ys ysL1 ysL2 yr*, cluster(id)
```

One obvious difficulty with this approach is the likely importance of firm-level unobserved heterogeneity. We have accounted for potential correlation between firms' errors over time with the cluster-robust VCE, but this does not address the potential impact of unobserved heterogeneity on the conditional mean.

We can apply the within transformation to take account of this aspect of the data: **Try it out:**

```
xtreg n nL1 nL2 w wL1 k kL1 kL2 ys ysL1 ysL2 yr*, fe cluster(id)
```

If we were to estimate this model ignoring its dynamic panel nature, we could merely apply `regress` with panel-clustered standard errors:

Try it out:

```
regress n nL1 nL2 w wL1 k kL1 kL2 ys ysL1 ysL2 yr*, cluster(id)
```

One obvious difficulty with this approach is the likely importance of firm-level unobserved heterogeneity. We have accounted for potential correlation between firms' errors over time with the cluster-robust VCE, but this does not address the potential impact of unobserved heterogeneity on the conditional mean.

We can apply the within transformation to take account of this aspect of the data: **Try it out:**

```
xtreg n nL1 nL2 w wL1 k kL1 kL2 ys ysL1 ysL2 yr*, fe cluster(id)
```

The fixed effects estimates will suffer from Nickell bias, which may be severe given the short timeseries available.

	OLS		FE	
nL1	1.045***	(20.17)	0.733***	(12.28)
nL2	-0.0765	(-1.57)	-0.139	(-1.78)
w	-0.524**	(-3.01)	-0.560***	(-3.51)
k	0.343***	(7.06)	0.388***	(6.82)
ys	0.433*	(2.42)	0.469**	(2.74)
<i>N</i>	751		751	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In the original OLS regression, the lagged dependent variable was positively correlated with the error, biasing its coefficient upward. In the fixed effects regression, its coefficient is biased downward due to the negative sign on ν_{t-1} in the transformed error. The OLS estimate of the first lag of n is 1.045; the fixed effects estimate is 0.733.

Given the opposite directions of bias present in these estimates, consistent estimates should lie between these values, which may be a useful check. As the coefficient on the second lag of n cannot be distinguished from zero, the first lag coefficient should be below unity for dynamic stability.

In the original OLS regression, the lagged dependent variable was positively correlated with the error, biasing its coefficient upward. In the fixed effects regression, its coefficient is biased downward due to the negative sign on ν_{t-1} in the transformed error. The OLS estimate of the first lag of n is 1.045; the fixed effects estimate is 0.733.

Given the opposite directions of bias present in these estimates, consistent estimates should lie between these values, which may be a useful check. As the coefficient on the second lag of n cannot be distinguished from zero, the first lag coefficient should be below unity for dynamic stability.

To deal with these two aspects of the estimation problem, we might use the Anderson–Hsiao estimator to the first-differenced equation, instrumenting the lagged dependent variable with the twice-lagged level: **Try it out:**

```
ivregress 2sls D.n (D.nL1 = nL2) D.(nL2 w wL1 k kL1 kL2 ///  
ys ysL1 ysL2 yr1979 yr1980 yr1981 yr1982 yr1983 )
```

	A-H	
D.nL1	2.308	(1.17)
D.nL2	-0.224	(-1.25)
D.w	-0.810**	(-3.10)
D.k	0.253	(1.75)
D.y _s	0.991*	(2.14)
<i>N</i>	611	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Although these results should be consistent, they are quite disappointing. The coefficient on lagged n is outside the bounds of its OLS and FE counterparts, and much larger than unity, a value consistent with dynamic stability. It is also very imprecisely estimated.

	A-H	
D.nL1	2.308	(1.17)
D.nL2	-0.224	(-1.25)
D.w	-0.810**	(-3.10)
D.k	0.253	(1.75)
D.y _s	0.991*	(2.14)
<i>N</i>	611	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Although these results should be consistent, they are quite disappointing. The coefficient on lagged n is outside the bounds of its OLS and FE counterparts, and much larger than unity, a value consistent with dynamic stability. It is also very imprecisely estimated.

The difference GMM approach deals with this inherent endogeneity by transforming the data to remove the fixed effects. The standard approach applies the first difference (FD) transformation, which as discussed earlier removes the fixed effect at the cost of introducing a correlation between $\Delta y_{i,t-1}$ and $\Delta \nu_{it}$, both of which have a term dated $(t - 1)$. This is preferable to the application of the within transformation, as that transformation makes every observation in the transformed data endogenous to every other for a given individual.

The one disadvantage of the first difference transformation is that it magnifies gaps in unbalanced panels. If some value of y_{it} is missing, then both Δy_{it} and $\Delta y_{i,t-1}$ will be missing in the transformed data. This motivates an alternative transformation: the forward orthogonal deviations (FOD) transformation, proposed by Arellano and Bover (*J. Econometrics*, 1995).

The difference GMM approach deals with this inherent endogeneity by transforming the data to remove the fixed effects. The standard approach applies the first difference (FD) transformation, which as discussed earlier removes the fixed effect at the cost of introducing a correlation between $\Delta y_{i,t-1}$ and $\Delta \nu_{it}$, both of which have a term dated $(t - 1)$. This is preferable to the application of the within transformation, as that transformation makes every observation in the transformed data endogenous to every other for a given individual.

The one disadvantage of the first difference transformation is that it magnifies gaps in unbalanced panels. If some value of y_{it} is missing, then both Δy_{it} and $\Delta y_{i,t-1}$ will be missing in the transformed data. This motivates an alternative transformation: the forward orthogonal deviations (FOD) transformation, proposed by Arellano and Bover (*J. Econometrics*, 1995).

In contrast to the within transformation, which subtracts the average of all observations' values from the current value, and the FD transformation, that subtracts the previous value from the current value, the FOD transformation subtracts the average of all available *future* observations from the current value. While the FD transformation drops the first observation on each individual in the panel, the FOD transformation drops the last observation for each individual. It is computable for all periods except the last period, even in the presence of gaps in the panel.

The FOD transformation is not available in any of official Stata's DPD commands, but it is available in David Roodman's `xtabond2` implementation of the DPD estimator, available from SSC.

In contrast to the within transformation, which subtracts the average of all observations' values from the current value, and the FD transformation, that subtracts the previous value from the current value, the FOD transformation subtracts the average of all available *future* observations from the current value. While the FD transformation drops the first observation on each individual in the panel, the FOD transformation drops the last observation for each individual. It is computable for all periods except the last period, even in the presence of gaps in the panel.

The FOD transformation is not available in any of official Stata's DPD commands, but it is available in David Roodman's `xtabond2` implementation of the DPD estimator, available from SSC.

To illustrate the use of the AB estimator, we may reestimate the model with `xtabond2`, assuming that the only endogeneity present is that involving the lagged dependent variable. **Try it out:**

```
xtabond2 n L(1/2).n L(0/1).w L(0/2).(k ys) yr*, gmm(L.n) ///
iv(L(0/1).w L(0/2).(k ys) yr*) nolevel robust small
```

Note that in `xtabond2` syntax, every right-hand variable generally appears twice in the command, as instruments must be explicitly specified when they are instrumenting themselves. In this example, all explanatory variables except the lagged dependent variable are taken as “IV-style” instruments, entering the Z matrix as a single column. The lagged dependent variable is specified as a “GMM-style” instrument, where all available lags will be used as separate instruments. The `noleveleq` option is needed to specify the AB estimator.

To illustrate the use of the AB estimator, we may reestimate the model with `xtabond2`, assuming that the only endogeneity present is that involving the lagged dependent variable. **Try it out:**

```
xtabond2 n L(1/2).n L(0/1).w L(0/2).(k ys) yr*, gmm(L.n) ///
iv(L(0/1).w L(0/2).(k ys) yr*) nolevel robust small
```

Note that in `xtabond2` syntax, every right-hand variable generally appears twice in the command, as instruments must be explicitly specified when they are instrumenting themselves. In this example, all explanatory variables except the lagged dependent variable are taken as “IV-style” instruments, entering the \mathbf{Z} matrix as a single column. The lagged dependent variable is specified as a “GMM-style” instrument, where all available lags will be used as separate instruments. The `noleveleq` option is needed to specify the AB estimator.

	A-B	
L.n	0.686***	(4.67)
L2.n	-0.0854	(-1.50)
w	-0.608**	(-3.36)
k	0.357***	(5.95)
ys	0.609***	(3.47)
<i>N</i>	611	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In these results, 41 instruments have been created, with 17 corresponding to the “IV-style” regressors and the rest computed from lagged values of n . Note that the coefficient on the lagged dependent variable now lies within the range for dynamic stability. In contrast to that produced by the Anderson–Hsiao estimator, the coefficient is quite precisely estimated.

A-B		
L.n	0.686***	(4.67)
L2.n	-0.0854	(-1.50)
w	-0.608**	(-3.36)
k	0.357***	(5.95)
ys	0.609***	(3.47)
<i>N</i>	611	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In these results, 41 instruments have been created, with 17 corresponding to the “IV-style” regressors and the rest computed from lagged values of n . Note that the coefficient on the lagged dependent variable now lies within the range for dynamic stability. In contrast to that produced by the Anderson–Hsiao estimator, the coefficient is quite precisely estimated.

There are 25 overidentifying restrictions in this instance, as shown in the first column below. The `hansen_df` represents the degrees of freedom for the Hansen J test of overidentifying restrictions. The p -value of that test is shown as `hansenp`.

	All lags		lags 2-5		lags 2-4	
L.n	0.686***	(4.67)	0.835*	(2.59)	1.107***	(3.94)
L2.n	-0.0854	(-1.50)	0.262	(1.56)	0.231	(1.32)
w	-0.608**	(-3.36)	-0.671**	(-3.18)	-0.709**	(-3.26)
k	0.357***	(5.95)	0.325***	(4.95)	0.309***	(4.55)
ys	0.609***	(3.47)	0.640**	(3.07)	0.698***	(3.45)
hansen_df	25		16		13	
hansenp	0.177		0.676		0.714	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In this table, we can examine the sensitivity of the results to the choice of “GMM-style” lag specification. In the first column, all available lags of the level of n are used. In the second column, the `lag(2 5)` option is used to restrict the maximum lag to 5 periods, while in the third column, the maximum lag is set to 4 periods. Fewer instruments are used in those instances, as shown by the smaller values of `sar_df`.

The p -value of Hansen’s J is also considerably larger for the restricted-lag cases. On the other hand, the estimate of the lagged dependent variable’s coefficient appears to be quite sensitive to the choice of lag length.

In this table, we can examine the sensitivity of the results to the choice of “GMM-style” lag specification. In the first column, all available lags of the level of n are used. In the second column, the `lag(2 5)` option is used to restrict the maximum lag to 5 periods, while in the third column, the maximum lag is set to 4 periods. Fewer instruments are used in those instances, as shown by the smaller values of `sar_df`.

The p -value of Hansen’s J is also considerably larger for the restricted-lag cases. On the other hand, the estimate of the lagged dependent variable’s coefficient appears to be quite sensitive to the choice of lag length.

We illustrate estimating this equation with both the FD transformation and the forward orthogonal deviations (FOD) transformation:

	First diff		FOD	
L.n	0.686***	(4.67)	0.737***	(5.14)
L2.n	-0.0854	(-1.50)	-0.0960	(-1.38)
w	-0.608**	(-3.36)	-0.563***	(-3.47)
k	0.357***	(5.95)	0.384***	(6.85)
ys	0.609***	(3.47)	0.469**	(2.72)
hansen_df	25		25	
hansenp	0.177		0.170	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The results appear reasonably robust to the choice of transformation, with slightly more precise estimates for most coefficients when the FOD transformation is employed.

We illustrate estimating this equation with both the FD transformation and the forward orthogonal deviations (FOD) transformation:

	First diff		FOD	
L.n	0.686***	(4.67)	0.737***	(5.14)
L2.n	-0.0854	(-1.50)	-0.0960	(-1.38)
w	-0.608**	(-3.36)	-0.563***	(-3.47)
k	0.357***	(5.95)	0.384***	(6.85)
ys	0.609***	(3.47)	0.469**	(2.72)
hansen_df	25		25	
hansenp	0.177		0.170	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The results appear reasonably robust to the choice of transformation, with slightly more precise estimates for most coefficients when the FOD transformation is employed.

We might reasonably consider, as did Blundell and Bond (*J. Econometrics*, 1998), that wages and the capital stock should not be taken as strictly exogenous in this context, as we have in the above models.

Reestimate the equation producing “GMM-style” instruments for all three variables, with both one-step and two-step VCE:

Try it out:

```
xtabond2 n L(1/2).n L(0/1).w L(0/2).(k ys) yr*, gmm(L.(n w k)) ///  
iv(L(0/2).ys yr*) nolevel robust small
```

We might reasonably consider, as did Blundell and Bond (*J. Econometrics*, 1998), that wages and the capital stock should not be taken as strictly exogenous in this context, as we have in the above models.

Reestimate the equation producing “GMM-style” instruments for all three variables, with both one-step and two-step VCE:

Try it out:

```
xtabond2 n L(1/2).n L(0/1).w L(0/2).(k ys) yr*, gmm(L.(n w k)) ///
iv(L(0/2).ys yr*) nolevel robust small
```

	One-step		Two-step	
L.n	0.818***	(9.51)	0.824***	(8.51)
L2.n	-0.112*	(-2.23)	-0.101	(-1.90)
w	-0.682***	(-4.78)	-0.711***	(-4.67)
k	0.353**	(2.89)	0.377**	(2.79)
ys	0.651***	(3.43)	0.662***	(3.89)
hansen_df	74		74	
hansenp	0.487		0.487	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The results from both one-step and two-step estimation appear reasonable. Interestingly, only the coefficient on *ys* appears to be more precisely estimated by the two-step VCE. With no restrictions on the instrument set, 74 overidentifying restrictions are defined, with 90 instruments in total.

	One-step		Two-step	
L.n	0.818***	(9.51)	0.824***	(8.51)
L2.n	-0.112*	(-2.23)	-0.101	(-1.90)
w	-0.682***	(-4.78)	-0.711***	(-4.67)
k	0.353**	(2.89)	0.377**	(2.79)
ys	0.651***	(3.43)	0.662***	(3.89)
hansen_df	74		74	
hansenp	0.487		0.487	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The results from both one-step and two-step estimation appear reasonable. Interestingly, only the coefficient on *ys* appears to be more precisely estimated by the two-step VCE. With no restrictions on the instrument set, 74 overidentifying restrictions are defined, with 90 instruments in total.

To illustrate system GMM, we follow Blundell and Bond, who used the same `abdata` dataset on a somewhat simpler model, dropping the second lags and removing sectoral demand. We consider wages and capital as potentially endogenous, with GMM-style instruments.

Estimate the one-step BB model.

Try it out:

```
xtabond2 n L.n L(0/1).(w k) yr*, gmm(L.(n w k)) iv(yr*, equation(level)) ///  
robust small
```

We indicate here with the `equation(level)` suboption that the year dummies are only to be considered instruments in the level equation. As the default for `xtabond2` is the BB estimator, we omit the `noleveleq` option that has called for the AB estimator in earlier examples.

To illustrate system GMM, we follow Blundell and Bond, who used the same `abdata` dataset on a somewhat simpler model, dropping the second lags and removing sectoral demand. We consider wages and capital as potentially endogenous, with GMM-style instruments.

Estimate the one-step BB model.

Try it out:

```
xtabond2 n L.n L(0/1).(w k) yr*, gmm(L.(n w k)) iv(yr*, equation(level)) ///  
robust small
```

We indicate here with the `equation(level)` suboption that the year dummies are only to be considered instruments in the level equation. As the default for `xtabond2` is the BB estimator, we omit the `noleveleq` option that has called for the AB estimator in earlier examples.

	n	
L.n	0.936***	(35.21)
w	-0.631***	(-5.29)
k	0.484***	(8.89)
hansen_df	100	
hansenp	0.218	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We find that the α coefficient is much higher than in the AB estimates, although it may be distinguished from unity. 113 instruments are created, with 100 degrees of freedom in the test of overidentifying restrictions.

		n
L.n	0.936***	(35.21)
w	-0.631***	(-5.29)
k	0.484***	(8.89)
hansen_df	100	
hansenp	0.218	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We find that the α coefficient is much higher than in the AB estimates, although it may be distinguished from unity. 113 instruments are created, with 100 degrees of freedom in the test of overidentifying restrictions.

A second empirical exercise

We also illustrate DPD estimation using the Penn World Table cross-country panel. We specify a model for k_c (the consumption share of real GDP per capita) depending on its own lag, c_{gnp} , and a set of time fixed effects.

We first estimate the two-step ‘difference GMM’ form of the model with (cluster-)robust VCE, using data for 1991–2007. We could use `testparm i.year` after estimation to evaluate the joint significance of time effects (listing of which has been suppressed).

A second empirical exercise

We also illustrate DPD estimation using the Penn World Table cross-country panel. We specify a model for k_c (the consumption share of real GDP per capita) depending on its own lag, c_{gnp} , and a set of time fixed effects.

We first estimate the two-step ‘difference GMM’ form of the model with (cluster-)robust VCE, using data for 1991–2007. We could use `testparm i.year` after estimation to evaluate the joint significance of time effects (listing of which has been suppressed).

```
. use /Users/cfbaum/Dropbox/baum/IMFOct2013/Datasets/pwt6_3, clear
(Penn World Tables 6.3, August 2009)
. keep if inrange(year, 1991, 2007)
(7790 observations deleted)
. xtabond2 kc L.kc cgnp i.year, gmm(L.kc openc cgnp, lag(2 9)) ///
> iv(i.year) twostep robust noleveleq nodiffsgargan
1991b.year dropped due to collinearity
1992.year dropped due to collinearity
Warning: Number of instruments may be large relative to number of observations.
...
Dynamic panel-data estimation, two-step difference GMM
```

Group variable: iso	Number of obs	=	2794
Time variable : year	Number of groups	=	188
Number of instruments = 283	Obs per group: min	=	12
Wald chi2(17) = 120.15	avg	=	14.86
Prob > chi2 = 0.000	max	=	15

kc	Coef.	Corrected Std. Err.	z	P> z	[95% Conf. Interval]	
kc L1.	.5744329	.0671504	8.55	0.000	.4428206	.7060452
cgnp	.0215235	.1296622	0.17	0.868	-.2326097	.2756568
....						

(continued)

Instruments for first differences equation

Standard

D.(1991b.year 1992.year 1993.year 1994.year 1995.year 1996.year 1997.year
1998.year 1999.year 2000.year 2001.year 2002.year 2003.year 2004.year
2005.year 2006.year 2007.year)

GMM-type (missing=0, separate instruments for each period unless collapsed)
L(2/9).(L.kc openc cgnp)

Arellano-Bond test for AR(1) in first differences: z = -4.71 Pr > z = 0.000

Arellano-Bond test for AR(2) in first differences: z = 0.38 Pr > z = 0.704

Sargan test of overid. restrictions: chi2(266) = 575.10 Prob > chi2 = 0.000

(Not robust, but not weakened by many instruments.)

Hansen test of overid. restrictions: chi2(266) = 184.05 Prob > chi2 = 1.000

(Robust, but weakened by many instruments.)

Given the relatively large number of time periods available, I have specified that the GMM instruments only be constructed for lags 2–9 to keep the number of instruments manageable. I am treating `openc` as a GMM-style instrument. The autoregressive coefficient is 0.574, and the `cgnp` coefficient is positive but insignificant. Although not shown, the test for joint significance of the time effects has p-value 0.335.

We could also fit this model with the ‘system GMM’ estimator, which will be able to utilize one more observation per country in the level equation, and estimate a constant term in the relationship. I am treating lagged `openc` as a IV-style instrument in this specification.

Given the relatively large number of time periods available, I have specified that the GMM instruments only be constructed for lags 2–9 to keep the number of instruments manageable. I am treating $open_c$ as a GMM-style instrument. The autoregressive coefficient is 0.574, and the $cgnp$ coefficient is positive but insignificant. Although not shown, the test for joint significance of the time effects has p-value 0.335.

We could also fit this model with the ‘system GMM’ estimator, which will be able to utilize one more observation per country in the level equation, and estimate a constant term in the relationship. I am treating lagged $open_c$ as a IV-style instrument in this specification.

```
. xtabond2 kc L.kc cgnp i.year, gmm(L.kc cgnp, lag(2 8)) ///
> iv(i.year L.openc) twostep robust nodiffsargan
1991b.year dropped due to collinearity
2007.year dropped due to collinearity
Warning: Number of instruments may be large relative to number of observations.
...
```

Dynamic panel-data estimation, two-step system GMM

Group variable: iso	Number of obs	=	2982
Time variable : year	Number of groups	=	188
Number of instruments = 207	Obs per group: min	=	13
Wald chi2(17) = 5852.66	avg	=	15.86
Prob > chi2 = 0.000	max	=	16

kc	Coef.	Corrected Std. Err.	z	P> z	[95% Conf. Interval]	
kc L1.	.988957	.0164111	60.26	0.000	.9567917	1.021122
cgnp	.0383551	.0184993	2.07	0.038	.0020971	.0746131
...						

(continued)

Instruments for first differences equation

Standard

D.(1991b.year 1992.year 1993.year 1994.year 1995.year 1996.year 1997.year
 1998.year 1999.year 2000.year 2001.year 2002.year 2003.year 2004.year
 2005.year 2006.year 2007.year L.openc)

GMM-type (missing=0, separate instruments for each period unless collapsed)
 L(2/8).(L.kc cgnp)

Instruments for levels equation

Standard

1991b.year 1992.year 1993.year 1994.year 1995.year 1996.year 1997.year
 1998.year 1999.year 2000.year 2001.year 2002.year 2003.year 2004.year
 2005.year 2006.year 2007.year L.openc

_cons

GMM-type (missing=0, separate instruments for each period unless collapsed)
 DL.(L.kc cgnp)

Arellano-Bond test for AR(1) in first differences: z = -5.47 Pr > z = 0.000

Arellano-Bond test for AR(2) in first differences: z = 0.49 Pr > z = 0.621

Sargan test of overid. restrictions: chi2(189) = 352.18 Prob > chi2 = 0.000
 (Not robust, but not weakened by many instruments.)

Hansen test of overid. restrictions: chi2(189) = 180.40 Prob > chi2 = 0.661
 (Robust, but weakened by many instruments.)

Note that the autoregressive coefficient is much larger: 0.989 in this context. The c_{gnp} coefficient is positive and significant. The Hansen test of this model has some power, in contrast to the difference GMM specification.

We can also estimate the model using the forward orthogonal deviations (FOD) transformation of Arellano and Bover, as described in Roodman's paper. The first-difference transformation applied in DPD estimators has the unfortunate feature of magnifying any gaps in the data, as one period of missing data is replaced with two missing differences. FOD transforms each observation by subtracting the average of all *future* observations, which will be defined (regardless of gaps) for all but the last observation in each panel. To illustrate:

Note that the autoregressive coefficient is much larger: 0.989 in this context. The c_{gnp} coefficient is positive and significant. The Hansen test of this model has some power, in contrast to the difference GMM specification.

We can also estimate the model using the forward orthogonal deviations (FOD) transformation of Arellano and Bover, as described in Roodman's paper. The first-difference transformation applied in DPD estimators has the unfortunate feature of magnifying any gaps in the data, as one period of missing data is replaced with two missing differences. FOD transforms each observation by subtracting the average of all *future* observations, which will be defined (regardless of gaps) for all but the last observation in each panel. To illustrate:

```
. xtabond2 kc L.kc cgnp i.year, gmm(L.kc cgnp, lag(2 8)) ///
> iv(i.year L.openc) twostep robust nodiffsargan orthog
1991b.year dropped due to collinearity
1992.year dropped due to collinearity
Warning: Number of instruments may be large relative to number of observations.
...
```

Dynamic panel-data estimation, two-step system GMM

Group variable: iso	Number of obs	=	2982
Time variable : year	Number of groups	=	188
Number of instruments = 207	Obs per group: min	=	13
Wald chi2(17) = 9421.59	avg	=	15.86
Prob > chi2 = 0.000	max	=	16

kc	Coef.	Corrected Std. Err.	z	P> z	[95% Conf. Interval]	
kc L1.	.9844427	.0150613	65.36	0.000	.9549231	1.013962
cgnp	.00181	.0229043	0.08	0.937	-.0430816	.0467016
...						

(continued)

Instruments for orthogonal deviations equation

Standard

FOD.(1991b.year 1992.year 1993.year 1994.year 1995.year 1996.year
 1997.year 1998.year 1999.year 2000.year 2001.year 2002.year 2003.year
 2004.year 2005.year 2006.year 2007.year L.openc)

GMM-type (missing=0, separate instruments for each period unless collapsed)

L(2/8).(L.kc cgnp)

Instruments for levels equation

Standard

1991b.year 1992.year 1993.year 1994.year 1995.year 1996.year 1997.year
 1998.year 1999.year 2000.year 2001.year 2002.year 2003.year 2004.year
 2005.year 2006.year 2007.year L.openc

_cons

GMM-type (missing=0, separate instruments for each period unless collapsed)

DL.(L.kc cgnp)

Arellano-Bond test for AR(1) in first differences: z = -5.47 Pr > z = 0.000

Arellano-Bond test for AR(2) in first differences: z = 0.52 Pr > z = 0.605

Sargan test of overid. restrictions: chi2(189) = 348.68 Prob > chi2 = 0.000

(Not robust, but not weakened by many instruments.)

Hansen test of overid. restrictions: chi2(189) = 178.45 Prob > chi2 = 0.698

(Robust, but weakened by many instruments.)

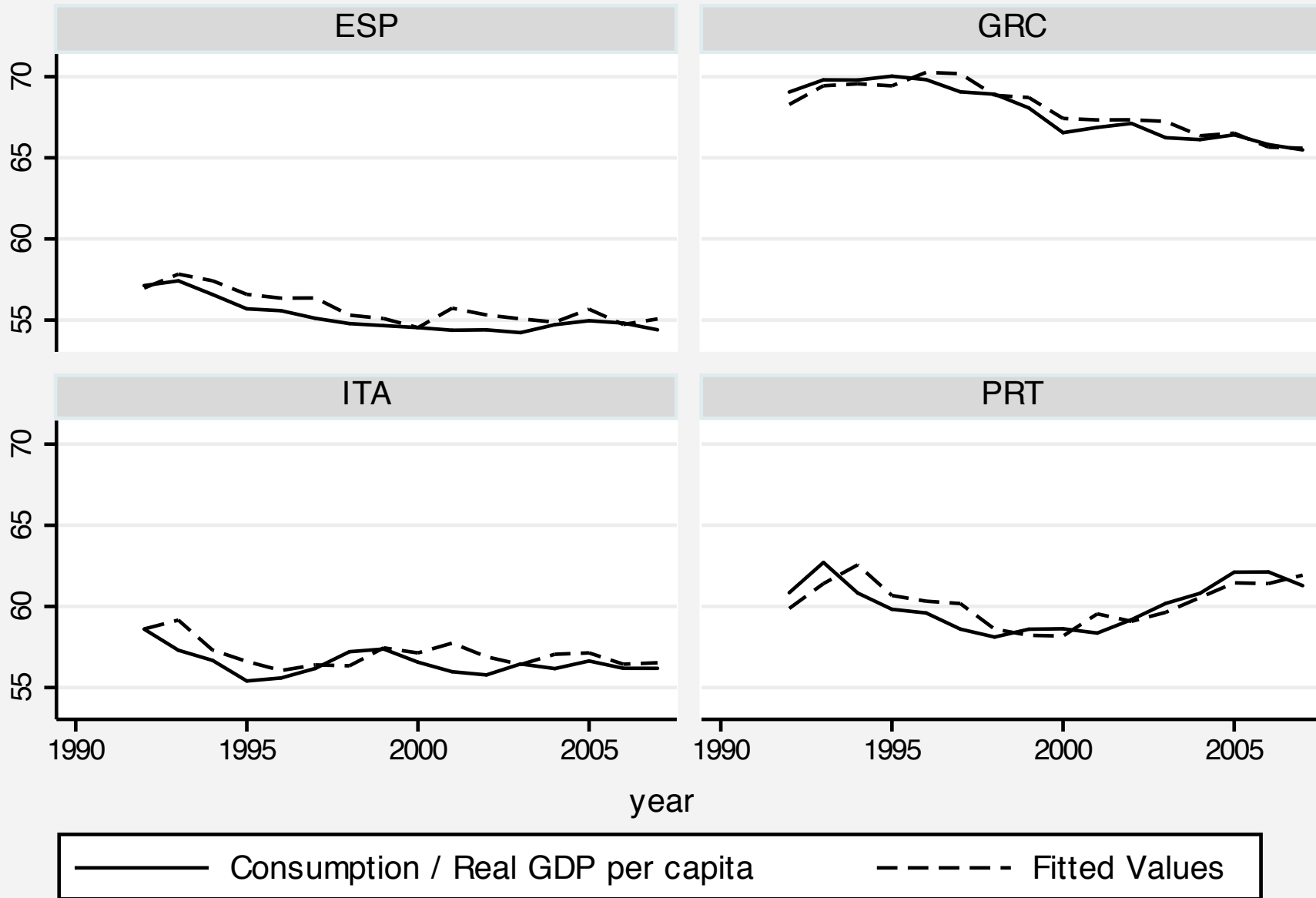
Using the FOD transformation, the autoregressive coefficient is a bit smaller, and the $cgnp$ coefficient loses its significance.

After any DPD estimation command, we may save predicted values or residuals and graph them against the actual values:

Using the FOD transformation, the autoregressive coefficient is a bit smaller, and the c_{gnp} coefficient loses its significance.

After any DPD estimation command, we may save predicted values or residuals and graph them against the actual values:

```
. predict double kchat if inlist(country, "Italy", "Spain", "Greece", "Portugal  
> ")  
(option xb assumed; fitted values)  
(1619 missing values generated)  
. label var kc "Consumption / Real GDP per capita"  
. xtline kc kchat if !mi(kchat), scheme(s2mono)
```

Graphs by ISO country code

Although the DPD estimators are linear estimators, they are highly sensitive to the particular specification of the model and its instruments: more so in my experience than any other regression-based estimation approach.

There is no substitute for experimentation with the various parameters of the specification to ensure that your results are reasonably robust to variations in the instrument set and lags used. A very useful reference for DPD modeling is David Roodman's paper "How to do `xtabond2`", available from <http://ideas.repec.org>.

Although the DPD estimators are linear estimators, they are highly sensitive to the particular specification of the model and its instruments: more so in my experience than any other regression-based estimation approach.

There is no substitute for experimentation with the various parameters of the specification to ensure that your results are reasonably robust to variations in the instrument set and lags used. A very useful reference for DPD modeling is David Roodman's paper "How to do `xtabond2`", available from <http://ideas.repec.org>.