Dynamic Linear Panel Data

Example Questions and Solutions

230347: Advanced Microeconometrics

Question 1

Nickell Bias

Consider the model:

$$y_{it} = \rho y_{it-1} + \alpha_i + \varepsilon_{it} \qquad i = 1, \ldots, N \quad t = 1, 2$$

Notice that $T = 2$. Assume that y_{i0} is observed for all i. $\mathbb{E}[\alpha_i] = \mathbb{E}[\varepsilon_{it}] = \mathbb{E}[\alpha_i \varepsilon_{it}] = 0$. Assume ε_{it} is iid with variance $\sigma_\varepsilon^2 > 0$. Assume further that $|\rho| < 1$. The within transformation of the above model yields:

$$y_{it} - \underbrace{\frac{y_{i1} + y_{i2}}{2}}_{=\bar{y}_i} = \rho \left(y_{it-1} - \underbrace{\frac{y_{i0} + y_{i1}}{2}}_{=\bar{y}_{i,-1}} \right) + \varepsilon_{it} - \underbrace{\frac{\varepsilon_{i1} + \varepsilon_{i2}}{2}}_{=\bar{\varepsilon}_i} \qquad i = 1, \ldots, N \quad t = 1, 2$$

(i) Write down the within estimator for ρ (call it $\widehat{\rho}$) and show that:

$$\widehat{\rho} - \rho = \frac{\frac{1}{2N} \sum_{i=1}^N \sum_{t=1}^2 \left(y_{it-1} - \bar{y}_{i,-1}\right) \left(\varepsilon_{it} - \bar{\varepsilon}_i\right)}{\frac{1}{2N} \sum_{i=1}^N \sum_{t=1}^2 \left(y_{it-1} - \bar{y}_{i,-1}\right)^2}$$

(ii) Show that the above is equal to:

$$\widehat{\rho} - \rho = \frac{\frac{1}{2N} \sum_{i=1}^N \left(y_{i1} - y_{i0}\right) \left(\varepsilon_{i2} - \varepsilon_{i1}\right)}{\frac{1}{2N} \sum_{i=1}^N \left(y_{i1} - y_{i0}\right)^2}$$

(iii) Show that the probability limit of the numerator of the above expression as $N \to \infty$ is $-\frac{\sigma_\varepsilon^2}{2}$.

(iv) Show that the probability limit of the denominator of the above expression as $N \to \infty$ is $\frac{\sigma_\varepsilon^2}{1+\rho}$. For this you can use that since y_{it} is stationary ($|\rho| < 1$), $y_{it} = \frac{\alpha_i}{1-\rho} + \sum_{j=0}^\infty \rho^j \varepsilon_{it-j}$.

(v) Use these to show that $\operatorname*{plim}_{N\to\infty} \widehat{\rho}_N - \rho = -\frac{(1+\rho)}{2}$

Solution

(i) The within estimator for ρ is:

$$\widehat{\rho} = \frac{\sum_{i=1}^N \sum_{t=1}^2 \left(y_{it-1} - \bar{y}_{i,-1}\right) \left(y_{it} - \bar{y}_i\right)}{\sum_{i=1}^N \sum_{t=1}^2 \left(y_{it-1} - \bar{y}_{i,-1}\right)^2}$$

Inserting $\rho(y_{it-1} - \bar{y}_{i,-1}) + \varepsilon_{it} - \bar{\varepsilon}_i$ for $y_{it} - \bar{y}_i$ and cancelling terms:

$$\begin{aligned}
\widehat{\rho} &= \frac{\sum_{i=1}^{N}\sum_{t=1}^{2}(y_{it-1} - \bar{y}_{i,-1})\left[\rho(y_{it-1} - \bar{y}_{i,-1}) + \varepsilon_{it} - \bar{\varepsilon}_i\right]}{\sum_{i=1}^{N}\sum_{t=1}^{2}(y_{it-1} - \bar{y}_{i,-1})^2} \\
\widehat{\rho} - \rho &= \frac{\frac{1}{2N}\sum_{i=1}^{N}\sum_{t=1}^{2}(y_{it-1} - \bar{y}_{i,-1})(\varepsilon_{it} - \bar{\varepsilon}_i)}{\frac{1}{2N}\sum_{i=1}^{N}\sum_{t=1}^{2}(y_{it-1} - \bar{y}_{i,-1})^2}
\end{aligned}$$

(ii) Since $T = 2$ we can rewrite the within transformation in first differences. One i term in the numerator is:

$$\begin{aligned}
&(y_{i0} - \bar{y}_{i,-1})(\varepsilon_{i1} - \bar{\varepsilon}_i) + (y_{i1} - \bar{y}_{i,-1})(\varepsilon_{i2} - \bar{\varepsilon}_i) \\
=&\left(y_{i0} - \frac{y_{i0} + y_{i1}}{2}\right)\left(\varepsilon_{i1} - \frac{\varepsilon_{i1} + \varepsilon_{i2}}{2}\right) + \left(y_{i1} - \frac{y_{i0} + y_{i1}}{2}\right)\left(\varepsilon_{i2} - \frac{\varepsilon_{i1} + \varepsilon_{i2}}{2}\right) \\
=&\left(\frac{y_{i0} - y_{i1}}{2}\right)\left(\frac{\varepsilon_{i1} - \varepsilon_{i2}}{2}\right) + \left(\frac{y_{i1} - y_{i0}}{2}\right)\left(\frac{\varepsilon_{i2} - \varepsilon_{i1}}{2}\right) \\
=&\frac{1}{2}(y_{i1} - y_{i0})(\varepsilon_{i2} - \varepsilon_{i1})
\end{aligned}$$

One i term in the denominator is:

$$\begin{aligned}
(y_{i0} - \bar{y}_{i,-1})^2 + (y_{i1} - \bar{y}_{i,-1})^2 &= \left(y_{i0} - \frac{y_{i0} + y_{i1}}{2}\right)^2 + \left(y_{i1} - \frac{y_{i0} + y_{i1}}{2}\right)^2 \\
&= \left(\frac{y_{i0} - y_{i1}}{2}\right)^2 + \left(\frac{y_{i1} - y_{i0}}{2}\right)^2 \\
&= \frac{(y_{i1} - y_{i0})^2}{2}
\end{aligned}$$

The bias can then be written as:

$$\begin{aligned}
\widehat{\rho} - \rho &= \frac{\frac{1}{2N}\sum_{i=1}^{N}\sum_{t=1}^{2}(y_{it-1} - \bar{y}_{i,-1})(\varepsilon_{it} - \bar{\varepsilon}_i)}{\frac{1}{2N}\sum_{i=1}^{N}\sum_{t=1}^{2}(y_{it-1} - \bar{y}_{i,-1})^2} \\
&= \frac{\frac{1}{2N}\sum_{i=1}^{N}(y_{i1} - y_{i0})(\varepsilon_{i2} - \varepsilon_{i1})}{\frac{1}{2N}\sum_{i=1}^{N}(y_{i1} - y_{i0})^2}
\end{aligned}$$

(iii) Expanding the terms in the numerator:

$$\begin{aligned}
&\frac{1}{2N}\sum_{i=1}^{N} y_{i1}\varepsilon_{i2} - \frac{1}{2N}\sum_{i=1}^{N} y_{i1}\varepsilon_{i1} - \frac{1}{2N}\sum_{i=1}^{N} y_{i0}\varepsilon_{i2} + \frac{1}{2N}\sum_{i=1}^{N} y_{i0}\varepsilon_{i1} \\
=&\frac{1}{2N}\sum_{i=1}^{N} y_{i1}\varepsilon_{i2} - \frac{1}{2N}\sum_{i=1}^{N}(\rho y_{i0} + \alpha_i + \varepsilon_{i1})\varepsilon_{i1} - \frac{1}{2N}\sum_{i=1}^{N} y_{i0}\varepsilon_{i2} + \frac{1}{2N}\sum_{i=1}^{N} y_{i0}\varepsilon_{i1} \\
=&\frac{1}{2N}\sum_{i=1}^{N} y_{i1}\varepsilon_{i2} - \frac{1}{2N}\sum_{i=1}^{N}\rho y_{i0}\varepsilon_{i1} + \frac{1}{2N}\sum_{i=1}^{N}\alpha_i\varepsilon_{i1} + \frac{1}{2N}\sum_{i=1}^{N}\varepsilon_{i1}^2 - \frac{1}{2N}\sum_{i=1}^{N} y_{i0}\varepsilon_{i2} + \frac{1}{2N}\sum_{i=1}^{N} y_{i0}\varepsilon_{i1}
\end{aligned}$$

Taking the probality limit as $N \to \infty$, all terms except $\frac{1}{2N}\sum_{i=1}^{N}\varepsilon_{i1}^2$ go to zero. This remaining term converges to $-\frac{\sigma_\varepsilon^2}{2}$.

(iv) The denominator:

$$
\begin{aligned}
& \operatorname*{plim}_{N\to\infty} \frac{1}{2N}\sum_{i=1}^{N}\left(y_{i1}-y_{i0}\right)^2 \\
&= \operatorname*{plim}_{N\to\infty} \frac{1}{2N}\sum_{i=1}^{N}\left[\left(\frac{\alpha_i}{1-\rho}+\sum_{j=0}^{\infty}\rho^j\varepsilon_{i,1-j}\right)-\left(\frac{\alpha_i}{1-\rho}+\sum_{j=0}^{\infty}\rho^j\varepsilon_{i,0-j}\right)\right]^2 \\
&= \operatorname*{plim}_{N\to\infty} \frac{1}{2N}\sum_{i=1}^{N}\left(\sum_{j=0}^{\infty}\rho^j\varepsilon_{i,1-j}-\sum_{j=0}^{\infty}\rho^j\varepsilon_{i,0-j}\right)^2 \\
&= \operatorname*{plim}_{N\to\infty} \frac{1}{2N}\sum_{i=1}^{N}\left[\left(\sum_{j=0}^{\infty}\rho^j\varepsilon_{i,1-j}\right)^2-2\left(\sum_{j=0}^{\infty}\rho^j\varepsilon_{i,1-j}\right)\left(\sum_{j=0}^{\infty}\rho^j\varepsilon_{i,0-j}\right)+\left(\sum_{j=0}^{\infty}\rho^j\varepsilon_{i,0-j}\right)^2\right] \\
&= \frac{\sigma_\varepsilon^2}{2\left(1-\rho^2\right)}-\frac{\rho\sigma_\varepsilon^2}{\left(1-\rho^2\right)}+\frac{\sigma_\varepsilon^2}{2\left(1-\rho^2\right)} \\
&= \frac{\sigma_\varepsilon^2\left(1-\rho\right)}{\left(1-\rho^2\right)} \\
&= \frac{\sigma_\varepsilon^2\left(1-\rho\right)}{\left(1-\rho\right)\left(1+\rho\right)} \\
&= \frac{\sigma_\varepsilon^2}{1+\rho}
\end{aligned}
$$

(v) Dividing the answers in (ii) and (iii) yields the result:

$$\frac{-\frac{\sigma_\varepsilon^2}{2}}{\frac{\sigma_\varepsilon^2}{1+\rho}} = -\frac{(1+\rho)}{2}$$

Question 2

Acemoglu, Johnson, Robinson, Yared (2008)[1] estimate the following model:

$$d_{it} = \rho d_{it-1} + \gamma y_{it-1} + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \alpha_i + \delta_t + \varepsilon_{it}$$

where d_{it} is a measure of democracy for country i in year t, y_{it} is the log GDP per capita of the country, and $\boldsymbol{x}_{it}$ are additional control variables. A regression table from the paper is shown below:

[1] Acemoglu, D., Johnson, S., Robinson, J. A., & Yared, P. (2008). "Income and democracy". *American Economic Review*, 98(3), 808-42

TABLE 2—FIXED EFFECTS RESULTS USING FREEDOM HOUSE MEASURE OF DEMOCRACY

	Base sample, 1960–2000								
	Five-year data				Annual data	Ten-year data		Twenty-year data
	Pooled OLS (1)	Fixed effects OLS (2)	Anderson-Hsiao IV (3)	Arellano-Bond GMM (4)	Fixed effects OLS (5)	Fixed effects OLS (6)	Fixed effects OLS (7)	Arellano-Bond GMM (8)	Fixed effects OLS (9)
	Dependent variable is democracy								
Democracy_{t-1}	0.706	0.379	0.469	0.489		[0.00]	−0.025	0.226	−0.581
	(0.035)	(0.051)	(0.100)	(0.085)			(0.088)	(0.123)	(0.198)
Log GDP per capita_{t-1}	0.072	0.010	−0.104	−0.129	0.054	[0.33]	0.053	−0.318	−0.030
	(0.010)	(0.035)	(0.107)	(0.076)	(0.046)		(0.066)	(0.180)	(0.156)
Hansen *J* test				[0.26]				[0.07]	
AR(2) test				[0.45]				[0.96]	
Implied cumulative	0.245	0.016	−0.196	−0.252				−0.411	−0.019
effect of income	[0.00]	[0.76]	[0.33]	[0.09]				[0.09]	[0.85]
Observations	945	945	838	838	958	2895	457	338	192
Countries	150	150	127	127	150	148	127	118	118
R-squared	0.73	0.80			0.76	0.93	0.77		0.89

Notes: Pooled cross-sectional OLS regression in column 1, with robust standard errors clustered by country in parentheses. Fixed effects OLS regressions in columns 2, 5, 6, 7, and 9, with country dummies and robust standard errors clustered by country in parentheses. Implied cumulative effect of income represents the coefficient estimate of log GDP per capita $_{t-1}$/(1−democracy$_{t-1}$), and the *p*-value from a nonlinear test of the significance of this coefficient is in brackets. Column 3 uses the instrumental variables method of Theodore W. Anderson and Cheng Hsiao (1982), with clustered standard errors, and columns 4 and 8 use the GMM of Manuel Arellano and Stephen R. Bond (1991), with robust standard errors; in both methods we instrument for income using a double lag. Year dummies are included in all regressions. Dependent variable is Freedom House measure of democracy. Base sample is an unbalanced panel, 1960–2000, with data at five-year intervals, where the start date of the panel refers to the dependent variable (i.e., $t = 1960$, so $t - 1 = 1955$); column 6 uses annual data from the same sample; a country must be independent for five years before it enters the panel. Columns 7 and 8 use ten-year data from the same sample, where, as before, the start date of the panel refers to the dependent variable (i.e., $t = 1960$, so $t - 1 = 1950$); a country must be independent for ten years before it enters the panel. Column 9 uses twenty-year data from the same sample, where, as before, the start date of the panel refers to the dependent variable (i.e., $t = 1980$, so $t - 1 = 1960$); a country must be independent for twenty years before it enters the panel. In column 6, each right-hand-side variable has five annual lags; we report the *p*-value from an *F*-test for the joint significance of all five lags. For detailed data definitions and sources, see Table 1 and Appendix Table A1.

(i) Why would the estimated coefficient on y_{it-1} be larger in column 1 compared to column 2?

(ii) Why would the estimated coefficient on d_{it-1} be smaller in column 2 compared to column 4?

(iii) Why would the estimated coefficient on d_{it-1} be *even smaller* in column 7 compared to column 8 relative to those compared in question (ii)?

(iv) Why would the standard errors of the coefficients in column 4 be smaller than those in column 3?

(v) Interpret the result of the Hansen J test in column 4 (the p-value is shown). Note that Hansen J test is the same as the Sargan test.

(vi) Interpret the result of the AR(2) test in column 4 (the p-value is shown).

Example Solution

(i) If an unobservable time-invariant characteristic of a country that is positively correlated with y_{it-1} also affects its democracy score, d_{it}, then the coefficient on y_{it-1} will be biased upwards if we omit the fixed effect from the regression.

(ii) Using fixed effects with a lagged dependent variable with not many time observations per country (here on average 6 per country) causes Nickell bias which will bias the coefficient downwads, which is what we observe here.

(iii) Here the Nickell bias is even stronger because we have fewer time periods (on average only 3 per country).

(iv) The Arellano-Bond estimates are more efficient than the Anderson-Hsiao estimate as they use more moment conditions (further lags as additional instruments).

(v) The null hypothesis is that the instruments are valid. The large p-value indicates that we do not reject the null. There is no evidence that the instruments are invalid.

(vi) The large p-value indicates that we cannot reject the null of no second-order serial correlation.