Advanced Panel Data: Dynamic Panel

Chris Conlon

May 14, 2020

NYU Stern

Dynamic Panel Data

Dynamic Panel

- Suppose that we also want to include a lagged $y_{i,t-1}$

$$y_{it} = \rho y_{i,t-1} + x'_{it}\beta + \eta_i + \varepsilon_{it}$$

- We can treat η_i as a random effect or a fixed effect.

Dynamic Panel: Nickell (1981) Bias

Consider the within transform

$$(y_{it} - \overline{y}_i) = \rho(y_{i,t-1} - \overline{y}_i) + (x_{it} - \overline{x}_i)'\beta + (\varepsilon_{it} - \overline{\varepsilon}_i)$$

- This eliminates the fixed effect.
- But $Cov(y_{i,t-1} - \overline{y}_i, \varepsilon_{it} - \overline{\varepsilon}_i) \neq 0$. Why?
  - Both contain past and future values
  - There is a direct relationship between y and ε
  - Bias does not disappear as $N \to \infty$ (it does as $T \to \infty$).
  - For small T, dynamic panel model is inconsistent.

Dynamic Panel: Bias Alternative

$$y_{it} = \rho y_{i,t-1} + x'_{it}\beta + \eta_i + \varepsilon_{it}$$

- We require the following assumption (strict exogeneity):

$$E\left(\varepsilon_{it} | x_{i1}, \ldots, x_{iT}, \eta_i\right) = 0, \quad t = 1, \ldots, T$$

- But what about y_{it-1}?
  - It is correlated with $\varepsilon_{i,t-1}$ and η_i (by construction).
  - With serial correlation it is correlated with ε_{it}
  - This is the usual endogeneity concern.

Dynamic Panel: Differenced Model (Anderson-Hsiao)

How do we deal with endogeneity? With instruments!

$$y_{it} = \rho y_{i,t-1} + x_{it}'\beta + \eta_i + \varepsilon_{it}$$

Consider the first differences (s is a dummy time index):

$$E\left[x_{is}\left(\Delta y_{it} - \rho\Delta y_{i(t-1)} - \Delta x_{it}'\beta\right)\right] = 0$$

Idea:

- Under strict exogeneity of x_{it} we can use both lags and leads as instruments for $y_{i,t-1}$
- Excluded Instruments $x_{i,s}$ do not have a direct effect on $\Delta y_{i,t-1}$.
- These moments work even in presence of serially correlated errors.

Minimal Example: Anderson-Hsiao

Imagine we have only $T = 3$ periods:

$$y_3 - y_2 = \alpha\left(y_2 - y_1\right) + \beta_0\left(x_3 - x_2\right) + \beta_1\left(x_2 - x_1\right) + \left(\varepsilon_3 - \varepsilon_2\right)$$

- $E\left(x_{is}\Delta\varepsilon_{i3}\right) = 0$ has three instruments $\left(x_{i1}, x_{i2}, x_{i3}\right)$.
- The model is just identified with 3 parameters $\left(\alpha, \beta_0, \beta_1\right)$.
- The challenge with this approach is often that it suffers from weak instruments.
    - We need a lot of variation in x_{is} across t to make this work.
    - A constant trend isn't going to help either.

Becker, Grossman, Murphy (1994)

Study annual cigarette consumption with state-level data:

$$c_{it} = \theta c_{i,t-1} + \beta\theta c_{i,t+1} + \gamma p_{it} + \eta_i + \delta_t + v_{it}$$

A model of (forward looking) rational addiction:

- c_{it} = Annual per capita cigarette consumption in packs by state.
- p_{it} = Average cigarette price per pack.
- θ = Measure of the extent of addiction (for $\theta > 0$).
- β = Discount factor.
- Derived from forward looking model of habit formation FOC's.

Becker, Grossman, Murphy (1994)

$$c_{it} = \theta c_{i,t-1} + \beta\theta c_{i,t+1} + \gamma p_{it} + \eta_i + \delta_t + v_{it}$$

- Marginal utility of wealth can show up in γ or η_i.
- The errors v_{it} are unobserved life-cycle utility shifters, can be autocorrelated.
- Absent addiction $\theta = 0$ and serial correlation in prices, we would expect to find dependence over time in c_{it}.
- Conditional on $c_{i,t}|(c_{i,t-1}, c_{i,t+1})$ does not depend on $p_{i,t+1}$ or $p_{i,t+1}$.

Becker, Grossman, Murphy (1994)

$$c_{it} = \theta c_{i,t-1} + \beta\theta c_{i,t+1} + \gamma p_{it} + \eta_i + \delta_t + v_{it}$$

- Identify (θ, β, γ) from the assumption that prices are strictly exogenous
- Use lagged and future $p_{i,t+s}$ and $p_{i,t-s}$ as IV.
- Use the change in cigarette taxes.
- Consumers need to fully anticipate future price changes for this to work.

Becker, Grossman, Murphy (1994): Table 1

TABLE 1—DEFINITIONS, MEANS, AND STANDARD DEVIATIONS (SD) OF VARIABLES

Variable	Definition (mean, SD)
C_t	Per capita cigarette consumption in packs in fiscal year t, as derived from state tax-paid sales (mean = 126.171, SD = 31.794)
P_t	Average retail cigarette price per pack in January of fiscal year t in 1967 cents (mean = 29.812, SD = 3.184)
income	Per capita income on a fiscal-year basis, in hundreds of 1967 dollars (mean = 31.439, SD = 8.092)
ℓdtax	Index which measures the incentives to smuggle cigarettes long distance from Kentucky, Virginia, or North Carolina. The index is positively related to the difference between the state's excise tax and the excise taxes of the exporting states (mean = 0.160, SD = 15.572)
sdtexp	Index which measures short-distance (export) smuggling incentives. The index is a weighted average of differences between the exporting state's excise tax and excise taxes of neighboring states, with weights based on border populations (mean = −0.828, SD = 1.847)
sdtimp	Index which measures short-distance (import) smuggling incentives in a state. Similar to sdtexp (mean = 0.494, SD = 0.792)
tax	Sum of state and local excise taxes on cigarettes in 1967 cents per pack (mean = 6.582, SD = 2.651)

Becker, Grossman, Murphy (1994): Table 2

TABLE 2—ESTIMATES OF MYOPIC MODELS OF ADDICTION, DEPENDENT VARIABLE = C_t
(ASYMPTOTIC t STATISTICS IN PARENTHESES)

Independent variable	2SLS (i)	2SLS (ii)	2SLS (iii)	OLS (iv)
C_{t-1}	0.478 (12.07)	0.502 (14.68)	0.602 (21.43)	0.755 (64.84)
P_t	−1.603 (10.12)	−1.538 (10.48)	−1.269 (9.74)	−0.860 (8.33)
Y_t	0.942 (7.61)	0.903 (7.71)	0.741 (6.96)	0.493 (5.44)
ℓ dtax	−0.240 (7.33)	−0.233 (7.40)	0.212 (7.22)	−0.160 (6.17)
sdtimp	−1.541 (5.04)	−1.514 (5.09)	−1.372 (4.97)	−1.228 (4.84)
sdtexp	−3.659 (13.24)	−3.544 (13.88)	−3.059 (13.71)	−2.328 (13.15)
R^2:	0.969	0.970	0.976	0.979
Wu F ratio:	84.76	94.42	41.61	—
N:	1,415	1,415	1,371	1,415

Notes: Intercepts are not shown. Regressors include state and year dummy variables. Columns (i)–(iii) give two-stage least squares (2SLS) estimates with C_{t-1} treated as endogenous. Column (iv) gives an ordinary least-squares (OLS) estimate. The instruments in column (i) consist of the one-period lag of price plus the other explanatory variables in the model. Column (ii) adds the current and one-period lag values of the state cigarette tax to the instruments, and column (iii) further adds two additional lags of the price and tax variables. The Wu F ratios pertain to tests of the hypothesis that the OLS models corresponding to the first three columns are consistent. They all are significant at the 1-percent level.

Becker, Grossman, Murphy (1994): Table 3

TABLE 3—ESTIMATES OF RATIONAL MODELS OF ADDICTION, DEPENDENT VARIABLE = C_t (ASYMPTOTIC t STATISTICS IN PARENTHESES)

Independent variable	2SLS (i)	2SLS (ii)	2SLS (iii)	2SLS (iv)	OLS (v)
C_{t-1}	0.418	0.373	0.443	0.481	0.485
	(8.88)	(9.18)	(11.72)	(14.58)	(36.92)
C_{t+1}	0.135	0.236	0.169	0.228	0.423
	(2.45)	(5.04)	(3.79)	(5.87)	(28.61)
P_t	−1.388	−1.230	−1.227	−0.971	−0.412
	(8.94)	(9.11)	(9.11)	(8.36)	(4.98)
Y_t	0.837	0.761	0.746	0.608	0.302
	(7.34)	(7.44)	(7.31)	(6.72)	(4.21)
ℓdtax	−0.188	−0.150	−0.164	−0.127	−0.022
	(5.42)	(4.82)	(5.30)	(4.50)	(1.05)
sdtimp	−1.358	−1.222	−1.266	−1.090	−0.748
	(4.82)	(4.70)	(4.88)	(4.63)	(3.73)
sdtexp	−3.218	−2.892	−2.914	−2.401	−1.347
	(11.37)	(11.84)	(11.96)	(11.58)	(9.39)
R^2:	0.975	0.978	0.978	0.983	0.987
Wu F ratio:	87.15	85.13	82.63	46.62	—
N:	1,415	1,415	1,415	1,371	1,415

Notes: Intercepts are not shown. Regressors include state and year dummy variables. Columns (i)–(iv) give two-stage least-squares (2SLS) estimates with C_{t-1} and C_{t+1} treated as endogenous. Column (v) gives an ordinary least-squares (OLS) estimate. The instruments in column (i) consist of the one-period lag and lead of price plus the other explanatory variables in the model. Column (ii) adds the current and one-period lag values of the state cigarette tax to the instruments; column (iii) further adds the one-period lead of the tax; and column (iv) further adds two additional lags of the price and tax variables. The Wu F ratios pertain to tests of the hypothesis that the OLS models corresponding to the first four columns are consistent. They all are significant at the 1-percent level.

Becker, Grossman, Murphy (1994): Table 4

TABLE 4—PRICE ELASTICITIES FOR TWO-STAGE LEAST-SQUARES MODELS (APPROXIMATE t STATISTICS IN PARENTHESES)

Elasticity	(i)	(ii)	(iii)	(iv)
Long-run	−0.734	−0.743	−0.747	−0.788
	(13.06)	(12.43)	(12.43)	(10.67)
Own price:				
Anticipated	−0.373	−0.361	−0.346	−0.306
	(10.73)	(11.13)	(10.86)	(9.87)
Unanticipated	−0.349	−0.322	−0.316	−0.262
	(9.97)	(10.09)	(10.10)	(9.20)
Future price, unanticipated	−0.050	−0.084	−0.058	−0.068
	(2.37)	(4.90)	(3.70)	(5.14)
Past price, unanticipated	−0.155	−0.133	−0.152	−0.144
	(8.99)	(8.01)	(9.80)	(9.43)
Short-run	−0.407	−0.436	−0.387	−0.355
	(9.34)	(9.51)	(9.69)	(8.80)

Becker, Grossman, Murphy (1994): Table 5

TABLE 5—TWO-STAGE LEAST-SQUARES ESTIMATES OF RATIONAL-ADDICTION MODELS, FUTURE PRICE AND TAX EXCLUDED FROM SET OF INSTRUMENTS, DEPENDENT VARIABLE = C_t (ASYMPTOTIC t STATISTICS IN PARENTHESES)

Independent variable	Model (i)	Model (ii)	Model (iv)
C_{t-1}	−0.235	0.139	0.109
	(1.03)	(2.25)	(1.69)
C_{t+1}	1.601	0.737	0.887
	(3.75)	(6.62)	(8.55)
P_t	0.865	−0.472	−0.164
	(1.39)	(2.33)	(0.89)
Y_t	−0.217	0.397	0.258
	(−0.67)	(3.19)	(2.14)
ℓdtax	0.393	0.038	0.115
	(2.30)	(0.77)	(2.39)
sdtimp	0.630	−0.559	−0.297
	(0.86)	(1.94)	(0.98)
sdtexp	1.571	1.325	−0.631
	(1.20)	(3.33)	(1.75)
R^2:	0.926	0.979	0.976

Becker, Grossman, Murphy (1994): Table 6

TABLE 6—FUTURE CONSUMPTION COEFFICIENT (θ_f), PAST CONSUMPTION COEFFICIENT (θ_ℓ), AND RATIO OF LONG-RUN TO SHORT-RUN PRICE ELASTICITY, CORRECTED FOR FORECAST ERROR

k	θ_f	θ_ℓ	Ratio of long-run to short-run price elasticity
1.000	0.135	0.418	1.803
0.750	0.179	0.399	1.762
0.500	0.268	0.360	1.676
0.400	0.336	0.330	1.608
0.333	0.407	0.299	1.535

Notes: In the first column, k is the ratio of the partial covariance between expected future consumption and expected future price to the partial covariance between actual future consumption and actual future price, with current price, income, the three smuggling measures, and the state and time dummies held constant.

Becker, Grossman, Murphy (1994): Table 7

TABLE 7—CURRENT PRICE COEFFICIENTS, LAGGED CONSUMPTION COEFFICIENTS, LONG-RUN PRICE ELASTICITIES, AND SHORT-RUN PRICE ELASTICITIES IN RESTRICTED MODELS

		Panel A: *Future Price or Future Price and Future Tax Included as Instruments*					Panel B: *No Future Variables Included as Instruments*				
β	Model	Marginal significance level of restriction	P_t	C_{t-1}	Long-run price elasticity	Short-run price elasticity	Marginal significance level of restriction	P_t	C_{t-1}	Long-run price elasticity	Short-run price elasticity
0.70	(ii)	0.727	−1.220	0.360	−0.742	−0.445	0.000	−1.105	0.385	−0.755	−0.426
	(iv)	0.054	−0.925	0.426	−0.792	−0.395	0.000	−0.822	0.449	−0.820	−0.376
0.75	(ii)	0.548	−1.214	0.351	−0.743	−0.452	0.000	−1.084	0.378	−0.756	−0.430
	(iv)	0.021	−0.919	0.415	−0.792	−0.404	0.000	−0.803	0.440	−0.824	−0.384
0.80	(ii)	0.400	−1.208	0.342	−0.742	−0.458	0.000	−1.063	0.372	−0.759	−0.436
	(iv)	0.008	−0.913	0.404	−0.790	−0.413	0.000	−0.781	0.432	−0.829	−0.391
0.85	(ii)	0.285	−1.203	0.334	−0.743	−0.465	0.000	−1.044	0.366	−0.763	−0.442
	(iv)	0.003	−0.908	0.394	−0.791	−0.421	0.000	−0.761	0.424	−0.833	−0.398
0.90	(ii)	0.199	−1.199	0.326	−0.743	−0.472	0.000	−1.025	0.359	−0.761	−0.446
	(iv)	0.001	−0.904	0.385	−0.795	−0.431	0.000	−0.743	0.416	−0.837	−0.405
0.95	(ii)	0.136	−1.196	0.318	−0.743	−0.478	0.000	−1.007	0.353	−0.763	−0.451
	(iv)	0.000	−0.901	0.375	−0.791	−0.439	0.000	−0.725	0.409	−0.845	−0.414

Notes: All price and lagged consumption coefficients and all elasticities are statistically significant at all conventional levels of confidence. For panel A, the instruments in model (ii) are the one-period lag of price, the one-period lead of price, the current state excise tax, the one-period lag of the tax, and the exogenous variables in the demand function. Model (iv) adds the one-period lead of the tax and the two-period lags of the tax and price to the set of instruments. For panel B, the instruments in model (ii) are the one-period lag of price, the current tax, the one-period lag of the tax, and the exogenous variables in the demand function. Model (iv) adds the two-period lags of the tax and price to the set of instruments. The marginal significance levels of the restrictions are based on a Lagrange multiplier (LM) test.

Dynamic Panel: Arellano Bond

The main idea is that the instruments come from within the model!

$$y_{it} = \rho y_{i,t-1} + x'_{it}\beta + \eta_i + \varepsilon_{it}$$

Consider the first differences (s is a dummy time index):

$$E\left[x_{is}\left(\Delta y_{it} - \rho \Delta y_{i(t-1)} - \Delta x'_{it}\beta\right)\right] = 0$$

Idea:

- Now relax strict exogeneity.
- Can still use x_{is} as contemporaneous exogenous instrument.
- What is an excluded instrument for $\Delta y_{i,t-1}$?
    - Needs to be relevant
    - Still needs to be exogenous: not a direct determinant

Dynamic Panel: Arellano Bond

$$E\left[x_{is}\left(\Delta y_{it} - \rho \Delta y_{i(t-1)} - \Delta x'_{it}\beta\right)\right] = 0$$

Idea: Use higher lags of y_{it}:

- $[t = 2]$ or $[t = 1]$: no instruments,
- $[t = 3]$: valid instrument for $\Delta y_{i2} = (y_{i2} - y_{i1})$ is y_{i1}.
- $[t = 4]$: valid instruments for $\Delta y_{i3} = (y_{i3} - y_{i2})$ is (y_{i1}, y_{i2})
- $[t = 5]$: valid instruments for $\Delta y_{i4} = (y_{i4} - y_{i3})$ is (y_{i1}, y_{i2}, y_{i3}).
- $[t = T]$: valid instruments for $\Delta y_{iT-1} = (y_{i,T-1} - y_{i,T-2})$ is $(y_{i1}, \ldots, y_{i,T-2})$.

Thus there are $T/(T-1)/2$ instruments

Dynamic Panel: Arellano Bond

$$E\left[\mathbf{y}_{is}\left(\Delta y_{it} - \rho \Delta y_{i(t-1)} - \Delta x'_{it}\beta\right)\right] = 0$$
$$E\left[\Delta x_{it}\left(\Delta y_{it} - \rho \Delta y_{i(t-1)} - \Delta x'_{it}\beta\right)\right] = 0$$

- $\mathbf{y}_{is} = [y_{i1}, \ldots, y_{i,t-2}]$ for $t > 2$.
- Levels instrument Differences
- Thus there are $T/(T-1)/2$ instruments
- We can estimate with linear IV GMM: `pgmm` or `dynpanel`.
- The common complain is that instruments are still weak.

More Moments: Blundell and Bond

$$E\left[\mathbf{y}_{is}\left(\Delta y_{it} - \rho \Delta y_{i(t-1)} - \Delta x'_{it}\beta\right)\right] = 0$$
$$E\left[\Delta y_{i,t-1}\left(y_{it} - \rho y_{i(t-1)} - x'_{it}\beta\eta_i\right)\right] = 0$$
$$E\left[\Delta x_{it}\left(\Delta y_{it} - \rho \Delta y_{i(t-1)} - \Delta x'_{it}\beta\right)\right] = 0$$

- Differences also instrument Levels!
- Important when $\rho \to 1$ or when σ_u/σ_ϵ becomes large.
- These can also pin down y_{i0}, etc.
- This is known as `GMM-SYS`.

Recap: Dynamic Panel

- Arellano Bond estimators are quite popular for a number of reasons:
  - Easy to estimate (Linear IV)
  - Construct instruments out of the model itself.
- Blundell Bond estimators are somewhat less popular
  - System GMM isn't a single Linear IV problem anymore (requires full GMM).
  - Should be more efficient (more moment conditions)
  - The additional moments aren't any less reasonable that AB moments.
- But how much do we trust using lags and levels as IV for each other?

Thanks!