Lecture 5: Advanced Panel Data

Chris Conlon

NYU Stern

March 8, 2019

Recall the FE Assumptions

$$y_{it} = x_{it}'\beta + \eta_i + \varepsilon_{it}$$

- η_i is a fixed effect.
- To estimate everything consistently, we need $E[\varepsilon_{it}|x_{it}, \eta_i] = 0$
- Mostly this is not true. Instead usually treat η_i as a control variable or nuisance parameter.
  - A nuisance parameter is one that we estimate but don't care about interpreting.
  - If we only care about β then η_i is a nuisance parameter.
- With a control or nuisance parameter we only require that $E[\varepsilon_{it}|\eta_i] = E[\varepsilon_{it}|x_{it}, \eta_i]$ conditional mean independence.

Recall the FE Assumptions

$$y_{it} = x_{it}'\beta + \eta_i + \varepsilon_{it}$$

- η_i is a fixed effect.
- To estimate everything consistently, we need $E[\varepsilon_{it}|x_{it}, \eta_i] = 0$
- Mostly this is not true. Instead usually treat η_i as a control variable or nuisance parameter.
  - A nuisance parameter is one that we estimate but don't care about interpreting.
  - If we only care about β then η_i is a nuisance parameter.
- With a control or nuisance parameter we only require that $E[\varepsilon_{it}|\eta_i] = E[\varepsilon_{it}|x_{it}, \eta_i]$ conditional mean independence.
- Once we condition on η_i it is as if ε_{it} and x_{it} are uncorrelated.

Causal FE

- We can get away with conditional mean independence if we don't care about η_i.
- But suppose that we care about $\widehat{\eta_i}$?
  - Teacher FE
  - Physician/Hospital FE
  - Location/County FE
  - Suppose we take someone from the 10*th* percentile and move them to the 90*th* percentile

Causal FE

- Now we have to really believe $E[\varepsilon_{it}|x_{it}, \eta_i] = 0$
- We should worry about the conventional omitted variable bias problem.
- Suppose there exists a variable w_{it} so that:

$$y_{it} = x_{it}'\beta + w_{it}'\gamma + \eta_i + \varepsilon_{it}$$

- Recall the conditions for OVB
  - w_{it} is correlated with x_{it}
  - w_{it} is a determinant of y_{it}
- New one: w_{it} is correlated with η_i
  - This is easy to satisfy!
  - w_{it} needs to be uncorrelated with anything about the individual i.

Example: Test Scores

- Students s, Teachers t
- Want to measure effect of Teachers on Test Scores

$$TestScore_{st} = \beta x_s + \gamma w_t + \eta_t + \varepsilon_{st}$$

- We observe some features of students but not all of them (parent's education, household income, language spoken at home).
- We also observe some school specific variables w_t but not all of them (district spending per pupil, % free lunch, etc.).
- But we don't observe other things (jackhammering outside the classroom, which students have disruptive home lives,etc.).
  - If the mean of those things varies across teachers $\rightarrow$ we are screwed!
  - Can't get an accurate estimate of η_i.

Example: Test Scores

We need a better design:

- We probably need random assignment of students to teachers.
- Ideally we would be able to control for student and school unobservables.
- Might want to see many students match with many teachers.

Is quality (or productivity) correlated with marketshare?

$$\ln(N_h) = \beta_0^s + \beta_1^s q_h + \gamma_M^S + \varepsilon_h^s$$

- N_h measures market size for hospital h
- γ_M^s are market FE
- q_h is measure of hospital quality
- Goal: Is $\beta_1^s > 0$ or not. $\beta_1^s < 0$ is usually only Soviet countries or 1970's steel.
- $\beta_1^s > 0$ means allocation towards productive firms (or just returns to scale?)

$$\Delta_h = \beta_0^d + \beta_1^d q_h + \gamma_M^d + \varepsilon_h^d$$

$$\Delta_h = \frac{N_{h,2010} - N_{h,2008}}{\frac{1}{2}\left(N_{h,2010} + N_{h,2008}\right)}$$

- $\beta_1^d > 0$ means growth towards productive firms (not just returns to scale)
- Same idea but now we capture dynamics.
- Patients may still be attracted to unobservables correlated with quality.

Table 1—Static and Dynamic Allocation Metrics across Conditions

Condition:	AMI (1)	Heart failure (2)	Pneumonia (3)	Hip/knee (4)
Panel A. Composition of all Medicare discharges in 2008				
Number of patients in 2008	263,485	545,363	475,756	350,536
Share through emergency dept.	0.71	0.76	0.76	0.02
Share of all Medicare discharges	0.03	0.06	0.05	0.04
Share of Medicare hospital spending	0.04	0.05	0.04	0.05
Number of hospitals in 2008	4,257	4,547	4,607	3,297
Panel B. Static allocation: patients in 2008				
Patients (index events)	190,189	308,122	354,319	267,557
Average number of patients per hospital	65.8	76.6	81.9	101.7
SD of patients per hospital	67.6	78.2	70.8	118.0
Hospitals	2,890	4,023	4,325	2,632
Average number of hospitals per market	9.4	13.1	14.1	8.6
Panel C. Dynamic allocation: growth in patients from 2008 to 2010				
Average growth rate across hospitals	−0.17	−0.10	−0.13	−0.03
SD across hospitals	0.42	0.38	0.36	0.46
Hospitals	2,890	4,023	4,325	2,632

Notes: Panel A is calculated on a 100 percent sample of age 65 and older fee-for-service Medicare patients in 2008 and counts all patients with the condition, not just the index events that are the subject of the remainder of this study and panels B and C. The sample in panels B and C is all hospitals that had at least 1 index admission in 2008 for the condition shown in the column heading and had a valid risk-adjusted survival rate for that condition (risk-adjusted readmission for hip/knee replacement). There are 306 hospital markets, called Hospital Referral Regions (HRRs). ulated based on the formula in equation (3) that restricts values to between −2 and 2.

Table 2—Summary Statistics on Quality Metrics across Conditions

Condition:	AMI (1)	Heart failure (2)	Pneumonia (3)	Hip/knee (4)
Panel A. Risk-adjusted survival rates (30 days): patients in 2006–2008				
Average 30-day survival rate	0.82	0.89	0.88	
SD of risk-adjusted measure	(0.03)	(0.02)	(0.02)	
Hospitals in risk-adjusted measure	2,890	4,023	4,325	
Panel B. Risk-adjusted readmission rates (30 days): patients in 2006–2008				
Average 30-day readmission rate	0.21	0.21	0.16	0.06
SD of risk-adjusted measure	(0.03)	(0.02)	(0.02)	(0.02)
Hospitals in risk-adjusted measure	2,322	3,904	4,264	2,632
Panel C. Processes of care: shares of patients receiving appropriate treatments in 2006–2008				
Average score	0.93	0.83	0.88	
SD	(0.05)	(0.14)	(0.07)	
Hospitals	2,398	3,666	3,920	
Average number of processes reported	4.40	3.30	6.22	
Panel D. Patient survey: survey covers all patients in 2008 (not limited to particular condition)				
Average overall rating (1–3, higher is better)	2.53	2.53	2.53	2.53
SD	(0.14)	(0.14)	(0.14)	(0.14)
Hospitals	3,498	3,598	3,610	3,061

Notes: Sample restrictions are specific to the condition and quality metric; see text for more details of the metric definitions and sample restrictions. Summary statistics are reported across hospitals. In panels A and B, the standard deviations are of the risk-adjusted measures and are empirical-Bayes-adjusted to account for measurement error (see online Appendix Section C.3.1). In panel D, the number of hospitals differs across conditions even though
etric is not condition-specific because we calculate the ratings on the subset of hospitals that
patient with the condition in 2008.

Table 3—Correlation of Quality Metrics within Condition

Metric	AMI Risk-adj survival (1)	AMI Risk-adj readm (2)	AMI Process of care Z (3)	AMI Patient survey Z (4)	HF Risk-adj survival (5)	HF Risk-adj readm (6)	HF Process of care Z (7)	HF Patient survey Z (8)
Risk-adjusted survival	1.00 [2,890]				1.00 [4,023]			
Risk-adjusted readmission	0.03 [2,322]	1.00 [2,322]			0.35 [3,904]	1.00 [3,904]		
Process of care Z-score	0.24 [2,346]	−0.25 [2,214]	1.00 [2,398]		0.17 [3,607]	−0.15 [3,578]	1.00 [3,666]	
Patient survey Z-score	−0.06 [2,799]	−0.26 [2,293]	0.18 [2,370]	1.00 [3,498]	−0.18 [3,447]	−0.36 [3,398]	0.01 [3,392]	1.00 [3,598]

Metric	Pneumonia Risk-adj survival	Pneumonia Risk-adj readm	Pneumonia Process of care Z	Pneumonia Patient survey Z	Hip/knee replacement Risk-adj survival	Hip/knee replacement Risk-adj readm	Hip/knee replacement Process of care Z	Hip/knee replacement Patient survey Z
Risk-adjusted survival	1.00 [4,325]							
Risk-adjusted readmission	0.08 [4,264]	1.00 [4,264]				1.00 [2,632]		
Process of care Z-score	0.08 [3,871]	−0.18 [3,847]	1.00 [3,920]					
Patient survey Z-score	−0.03 [3,527]	−0.36 [3,503]	0.18 [3,512]	1.00 [3,610]		−0.23 [2,542]		1.00 [3,061]

to calculate correlation in brackets. All quality metrics are condition-specific except the
s only available as an all-patient average. Correlations involving risk-adjusted survival and

Healthcare Exceptionalism

Table 4—Allocation across Conditions

Measure/condition	Static allocation AMI (1)	HF (2)	Pneu (3)	Hip/knee (4)	Dynamic allocation AMI (5)	HF (6)	Pneu (7)	Hip/knee (8)
Risk-adjusted survival								
Coef. on survival rate	17.496	15.360	5.140		1.533	0.774	1.220	
	(0.995)	(1.320)	(0.777)		(0.379)	(0.501)	(0.354)	
Hospitals	2,890	4,023	4,325		2,890	4,023	4,325	
Risk-adjusted readmission								
Coef. on readmission rate	−9.162	−10.346	0.499	−21.037	−1.428	−2.300	−1.138	−1.112
	(1.621)	(1.782)	(1.575)	(2.027)	(0.611)	(0.651)	(0.679)	(0.836)
Hospitals	2,322	3,904	4,264	2,632	2,322	3,904	4,264	2,632
Process of care Z-score								
Coef. on process Z-score	0.319	0.332	0.211		0.048	0.043	0.026	
	(0.026)	(0.016)	(0.015)		(0.010)	(0.009)	(0.009)	
Hospitals	2,398	3,666	3,920		2,398	3,666	3,920	
Patient survey Z-score								
Coef. on survey Z-score	−0.321	−0.252	−0.210	0.057	−0.065	−0.003	0.007	0.037
	(0.052)	(0.038)	(0.030)	(0.051)	(0.015)	(0.011)	(0.011)	(0.022)
Hospitals	3,498	3,598	3,610	3,061	3,498	3,598	3,610	3,061

Notes: The static allocation results are estimated using equation (1), a hospital-level regression of log-patients in 2008 on market fixed effects and the quality measure named in the row. The dynamic allocation results are estimated using equation (2), which is an identical regression except for the dependent variable, which is now growth in patients from 2008 to 2010. Growth is defined as in equation (3). Standard errors are bootstrapped with 300 replications and are clustered at the market level. Risk-adjusted survival and readmission are reported in percentage … 10 percentage points); process of care and patient survey metrics are reported in stan- … value of 1 is one standard deviation).

Healthcare Exceptionalism: Production Function

Hospital Production Function:

$$y_p^s = a_h + \sum_k \lambda_k r_{pk} + \mu x_p + \xi_p$$

- a_h is hospital productivity (a FE) and variable of interest
- y_p is a patient outcome (survival-days, etc.)
- x_p are (log) hospital inputs
- r_{pk} are patient risk factors.
- This has interpretation as a production function. Why?

TABLE 7—ALLOCATION OF AMI WITH RESPECT TO AMI PRODUCTIVITY AND ITS COMPONENTS

Measure	Static allocation for AMI				Dynamic allocation for AMI			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Productivity (fed $)	17.637				1.491			
	(1.118)				(0.420)			
Productivity (resources)		17.540				1.471		
		(1.013)				(0.386)		
Risk-adjusted ln(fed $)			1.447				0.246	
			(0.169)				(0.064)	
Risk-adjusted ln(resources)				0.620				0.468
				(0.406)				(0.162)
Risk-adjusted survival			17.940	19.789			1.479	1.559
			(1.192)	(1.297)			(0.446)	(0.441)
Hospitals	2,890	2,890	2,890	2,890	2,890	2,890	2,890	2,890

Notes: This table extends the analysis of Table 4 but is limited to static and dynamic allocation for AMI. It shows how allocation is related to AMI productivity or its two components (risk-adjusted survival and risk-adjusted log inputs). Productivity is defined as risk- and inputs-adjusted survival; see Section IIIC and equation (9). We consider two input measures, "federal expenditures" and "resources," also defined in the text. Standard errors are bootstrapped with 300 replications and are clustered at the market level. The standard deviation of productivity is 0.03 (Fed $ or Resources), of risk-adjusted log-inputs is 0.22 (Fed $) and 0.07 (Resources), and of risk-adjusted survival is 0.04—this number differs from that of Table 2 because it comes from estimating the joint distribution of survival and inputs, not survival alone.

Healthcare Exceptionalism: EB Adjustment

Table A2 - Sensitivity of Allocation Results to Empirical Bayes Adjustment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Static Allocation				Dynamic Allocation			
Condition	AMI	HF	Pneu	Hip/Knee	AMI	HF	Pneu	Hip/Knee
Panel A - Risk-Adjusted Survival								
Baseline (EB-Adjusted)	17.496	15.360	5.140		1.533	0.774	1.220	
	(0.995)	(1.320)	(0.777)		(0.379)	(0.501)	(0.354)	
Raw (No EB Adjustment)	6.833	3.761	1.957		0.645	0.084	0.340	
	(0.342)	(0.425)	(0.403)		(0.175)	(0.199)	(0.175)	
Hospitals	2,890	4,023	4,325		2,890	4,023	4,325	
Raw SD / Corrected SD	1.597	1.788	1.547		1.597	1.788	1.547	
Panel B - Risk-Adjusted Readmission								
Baseline (EB-Adjusted)	-9.162	-10.346	0.499	-21.037	-1.428	-2.300	-1.138	-1.112
	(1.621)	(1.782)	(1.575)	(2.027)	(0.611)	(0.651)	(0.679)	(0.836)
Raw (No EB Adjustment)	-1.699	-1.043	0.755	-6.492	-0.307	-0.217	-0.189	-0.431
	(0.395)	(0.346)	(0.427)	(0.727)	(0.197)	(0.162)	(0.195)	(0.382)
Hospitals	2,322	3,904	4,264	2,632	2,322	3,904	4,264	2,632
Raw SD / Corrected SD	1.870	2.132	1.864	1.794	1.870	2.132	1.864	1.794

This table shows the sensitivity of the allocation results of Table 4 to the empirical Bayes adjustment procedure. In each panel, we first repeat the baseline allocation results in which the quality metric is empirical-Bayes-adjusted. We then show the same allocation models using the raw quality metric without empirical Bayes adjustment. Lastly, we show the ratio of the raw standard deviation of the quality measure to its standard deviation after correcting for). Standard errors are bootstrapped with 300 replications and are

- 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

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

- This eliminates the fixed effect.
- But $Cov(y_{i,t-1} - \bar{y}_i, \varepsilon_{it} - \bar{\varepsilon}_i) \neq 0$. 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 (x_{i1}, x_{i2}, x_{i3}).
- The model is just identified with 3 parameters $(\alpha, \beta_0, \beta_1)$.
- The challenge with this approach is often that it suffers from weak instruments.

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	0.502	0.602	0.755
	(12.07)	(14.68)	(21.43)	(64.84)
P_t	−1.603	−1.538	−1.269	−0.860
	(10.12)	(10.48)	(9.74)	(8.33)
Y_t	0.942	0.903	0.741	0.493
	(7.61)	(7.71)	(6.96)	(5.44)
ℓdtax	−0.240	−0.233	0.212	−0.160
	(7.33)	(7.40)	(7.22)	(6.17)
sdtimp	−1.541	−1.514	−1.372	−1.228
	(5.04)	(5.09)	(4.97)	(4.84)
sdtexp	−3.659	−3.544	−3.059	−2.328
	(13.24)	(13.88)	(13.71)	(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

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)
ℓ dtax	−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

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)

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)

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

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

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_{i2}, y_{i1})
- $[t = 5]$: valid instruments for $\Delta y_{i5} = (y_{i5} - y_{i4})$ is $(y_{i1}, \ldots, y_{i4})$.
- $[t = T]$: valid instruments for $\Delta y_{iT} = (y_{iT} - y_{i,T-1})$ is $(y_{i1}, \ldots, y_{i,T-1})$.

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.

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

- 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.