



Advanced Econometrics

Take-home exam

Gerdie Everaert

2023-2024

Main theme: dynamic panel data models

0. To warm up: OLS estimation
1. Fixed Effects (FE) estimation
 - ▶ Implement the FE estimator in R/Matlab
 - ▶ Simulate its properties in a dynamic setting
 - ▶ Use simulation methods for bias correction and inference
2. Generalized Method of Moments (GMM) estimation
 - ▶ Implement the GMM function in R/Matlab
 - ▶ Simulate its properties in a dynamic setting

To warm up: OLS estimation

- ▶ Sample R and Matlab **functions** implementing the basic OLS estimator are available on Ufora
 - ▶ inputs: y and X
 - ▶ output: OLS coefs, se, t -stats, p -values
- ▶ Replicate the OLS results (see next slide) for the following cigarette demand model (Baltagi, 2008)

$$\ln C_{it} = \alpha + \rho \ln C_{i,t-1} + \beta_1 \ln P_{it} + \beta_2 \ln Pn_{it} + \beta_3 \ln Y_{it} + \mu_{it}$$

with C_{it} the number of packs of cigarettes sold per person in state i in year t , P_{it} the price per pack, Y_{it} disposable income and Pn_{it} the minimum price found in neighboring states

- ▶ data for 46 states over the period 1963-92 are available on Ufora ('Data_Baltagi.xlsx')

└ To warm up

Table 8.1 Pooled Estimation Results.* Cigarette Demand Equation 1963–92

	$\ln C_{i,t-1}$	$\ln P_{it}$	$\ln Pn_{it}$	$\ln Y_{it}$
OLS	0.97 (157.7)	-0.090 (6.2)	0.024 (1.8)	-0.03 (5.1)
Within	0.83 (66.3)	-0.299 (12.7)	0.034 (1.2)	0.10 (4.2)
2SLS	0.85 (25.3)	-0.205 (5.8)	0.052 (3.1)	-0.02 (2.2)
2SLS-KR	0.71 (22.7)	-0.311 (13.9)	0.071 (3.7)	-0.02 (1.5)
Within-2SLS	0.60 (17.0)	-0.496 (13.0)	-0.016 (0.5)	0.19 (6.4)
FD-2SLS	0.51 (9.5)	-0.348 (12.3)	0.112 (3.5)	0.10 (2.9)
FD-2SLS-KR	0.49 (18.7)	-0.348 (18.0)	0.095 (4.7)	0.13 (9.0)
GMM-two-step	0.70 (10.2)	-0.396 (6.0)	-0.105 (1.3)	0.13 (3.5)
System GMM	0.70 (8.8)	-0.415 (4.3)	-0.003 (0.1)	0.09 (3.4)

* Numbers in parentheses are t -statistics. All regressions except OLS and 2SLS include time dummies.

Source: Some of the results in this table are reported in Baltagi, Griffin and Xiong (2000).

Implementing the FE estimator

- ▶ Write a Matlab/R function that implements the FE estimator
 - ▶ inputs: y and X
 - ▶ output: FE coefs, se, t -stats (p -values)
- ▶ Note that it may be (computationally) convenient to
 - ▶ set up y as a $(T \times N)$ matrix and X as a $(T \times N \times K)$ matrix
 - ▶ loop over the N -dimension to compute

$$X' M_D X = \sum_{i=1}^N X_i' M_d X_i, \quad X' M_D y = \sum_{i=1}^N X_i' M_d y_i$$

with y_i and X_i the data for cross-sectional unit i and
 $M_d = I_T - T^{-1} \iota_T \iota_T'$

- ▶ Replicate the FE results reported by Baltagi
 (you can add time effects using dummies)

Simulating the properties of the FE estimator

Set up a Monte Carlo experiment to numerically illustrate the properties of the FE estimator in a dynamic setting

- ▶ Generate data from $y_{it} = \rho y_{i,t-1} + \varepsilon_{it}$ with $\varepsilon_{it} \sim N(0, 1)$
 - ▶ use a burn-in of 25 periods (i.e., initialize $y_{i,-25} = 0$, generate y_{it} for $t = -24, \dots, 0, 1, \dots, T$ and discard the first 25 draws)
 - ▶ by using $y_{i,0}, \dots, y_{i,T}$ and after taking a lag, this implies that you have T effective observations
 - ▶ use 3 different values of $\rho = 0, 0.5, 0.9$
 - ▶ use all combinations of $T = 4, 10, 20, 50$ and $N = 10, 100$
- ▶ Evaluate the properties of the FE estimator $\hat{\rho}_{FE}$ in the model $y_{it} = \alpha_i + \rho y_{i,t-1} + \varepsilon_{it}$
 - ▶ compare the mean of $\hat{\rho}_{FE}$ to the population parameter ρ
 - ▶ compare the mean of the estimated standard error $\widehat{se}(\hat{\rho}_{FE})$ to the true standard error $se(\hat{\rho}_{FE})$
 - ▶ compute the size of a t -test for ρ based on the analytical se

Use simulation methods for bias correction and inference

- ▶ Implement a half-panel jackknife bias-corrected FE estimator $\hat{\rho}_{FEj}$
 - ▶ use the above Monte Carlo setup to compare the mean of $\hat{\rho}_{FEj}$ to the population parameter ρ
- ▶ Implement a bootstrap to compute standard errors for the jackknife FE estimator
 - ▶ think about an appropriate resampling scheme
 - ▶ compare the mean of the bootstrapped standard error $\hat{se}^b(\hat{\rho}_{FEj})$ to the true standard error $se(\hat{\rho}_{FEj})$
 - ▶ compute the size of a t -test for ρ based on the bootstrapped standard errors

Implementing the panel GMM estimator

Write a Matlab/R function that implements the GMM estimator

- ▶ The inputs include
 - ▶ y a $(T \times N)$ vector including the dependent variable
 - ▶ X a $(T \times N \times K)$ vector including K explanatory variables (include constant and dummies if desired)
 - ▶ Z a $(T \times N \times R)$ vector including $R \geq K$ instruments (including exogenous explanatory variables)
- ▶ The results should include
 - ▶ One-step GMM estimates, stde, t -stats and p -values
 - ▶ For overidentified models
 - ▶ Two-step GMM estimates, stde, t -stats and p -values
 - ▶ Sargan/Hansen test
 - ▶ Tailor the code to first-difference panel setting

Simulating the properties of the GMM estimator

Use the above Monte Carlo experiment to numerically illustrate the properties of the GMM estimator in a dynamic setting

- ▶ Take first-differences to obtain: $\Delta y_{it} = \rho \Delta y_{i,t-1} + \Delta \varepsilon_{it}$
- ▶ Set up two different GMM estimators by using as instrument(s) for $\Delta y_{i,t-1}$:
 - ▶ $y_{i,t-2}$
 - ▶ $y_{i,t-2}$ and $y_{i,t-3}$
- ▶ Theoretically argue whether/when these instruments are
 - ▶ strong or weak
 - ▶ valid

Simulating the properties of the GMM estimator

- ▶ Evaluate the properties of the one and two-step GMM estimator
 - ▶ compare the mean of $\hat{\rho}_{GMM}$ to the population parameter ρ
 - ▶ compare the true standard deviation to that of the FE estimator
 - ▶ compare the mean of the estimated standard error $\widehat{se}(\hat{\rho}_{GMM})$ to the true standard error $se(\hat{\rho}_{GMM})$
 - ▶ compute the size of a t -test for ρ

Practical arrangements

- ▶ Deadline for uploading your files on Ufora: Wednesday June 26 (exam is on June 28)
- ▶ These files should contain your R/Matlab code + a PDF file with a summary of the results (and answers to the questions in the case)