Review Questions: Advanced Panel Data Models
Handout 2 – Master in Data Science for Decision Making

Barcelona School of Economics

Winter 2026

Instructions

Answer all questions. For conceptual questions, provide clear explanations demonstrating your understanding. For mathematical questions, show all relevant steps. These questions are designed to help you assess whether you truly understand the material—if you find yourself uncertain, revisit the corresponding sections of the handout.

1 Dynamic Panel Data Models

Question 1. Defining Dynamic Models Consider the following two panel data specifications:

$$\text{Model A:} \quad y_{it} = c_i + \beta_1 X_{it} + \beta_2 X_{i,t-1} + v_{it} \tag{1}$$

$$\text{Model B:} \quad y_{it} = c_i + \beta X_{it} + \rho y_{i,t-1} + v_{it} \tag{2}$$

(a) Which model is a dynamic panel data model? Explain why.

(b) What is the key characteristic that makes a panel data model "dynamic"?

(c) Give an economic example where Model B would be more appropriate than Model A.

Question 2. Strict Exogeneity Failure In the dynamic panel data model:

$$y_{it} = c_i + X_{it}\beta + \rho y_{i,t-1} + v_{it}, \quad |\rho| < 1$$

where $E(v_{it}|y_{i,t-1}, X_{it}) = 0$ and v_{it} is serially uncorrelated.

(a) Explain why $y_{i,t-1}$ is **not** strictly exogenous.

(b) Write out the strict exogeneity condition and show explicitly which part of it fails.

(c) Why does this violation matter for the estimators we learned in static panel models?

Question 3. Random Effects Inconsistency Consider a dynamic panel with random effects, where c_i is assumed uncorrelated with X_{it}.

(a) Even under the random effects assumption, explain why $y_{i,t-1}$ is correlated with c_i.

(b) What is the composite error term in the RE model, and why does the correlation in part (a) cause problems?

(c) Is OLS on the pooled data consistent in this setting? Why or why not?

Question 4. The Nickell Bias – Conceptual Understanding The "Nickell bias" is a fundamental problem in dynamic panel data estimation with fixed effects.

(a) Explain intuitively why the within transformation $\tilde{y}_{i,t-1} = y_{i,t-1} - \bar{y}_i$ creates endogeneity.

(b) The Nickell bias is of order $O(T^{-1})$. What does this mean in practical terms?

(c) If you have a panel with $N = 1000$ and $T = 5$, would you be concerned about Nickell bias? What if $T = 50$?

(d) Does the Nickell bias only affect the estimate of ρ, or does it also affect the estimate of β? Explain.

Question 5. Nickell Bias – Mathematical Derivation Consider the simple AR(1) panel model without covariates:
$$y_{it} = c_i + \rho y_{i,t-1} + v_{it}$$

After the within transformation:
$$\tilde{y}_{it} = \rho \tilde{y}_{i,t-1} + \tilde{v}_{it}$$

where $\tilde{y}_{it} = y_{it} - \bar{y}_i$ and $\tilde{v}_{it} = v_{it} - \bar{v}_i$.

(a) Show that $\bar{y}_i = \frac{1}{T} \sum_{s=1}^{T} y_{is}$ contains v_{it}.

(b) Use this to explain why $\text{Cov}(\tilde{y}_{i,t-1}, \tilde{v}_{it}) \neq 0$.

(c) What sign do you expect for this covariance? (Hint: think about whether $\tilde{y}_{i,t-1}$ and $\tilde{v}_{it}$ move in the same or opposite directions.)

Question 6. Anderson-Hsiao Estimator The Anderson-Hsiao (AH) estimator proposes using $y_{i,t-2}$ as an instrument for the first-differenced model:

$$\Delta y_{it} = \rho \Delta y_{i,t-1} + \Delta X_{it}\beta + \Delta v_{it}$$

(a) What are the two conditions required for a valid instrumental variable?

(b) Verify that $y_{i,t-2}$ satisfies the exogeneity condition. Under what assumption about v_{it}?

(c) Verify that $y_{i,t-2}$ satisfies the relevance condition.

(d) Why is the AH estimator only fully efficient when $T = 2$?

Question 7. Arellano-Bond Difference GMM The Arellano-Bond estimator extends the Anderson-Hsiao approach.

(a) What is the key innovation of Arellano-Bond compared to Anderson-Hsiao?

(b) For the equation at time $t = 5$, list all valid instruments from the lagged dependent variable.

(c) Why does using more instruments improve efficiency?

(d) The model is said to be "overidentified." What does this mean, and what test can you use to check the validity of the overidentifying restrictions?

Question 8. System GMM Blundell and Bond (1998) propose System GMM as an improvement over Difference GMM.

(a) Under what circumstances does Difference GMM perform poorly?

(b) Explain intuitively why $y_{i,t-2}$ becomes a weak instrument when ρ is close to 1.

(c) In practice, how would you decide between Difference GMM and System GMM?

Question 9. Applied Question: Democracy and Growth Acemoglu et al. (2019) estimate:

$$y_{ct} = \beta D_{ct} + \sum_{j=1}^{p} \gamma_j y_{c,t-j} + \alpha_c + \delta_t + \varepsilon_{ct}$$

where y_{ct} is log GDP per capita, D_{ct} is a democracy indicator, α_c are country fixed effects, and δ_t are year fixed effects.

(a) Why do the authors include country fixed effects α_c?

(b) Why do they include lagged GDP $(y_{c,t-j})$ in the model?

(c) The paper uses both within estimates and Arellano-Bond estimates. If $N = 175$ and $T = 50$, would you expect Nickell bias to be large or small? Explain.

(d) What is the interpretation of β in this model?

2 Nonlinear Panel Data Models

Question 10. Linear vs. Nonlinear Models

(a) In linear panel models, c_i enters additively: $E(y_{it}|c_i, X_{it}) = c_i + X_{it}\beta$. Why does this additive structure make estimation easier?

(b) Why might we prefer a nonlinear model (like logit) over a linear probability model when y_{it} is binary?

Question 11. The Incidental Parameters Problem – Intuition

(a) What are "incidental parameters" in a panel data context?

(b) Why can't the individual effects $c_1, \ldots, c_N$ be consistently estimated when T is fixed and $N \to \infty$?

(c) In linear models with fixed effects, why doesn't the inconsistent estimation of c_i contaminate the estimation of β? (Hint: think about the Frisch-Waugh-Lovell theorem.)

(d) Why does this "protection" fail in nonlinear models?

Question 12. Incidental Parameters Problem – Simple Example Consider the model $y_{it} \sim N(\alpha_i, \sigma^2)$ where α_i are individual-specific means.

(a) Write down the MLE for α_i and σ^2.

(b) Show that as $N \to \infty$ with T fixed, $\hat{\sigma}^2 \to_p \sigma^2 \frac{T-1}{T}$.

(c) What is the probability limit of $\hat{\sigma}^2$ when $T = 2$?

(d) Propose a simple correction to obtain a consistent estimator of σ^2.

Question 13. Sufficient Statistics and Conditional Likelihood

(a) Define what it means for a statistic τ to be "sufficient" for a parameter θ.

(b) Explain the key idea behind the conditional likelihood approach to solving the incidental parameters problem.

(c) Why doesn't this approach work for all nonlinear models?

(d) Name two models for which a sufficient statistic exists and two for which it doesn't.

Question 14. Fixed Effects Logit Model Consider the fixed effects logit model:

$$\Pr(y_{it} = 1 | X_{it}, c_i) = \Lambda(c_i + X_{it}\beta) = \frac{\exp(c_i + X_{it}\beta)}{1 + \exp(c_i + X_{it}\beta)}$$

(a) Write down the joint density $f(y_i | X_i, c_i, \beta)$ assuming independence across t.

(b) What is the sufficient statistic for c_i in this model?

(c) Why does conditioning on $\sum_{t=1}^{T} y_{it}$ eliminate the incidental parameter c_i?

(d) Can the same approach be used for the fixed effects probit model? Why or why not?

Question 15. Conditional Logit – Sample Selection When estimating a conditional logit model in Stata using `xtlogit, fe`, you might see the message: "324 groups dropped because of all positive or all negative outcomes."

(a) Why are individuals with $y_{it} = 0$ for all t (or $y_{it} = 1$ for all t) dropped from the analysis?

(b) What does this imply about the type of variation the conditional logit exploits?

(c) Could this sample selection introduce bias in certain applications? Give an example.

(d) If 30% of your sample is dropped due to this reason, would you be concerned? Why?

Question 16. Interpreting Logit Coefficients In the NLSY poverty example, the conditional logit estimates show that the coefficient on "mother" (having a child) is 0.582, corresponding to an odds ratio of 1.79.

(a) Interpret the odds ratio of 1.79 in plain language.

(b) The odds ratio for "spouse" is 0.47. What does this tell us about the relationship between marriage and poverty?

(c) Why can't we directly interpret logit coefficients as marginal effects on probabilities?

(d) Can we estimate the fixed effects c_i in the conditional logit model? Why might this be a limitation?

3 Synthesis Questions

Question 17. Comparing Estimation Challenges Fill in the following table comparing estimation challenges:

Estimator	Dynamic Panel	Nonlinear Panel (FE)
OLS/Pooled		
Random Effects		
Fixed Effects (Within)		
First Differences		

For each cell, write "Consistent," "Inconsistent," or "Depends" and briefly explain why.

Question 18. Choosing an Estimator For each scenario below, recommend an appropriate estimation strategy and justify your choice:

(a) You have $N = 500$ firms observed over $T = 4$ years. The dependent variable is continuous (investment), and you want to include lagged investment. You suspect firm fixed effects are correlated with the regressors.

(b) You have $N = 1000$ individuals observed over $T = 10$ periods. The dependent variable is binary (employment status), and you suspect individual fixed effects.

(c) You have $N = 100$ countries observed over $T = 50$ years. The dependent variable is continuous (GDP growth), and you include lagged GDP growth. The autoregressive coefficient is expected to be high ($\rho \approx 0.9$).

(d) You have $N = 2000$ individuals observed over $T = 3$ periods. The dependent variable is binary (health insurance coverage), and you want to allow for individual fixed effects.

Question 19. True or False For each statement, indicate whether it is True or False, and provide a brief explanation.

(a) In dynamic panels, the random effects estimator is consistent if c_i is uncorrelated with the regressors X_{it}.

(b) The Nickell bias disappears as $N \to \infty$ with T fixed.

(c) First differencing solves the endogeneity problem in dynamic panels.

(d) The Anderson-Hsiao estimator uses $y_{i,t-2}$ as an instrument because it is correlated with $\Delta y_{i,t-1}$ but uncorrelated with Δv_{it}.

(e) System GMM is always preferred to Difference GMM.

(f) The incidental parameters problem affects both linear and nonlinear panel models equally.

(g) In the conditional logit model, we can estimate the marginal effect of X on the probability $\Pr(y_{it} = 1)$.

(h) A probit model with fixed effects can be consistently estimated using conditional likelihood.

Question 20. Critical Thinking A researcher estimates a dynamic panel model using within (FE) estimation with $N = 200$ and $T = 8$. They find $\hat{\rho} = 0.3$ and conclude that persistence is "moderate."

(a) What concern would you raise about this estimate?

(b) The Nickell bias is typically negative. What does this suggest about the true value of ρ?

(c) What alternative estimator would you recommend, and why?

(d) If the researcher re-estimates using Arellano-Bond and finds $\hat{\rho}_{AB} = 0.5$, would this be consistent with your concerns? Explain.