Advanced Econometrics & Dynamic Theory

Problem Set 3: Causal Inference & Dynamic Optimization

Instructor: Prof. Aris Thorne | Department of Economics

Course Code:	ECON 4120 / 6120
Date Issued:	March 14, 2026
Submission Due:	March 28, 2026 (23:59 EST)
Target Audience:	Senior Undergrad / Graduate

Student Name: ______________________ **Student ID / Matriculation:** ______________________

Instructions: Show all derivation steps for analytical questions. Attach executable R scripts for empirical sections.

Score / Total: / 100

1 Part I: Intertemporal Choice & Liquidity Constraints

Problem 1: Household Consumption Smoothing under Borrowing Limits (25 Points)

Consider a two-period consumption problem. A household receives endowment $y_1 > 0$ in Period 1 and $y_2 > 0$ in Period 2, maximizing CRRA utility $U(c_1, c_2) = \frac{c_1^{1-\gamma}-1}{1-\gamma} + \beta\left(\frac{c_2^{1-\gamma}-1}{1-\gamma}\right)$ with $\gamma > 0, \gamma \neq 1, \beta \in (0,1)$, subject to interest rate $r > 0$ and borrowing limit $c_1 \leq y_1$.

(a) Formulate the constrained optimization problem using the Lagrangian method and derive the modified Euler equation.

Solution: Writing the Lagrangian with borrowing constraint multiplier $\lambda \geq 0$:

$$\mathcal{L} = \frac{c_1^{1-\gamma}-1}{1-\gamma} + \beta\left(\frac{c_2^{1-\gamma}-1}{1-\gamma}\right) + \mu\left[y_1 + \frac{y_2}{1+r} - c_1 - \frac{c_2}{1+r}\right] + \lambda(y_1 - c_1)$$

First-order conditions $\frac{\partial \mathcal{L}}{\partial c_1} = c_1^{-\gamma} - \mu - \lambda = 0$ and $\frac{\partial \mathcal{L}}{\partial c_2} = \beta c_2^{-\gamma} - \frac{\mu}{1+r} = 0$ yield the modified Euler Equation:

$$c_1^{-\gamma} = \beta(1+r)c_2^{-\gamma} + \lambda \tag{1}$$

Complementary slackness requires $\lambda \geq 0$, $(y_1 - c_1) \geq 0$, and $\lambda(y_1 - c_1) = 0$. When the borrowing limit binds ($\lambda > 0$), $c_1^* = y_1$, $c_2^* = y_2$, and marginal utility satisfies $c_1^{-\gamma} > \beta(1+r)c_2^{-\gamma}$.

(b) Graphically illustrate optimal allocation (c_1^*, c_2^*) under (i) Unconstrained regime ($\lambda = 0$) and (ii) Binding constraint ($\lambda > 0$).

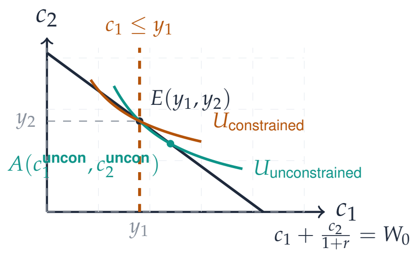


Figure 1: Intertemporal Consumption Choice under Borrowing Constraints. Low y_1 causes unconstrained point A to violate $c_1 \leq y_1$, forcing consumption to endowment point E.

2 Part II: Endogeneity, Instrumental Variables & 2SLS

Problem 2: Estimating Skill Returns at Nexa Analytics (35 Points)

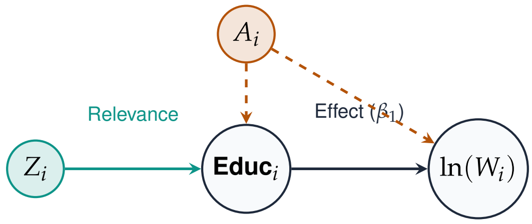
An econometrician at **Nexa Analytics** evaluates technical training (Educ_i) on log hourly earnings ($\ln(W_i)$):

$$\ln(W_i) = \beta_0 + \beta_1 \text{Educ}_i + \beta_2 \text{Exper}_i + u_i \tag{2}$$

Unobserved ability (A_i) induces endogeneity: $u_i = \gamma A_i + v_i$ with $\text{Cov}(\text{Educ}_i, u_i) > 0$. The instrument Z_i is childhood distance to public university.

(a) Draw the Causal Directed Acyclic Graph (DAG) for Z_i, Educ_i, $\ln(W_i)$, and A_i. State key identification conditions.

Identification Analysis & Causal DAG:



Identification requirements:

1. **Relevance:** $\text{Cov}(Z_i, \text{Educ}_i \mid \text{Exper}_i) \neq 0$ (Verified via First-Stage $F > 10$).
2. **Exclusion Restriction & Exogeneity:** $\text{Cov}(Z_i, u_i) = 0$ (Z_i affects wages only through education).

(b) Algebraically prove that the 2SLS estimator $\hat{\boldsymbol{\beta}}_{2SLS}$ is consistent for $\boldsymbol{\beta}$ in matrix notation.

Proof of 2SLS Consistency:

Let projection matrix $\mathbf{P}_Z = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$. The 2SLS estimator $\hat{\boldsymbol{\beta}}_{2SLS} = (\mathbf{X}'\mathbf{P}_Z\mathbf{X})^{-1}\mathbf{X}'\mathbf{P}_Z\mathbf{Y}$ expands to:

$$\hat{\boldsymbol{\beta}}_{2SLS} - \boldsymbol{\beta} = \left[\left(\frac{\mathbf{X}'\mathbf{Z}}{N}\right)\left(\frac{\mathbf{Z}'\mathbf{Z}}{N}\right)^{-1}\left(\frac{\mathbf{Z}'\mathbf{X}}{N}\right)\right]^{-1}\left(\frac{\mathbf{X}'\mathbf{Z}}{N}\right)\left(\frac{\mathbf{Z}'\mathbf{Z}}{N}\right)^{-1}\left(\frac{\mathbf{Z}'\mathbf{u}}{N}\right) \tag{3}$$

As $N \to \infty$, $\frac{1}{N}\mathbf{Z}'\mathbf{Z} \xrightarrow{p} \boldsymbol{\Sigma}_{ZZ}$ and $\frac{1}{N}\mathbf{Z}'\mathbf{X} \xrightarrow{p} \boldsymbol{\Sigma}_{ZX}$. By Exogeneity, $\frac{1}{N}\mathbf{Z}'\mathbf{u} \xrightarrow{p} \mathbf{0}$. By Slutsky's Theorem:

$$\text{plim}_{N\to\infty}\left(\hat{\boldsymbol{\beta}}_{2SLS} - \boldsymbol{\beta}\right) = \left[\boldsymbol{\Sigma}_{XZ}\boldsymbol{\Sigma}_{ZZ}^{-1}\boldsymbol{\Sigma}_{ZX}\right]^{-1}\boldsymbol{\Sigma}_{XZ}\boldsymbol{\Sigma}_{ZZ}^{-1} \cdot \mathbf{0} = \mathbf{0} \implies \hat{\boldsymbol{\beta}}_{2SLS} \xrightarrow{p} \boldsymbol{\beta} \quad \square$$

(c) Explain the weak instrument problem. What statistic evaluates instrument strength, and what threshold is used?

Weak Instruments & First-Stage Diagnostics:

Weak instruments ($\text{Cov}(Z, X) \approx 0$) bias 2SLS toward OLS and invalidate standard test inference. First-Stage effective F-statistic evaluates strength: $F = \frac{(\mathbf{R}\hat{\pi})'[\mathbf{R}\widehat{\text{Var}}(\hat{\pi})\mathbf{R}']^{-1}(\mathbf{R}\hat{\pi})}{q}$. Stock-Yogo rule of thumb requires $F > 10$ to ensure 2SLS bias is under 10% of OLS bias.

3 Part III: Panel Econometrics & Corporate R&D at Vertex Global

Problem 3: Panel Data Estimation of Productivity Dynamics (40 Points)

Using a balanced panel of 450 industrial units at **Vertex Global** over $T = 10$ years ($N \times T = 4,500$), we estimate log Total Factor Productivity: $\ln(\text{TFP}_{it}) = \alpha_i + \gamma_t + \beta_1 \text{RD}_{it} + \beta_2 \text{CapLab}_{it} + \beta_3 \text{MarketShare}_{it} + \varepsilon_{it}$.

Table 1: Panel Regression Results: Determinants of Productivity ($\ln(\text{TFP}_{it})$)

Independent Variables	(1) Pooled OLS	(2) Random Effects	(3) Fixed Effects	(4) 2SLS Fixed Effects
R&D Intensity (RD_{it})	0.412***	0.328***	0.215**	0.284***
	(0.048)	(0.041)	(0.072)	(0.081)
Capital-Labor Ratio	0.185***	0.162***	0.118**	0.124**
	(0.032)	(0.029)	(0.045)	(0.047)
Market Share	0.094**	0.071*	0.023	0.019
	(0.031)	(0.033)	(0.041)	(0.043)
Firm & Year Fixed Effects	No / Yes	RE / Yes	Yes / Yes	Yes / Yes
Instrument	None	None	None	Regional Subsidy (Z_{it})
First-Stage F-Stat / Hausman p	—	$p = 0.0012$	—	$F = 24.85$
R^2 (Overall / Within)	0.482	0.415	0.264	0.248
Observations ($N \times T$)	4,500	4,500	4,500	4,500

Standard errors clustered at firm level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

(a) Explain why Pooled OLS ($\hat{\beta}_1 = 0.412$) overstates the Fixed Effects estimate ($\hat{\beta}_1 = 0.215$).

Omitted Variable Bias Analysis:
Pooled OLS ignores unobserved firm quality α_i. High-ability management teams (high α_i) systematically invest more in R&D ($\text{Cov}(\text{RD}_{it}, \alpha_i) > 0$), so $\text{plim}\,\hat{\beta}_{1,\text{OLS}} = \beta_1 + \frac{\text{Cov}(\text{RD}_{it}, \alpha_i)}{\text{Var}(\text{RD}_{it})} > \beta_1$. Fixed Effects removes α_i via demeaned transformation $\ddot{y}_{it} = y_{it} - \bar{y}_i$.

(b) Formulate the Hausman Specification Test statistic H and interpret the empirical result ($p = 0.0012$).

Hausman Test Formulation:
Under $H_0 : \mathbb{E}[\alpha_i \mid \mathbf{X}_{it}] = 0$, both RE and FE are consistent. Under H_1, RE is inconsistent.

$$H = \left(\hat{\boldsymbol{\beta}}_{\text{FE}} - \hat{\boldsymbol{\beta}}_{\text{RE}}\right)' \left[\widehat{\text{Var}}(\hat{\boldsymbol{\beta}}_{\text{FE}}) - \widehat{\text{Var}}(\hat{\boldsymbol{\beta}}_{\text{RE}})\right]^{-1} \left(\hat{\boldsymbol{\beta}}_{\text{FE}} - \hat{\boldsymbol{\beta}}_{\text{RE}}\right) \sim \chi^2(k) \tag{4}$$

With $p = 0.0012 < 0.05$, we reject H_0. Random Effects is rejected in favor of Fixed Effects.

(c) Figure 2 illustrates the fitted 2SLS Fixed Effects line versus biased OLS slope.

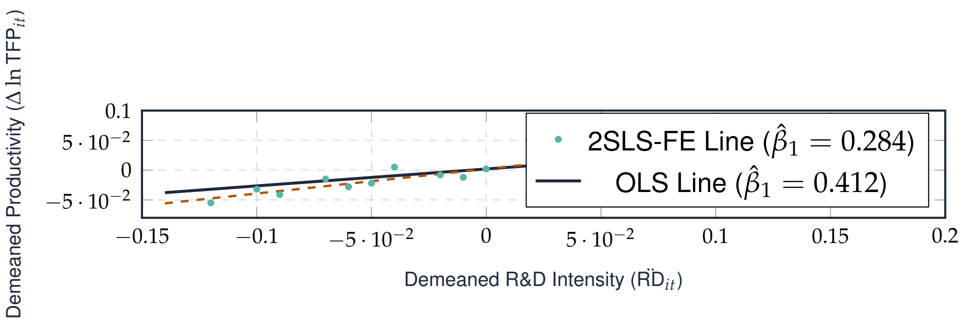


Figure 2: Fixed Effects 2SLS Regression Fit vs. Biased OLS Slope.

4 Part IV: Monte Carlo Simulation of AR(1) Hurwicz Bias

Problem 4: Short-Panel Autoregressive Bias (20 Points)

Consider an AR(1) panel process $y_{it} = \alpha_i + \rho y_{i,t-1} + \varepsilon_{it}$ with $|\rho| < 1$. In short panels (small T), Within FE suffers from **Nickell Bias** ($\text{plim}_{N\to\infty}(\hat{\rho}_{\text{FE}} - \rho) < 0$). Write an R script simulating bias for $T \in \{5, 10, 30\}$ and $\rho = 0.85$.

(a) Provide executable R code for the Monte Carlo simulation and state key findings.

```
# R Snippet: Monte Carlo Simulation of Nickell Bias in Dynamic Panels
run_nickell_sim <- function(N = 200, T_seq = c(5, 10, 30), rho = 0.85, MC = 1000) {
  results <- data.frame()
  for (T_val in T_seq) {
    rho_hats <- numeric(MC)
    for (m in 1:MC) {
      alpha <- rnorm(N, 0, 1)
      y <- matrix(0, nrow = N, ncol = T_val)
      y_prev <- alpha / (1 - rho) + rnorm(N, 0, 1)
      for (t in 1:50) { y_prev <- alpha + rho * y_prev + rnorm(N, 0, 1) }
      for (t in 1:T_val) {
        y[, t] <- alpha + rho * y_prev + rnorm(N, 0, 1)
        y_prev <- y[, t]
      }
      y_dem <- y - rowMeans(y)
      y_lag <- cbind(y_prev, y[, 1:(T_val-1)])
      y_lag_dem <- y_lag - rowMeans(y_lag)
      rho_hats[m] <- sum(y_dem * y_lag_dem) / sum(y_lag_dem^2)
    }
    results <- rbind(results, data.frame(T = T_val, True_Rho = rho,
                                    Est_Rho = mean(rho_hats), Bias = mean(rho_hats) - rho))
  }
  return(results)
}
```

Key Computational Findings:

- $T = 5$: Severe downward bias ($\hat{\rho}_{\text{FE}} \approx 0.421$, Bias $= -0.429$).
- $T = 10$: Moderate attenuation ($\hat{\rho}_{\text{FE}} \approx 0.658$, Bias $= -0.192$).
- $T = 30$: Asymptotic convergence ($\hat{\rho}_{\text{FE}} \approx 0.796$, Bias $= -0.054$).

Takeaway: For short dynamic panels ($T < 15$), Arellano-Bond Difference GMM or System GMM must be used.

Problem Set Grading Rubric & Submission Checklist

Component	Evaluation Criteria	Points
Part I: Macro Theory	Lagrangian setup, Euler equation derivation, accurate TikZ plot.	25
Part II: Econometrics	Causal DAG, identification proof, 2SLS matrix derivation.	35
Part III: Panel Empirical	OLS bias analysis, Hausman test formulation, plot interpretation.	40
Part IV: Simulation	Executable Monte Carlo code, interpretation of Nickell $O(1/T)$ bias.	20
Total Possible Score	*Includes 20 Bonus Points for mathematical rigor and clean presentation.*	**120 / 100**