ASSIGNMENT 2 (Arcidiacono and Miller, 2011)

Overview. This assignment concerns CCP-based estimation of dynamic oligopoly entry/exit games, based on Section 7.2 of Arcidiacono and Miller (2011, *Econometrica*). You will: (i) formally define the Markov Perfect Equilibrium and understand how finite dependence arises from terminal choices, (ii) solve for equilibrium and simulate data, (iii) estimate structural parameters using three different CCP methods and compare their performance, and (iv) compare CCP-based estimation to full solution methods that solve for equilibrium at each parameter guess. Throughout this assignment, we assume all state variables (including the demand shock s_t) are observed by the econometrician.

Instructions. Work in groups of about three. Each group should submit a single, self-contained report with well-documented code as an appendix. Hand written work will not be graded. Use the replication code provided in the `tutorial2/` folder. Questions will ask you to explain algorithms, modify parameters, and compare results. Report the specification of your computer (CPU, cores, memory) and computation times. All questions carry equal weight unless otherwise stated.

Model. Consider a dynamic infinite horizon entry-exit game with I potential firms in each market. Time is discrete with periods $t = 1, 2, \ldots$, and firms discount the future at rate $\beta \in (0, 1)$.

<u>State variables.</u> The state variables are $x_t = (x_1, x_{2t}, s_t)$ where:

- $x_1 \in \{1, 2, \ldots, 10\}$: Permanent market characteristic, drawn uniformly at the start.

- $x_{2t} = (d_{2,t-1}^{(1)}, \ldots, d_{2,t-1}^{(I)})$: Vector of incumbency status (whether each firm was active last period).

- $s_t \in \{1, 2, 3, 4, 5\}$: Demand shock following a first-order Markov chain with persistence π:
$$\Pr(s_{t+1}|s_t) = \begin{cases} \pi & \text{if } s_{t+1} = s_t \\ (1-\pi)/4 & \text{if } s_{t+1} \neq s_t \end{cases}$$

<u>Choice variables.</u> Each firm i chooses $d_t^{(i)} \in \{0, 1\}$ where:

- $d_t^{(i)} = 0$: Firm i exits or stays out of the market in period t.

- $d_t^{(i)} = 1$: Firm i enters or remains active in the market in period t.

Flow payoffs. The systematic component of current utility for firm i is:

$$U_0^{(i)}(x_t, s_t, d_t^{(-i)}) = 0 \quad \text{(exit/stay out)}$$

$$U_1^{(i)}(x_t, s_t, d_t^{(-i)}) = \theta_0 + \theta_1 x_1 + \theta_2 s_t + \theta_3 \sum_{i' \neq i} d_{2t}^{(i')} + \theta_4 \cdot \mathbf{1}\{d_{1,t-1}^{(i)} = 0\} \quad \text{(enter/stay)}$$

where $\theta_4 < 0$ is the entry cost paid only by entrants. Each firm also receives a private shock $\varepsilon_{jt}^{(i)}$ that is i.i.d. Type I extreme value across firms, choices, and time.

Information structure. Firms observe the common state x_t and their own private shocks $\varepsilon_t^{(i)} = (\varepsilon_{0t}^{(i)}, \varepsilon_{1t}^{(i)})$, but not the private shocks of their rivals. All state variables are also observed by the econometrician.

Parameters. Use the following baseline values from Arcidiacono and Miller (2011):

Parameter	Symbol	Value
Intercept	θ_0	0
Market characteristic	θ_1	0.05
Demand state	θ_2	0.25
Competition effect	θ_3	-0.20
Entry cost	θ_4	-1.5
Discount factor	β	0.9
Demand persistence	π	0.7
Number of firms	I	6

Question 1. Define the Markov Perfect Equilibrium for this game. Your answer should address:

(a) How rivals' strategies enter firm i's decision problem—specifically, how the flow utility $u_j^{(i)}(x_t)$ and transition probability $f_j^{(i)}(x_{t+1}|x_t)$ depend on rivals' conditional choice probabilities.

(b) The conditional value function $v_j^{(i)}(x_t)$ and the ex-ante value function $V^{(i)}(x_t)$, and how they relate to each other under Type I EV errors.

(c) Why equilibrium constitutes a fixed point problem and what object is being iterated on.

Question 2. This model exhibits finite dependence due to the terminal nature of exit.

(a) Explain how the terminal choice structure allows us to express $V(x')$ directly in terms of CCPs, without solving an infinite-horizon value function.

(b) Derive the identifying equation that links observed CCPs to structural parameters θ. Discuss how this simplifies both equilibrium computation and estimation relative to models without finite dependence.

(c) Suppose instead that both choices lead to continuation (e.g., "low investment" vs "high investment" with no exit). Explain why finite dependence fails in this case. What additional computational burden does this create for equilibrium computation and estimation?

Question 3. Implement the equilibrium computation and simulation.

(a) Describe the algorithm for computing equilibrium CCPs. What is being iterated, how are updates computed, and what determines convergence? Compare this to single-agent dynamic programming.

(b) Using the baseline parameters, solve for equilibrium. Report convergence diagnostics and computation time. Present the equilibrium CCPs and interpret the economic patterns (e.g., how do CCPs vary with competition, demand, and incumbency status?).

(c) Simulate panel data ($M = 3000$ markets, $T = 20$ periods). Report summary statistics on market structure and firm dynamics.

(d) Modify one or more parameters (e.g., entry cost, discount factor, demand persistence). Re-solve and re-simulate. Discuss how equilibrium behavior and market outcomes change.

(e) The equilibrium solver uses iterative methods that may converge to different fixed points depending on initialization. Run the solver from at least four different initial CCP values (e.g., $p_0 = 0.1, 0.5, 0.9$, and random). Do all initializations converge to the same equilibrium? If you find multiple equilibria, characterize their economic differences and discuss implications for estimation and counterfactual analysis.

Question 4. Arcidiacono and Miller (2011) propose three CCP-based estimation approaches: Two-Stage, CCP-Model, and CCP-Data.

(a) Explain each method. What is updated between iterations (if anything)? What are the theoretical trade-offs in terms of efficiency, robustness, and computational cost?

(b) Implement all three methods on your simulated data. Report parameter estimates, standard errors, and computation times in a single comparison table.

(c) Discuss which method performs best and under what circumstances each might be preferred.

Question 5. An alternative to CCP methods is the Nested Fixed Point (NFXP) approach, which solves for equilibrium at each parameter guess during likelihood optimization.

(a) Describe the NFXP algorithm. What is the "inner loop" and what is the "outer loop"? Why is this computationally expensive?

(b) Implement NFXP and estimate the model. Report estimates, standard errors, and computation time.

(c) Compare NFXP to the three CCP methods: accuracy of estimates, standard errors, and computation time. When might NFXP be preferred despite its cost?

Question 6. Assess the finite-sample performance of the estimators through Monte Carlo simulation.

(a) **Monte Carlo design.** Conduct a Monte Carlo study with at least 100 replications. For each replication: (i) simulate a new dataset with the baseline parameters, (ii) estimate using Two-Stage, CCP-Model, CCP-Data, and NFXP, and (iii) record point estimates and computation time.

(b) **Visualize sampling distributions.** For each structural parameter θ_j, create a figure showing the distribution of estimates across replications (e.g., histograms or kernel density plots). Include a vertical line at the true parameter value. Plot all four estimators in the same figure to facilitate comparison.

(c) **Summary statistics.** Report a table containing, for each parameter and each estimation method: mean estimate, bias (mean − true), standard deviation, and $\text{RMSE} = \sqrt{\text{bias}^2 + \text{variance}}$. Include mean computation time per replication.

(d) **Theoretical comparison.** Arcidiacono and Miller (2011, Section 4) and Aguirre-gabiria and Mira (2007, Section 3) discuss the asymptotic properties of CCP-based estimators. What does theory predict about efficiency rankings under correct specification? Do your Monte Carlo results align with these predictions? Discuss any discrepancies.

(e) **Sample size sensitivity.** Repeat the Monte Carlo study (with fewer replications if necessary) for $M \in \{500, 1000, 3000\}$ markets. How does RMSE change with sample size? Plot RMSE vs. M for each parameter and discuss what drives estimation precision in this setting.

Question 7. Investigate the robustness of the estimators to violations of key modeling assumptions.

(a) **Omitted state variables.** The demand shock s_t is serially correlated with persistence $\pi = 0.7$. Generate data using the full model where s_t is observed. Then re-estimate the model *excluding* s_t from the state space (i.e., estimate CCPs as functions of $(x_1, \text{incumbency}, n_{\text{rivals}})$ only). Compare the estimated parameters to the true values. Which parameters are most biased? Explain intuitively why omitting a serially correlated state variable creates problems similar to having serially correlated unobservables.

(b) **Serial correlation in private shocks.** The baseline model assumes i.i.d. Type I Extreme Value errors. Suppose instead that each firm has a persistent unobserved component ξ_{it} following an AR(1) process:

$$\xi_{it} = \rho \cdot \xi_{i,t-1} + \eta_{it}, \quad \text{where } \eta_{it} \sim N(0, 1 - \rho^2)$$

so that $\text{Var}(\xi_{it}) = 1$ in steady state. The firm's utility for entering/staying becomes:

$$\tilde{U}_1^{(i)} = U_1^{(i)} + \xi_{it} + \varepsilon_{1t}^{(i)}, \quad \text{where } \varepsilon_{1t}^{(i)} \sim \text{Type I EV (i.i.d.)}$$

Modify the simulation code to generate data with $\rho = 0.5$. (Note: choices are no longer simple logits—simulate by comparing realized utilities.) Estimate the model assuming i.i.d. shocks (ignoring ξ_{it}). Which parameters are most affected by this misspecification?

(c) **Discussion.** Based on your findings, discuss which assumptions appear most critical for reliable estimation. How might a researcher diagnose these issues with real data?

Question 8. Conduct a counterfactual policy analysis.

(a) Choose a policy experiment (e.g., reducing entry costs). Carefully describe the procedure for computing counterfactual outcomes. Why must we re-solve for equilibrium rather than simply plugging in new parameters?

(b) Implement the counterfactual and compare market outcomes under baseline vs. counterfactual scenarios. Discuss the economic implications.

(c) What are the key identification assumptions underlying this counterfactual? What limitations should be acknowledged?

Deliverables. Submit a report that:

(i) Answers each question with appropriate explanations and derivations.

(ii) Includes all tables, figures, and summary statistics requested.

(iii) Reports and interprets estimation results across all four methods (Three CCP + NFXP).

(iv) Discusses limitations and possible extensions.

Attach your code (well commented) as an appendix, highlighting any modifications you made to the provided code. Report computation times and computer specifications.

References.
Aguirregabiria, V. and Mira, P. (2007). Sequential Estimation of Dynamic Discrete Games. *Econometrica*, 75(1):1–53.
Arcidiacono, P. and Miller, R.A. (2011). Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity. *Econometrica*, 79(6):1823–1867.
Hotz, V.J. and Miller, R.A. (1993). Conditional Choice Probabilities and the Estimation of Dynamic Models. *The Review of Economic Studies*, 60(3):497–529.
Rust, J. (1987). Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica*, 55(5):999–1033.