NET 2023 Power Round

Advanced Division

March 2023

Instructions

This test consists of six questions. While you are free to attempt all six questions, we will only grade your four best-performing questions, regardless of how well you do on the other two. A question's point value is *not* informative of its difficulty; although questions have different point values, each question is weighted independently of its point value in your final cumulative score. After normalizing point-values of each question to the same weight, your cumulative score will be calculated as the sum of the scores of your four best-performing questions. You are encouraged to work together on these questions. Answer each question as clearly and succinctly as possible. You may write on a blank sheet of paper where you *clearly indicate* where your answer to each part is. If you are unsure of your answer, take your best guess: there is no penalty for incorrect answers. If you find yourself stuck on a question, skip it and return to it at the end if necessary. You will have two hours (120 minutes) to complete the exam. Remember, we do *not* share your answers or scores with Northwestern admissions, nor do we keep them for ourselves. You are not expected to know how to answer each question on the exam; rather, this test is designed to assess your economic and formal reasoning skills. Have fun, and good luck!

Problem 1: Comrade, to Infinity and Beyond!

In this question, we consider how game theory can be used to model cooperation in settings where naive applications of basic game theory may at first predict otherwise. This problem requires no mathematical prerequisites outside of the ability to sum infinite geometric series.

Part A *(6 points total)* Consider the game between two players, represented in the following payoff matrix where player 1's actions are represented in the rows and their payoffs are in the first coordinate.

	C	D
C	$(3,3)$	$(-1,4)$
D	$(4,-1)$	$(0,0)$

Table 1: Stage Game

(1) *(1 point)* Suppose you could force the individuals to choose an outcome. Which one would maximize payoffs for both players?

(2) *(3 point)* What is the *unique* Nash equilibrium of the game? Explain why this is a Nash equilibrium and why this is the only Nash equilibrium.

(3) *(2 points)* Comment on the friction between Part (A) and Part (B). Is this reasonable?

Part B *(2 points)* As concisely as possible, give one real-world example of a situation that we might want to model using a matrix like this. Is the prediction of the game reasonable in the context of economic intuition? What about in the context of observed behavior?

Part C *(7 Points Total)* Now assume instead that this game is played twice in a row.

(1) *(2 points)* Assume we are in the second period. Without knowing anything about the first period, what will happen in equilibrium in the second period? *Hint: Think of the sunk cost fallacy.*

(2) *(2 points)* Given your answer in the second period, what will individuals do in the first period? Explain.

(3) *(3 points)* Extrapolating on your answer here, is it possible for (C, C) to be played in any period of this game in equilibrium? What if the game ends in $N > 2$ periods? Explain.

Part D *(14 points total)* Assume now that individuals repeat this game infinitely many times, and discount the future at a rate of δ. For example, if both players cooperate in every period, then their payoff for both players would be would be

$$\sum_{t=0}^{\infty} 3\delta^t$$

while if they played (C, C) in even periods and (C, D) in odd periods, the payoffs for players 1 and 2, respectively, would be

$$\sum_{t=0}^{\infty} 3\delta^{2t} + \sum_{t=0}^{\infty}(-1)\delta^{2t+1} \text{ and } \sum_{t=0}^{\infty} 3\delta^{2t} + \sum_{t=0}^{\infty} 4\delta^{2t+1}$$

(1) *(2 points)* Show that playing (D, D) in every period regardless of what the other person is doing is a Nash equilibrium in this game. *(Hint: One way to do this is by showing that if one player players D in every period regardless of what the other player does, then they cannot do better than playing D in every period as well.)*

(2) *(1 point)* Assume instead that players play (C, D) in the first period, and (D, D) in every period after that. What is the payoff to player 2? Player 1?

(3) *(4 points)* Consider now the following *(grim-trigger) strategy* for player 2: player 2 will play C if, in every period before, player 1 played C (and will play C in the first period). If player 1 plays D in any period, player 2 will play D forever. Show that if player 2 is playing a grim trigger strategy, player 1 will want to play C in every period if $\delta = 0.9$.

(4) *(3 points)* Using your previous answers, explain why it is a Nash equilibrium for both players to play the grim trigger strategy when $\delta = 0.9$. What will the observed behavior be?

(5) *(2 points)* Find the smallest value of δ that can sustain (C, C) in every period.

(6) *(2 points)* Assume now that you are an economist defending the assumption that individuals are playing this game infinitely times against a skeptic who objects that individuals only live for finitely many periods. Give one (potential) justification for this assumption.

Part E *(2 points)* Compare and contrast your answers in Parts (C) and (D) in the context of Part (B).

Bonus *(4 points) (The Folk Theorem under Perfect Monitoring)* Show that, for some sufficiently large δ, a modified version of the grim trigger strategy can give both players any possible payoff between $(0, 3)$. Thus, for any possible mixed strategy equilibrium, there is an equilibrium where both players play grim-trigger strategies which is payoff equivalent. In particular, for sufficiently large discount rates, *any* observed behavior can be rationalized for sufficiently patient individuals!

Problem 2: Exchange Rates and Arbitrage (25 points)

(Inconsistencies in exchange and interest rates between countries lead to arbitrage. This problem compares two ways to resolve these inconsistencies.)

Part A *(1 point)* An exchange rate is the price of one currency in terms of another. If a particular jacket costs 800 Canadian dollars, and the exchange rate is \$1 CAD = \$0.75 USD, how much does that jacket cost in USD?

We can differentiate between certain exchange rates based on periods of time. The *spot exchange rate* is the current exchange rate between currencies. In contrast, the *forward exchange rate* is the exchange rate at a point in time in the future. Sometimes, disparities between these two rates create arbitrage: the possibility to make risk-free profit after accounting for transaction costs.

Part B *(4 points)* Suppose the spot exchange rate between CAD and USD is 0.75 and the forward exchange rate (one year into the future) is 0.8. Furthermore, suppose the annual real interest rate in Canada is 10% while it is 2% in the United States. Demonstrate how this creates arbitrage by describing a possible investment strategy to make risk-free profit.

Part C *(2 points)* What is the rate of return of the strategy you described in Part B?

Interest Rate Parity is an economic equilibrium in which arbitrage due to disparities in exchange and interest rates cannot occur. One way of achieving such an equilibrium is through the proper setting of forward exchange rates. This equilibrium is known as *covered interest rate parity.*

Part D *(2 points)* In the scenario described in Part B, what would the forward exchange rate have to be to achieve interest rate parity?

Part E *(4 points)* Based on your work from the previous part, develop an equation for F_1, the covered interest rate parity forward exchange rate one year into the future. Your answer should be in terms of the interest rate in the domestic country, i_D, the interest rate in the foreign country, i_F, and the spot exchange rate, S.

Part F *(3 points)* Now generalize your equation to any period of time. That is, develop an equation for F_t, the equilibrium forward exchange rate t years in the future.

Alternatively, equilibrium can be achieved without the use of forward exchange rates. Instead, it involves adjusting the spot exchange rate to adapt to predicted changes in interest rates between countries. This is known as *uncovered interest rate parity.*

Part G *(2 points)* Return to the scenario in Part B. To prevent arbitrage, what should a policymaker set the spot exchange rate to?

Part H *(4 points)* Based on your work from the previous part, develop an equation for the uncovered interest rate parity spot exchange rate, S. Your answer should be in terms of the interest rate in the domestic country i_D, the interest rate in the foreign country i_F, and the expected forward exchange rate F.

Part I *(3 points)* When are covered interest rate parity and uncovered interest rate parity theoretically the same? Provide either an intuitive or mathematical justification. (Hint: your answer should involve forward and spot exchange rates.)

Problem 3: If Everyone Jumped Off a Bridge, You Would Too! (20 points)

In this problem, we consider how agents learn from one another and build a model (of recommendations) to examine why bad businesses may flourish in some conditions. This problem requires no mathematics besides an intuitive understanding of conditional probability.

Preliminaries Assume that there are two colleges, Northwestern and UChicago. One school is good, and one school is bad (pretend we do not know which is which). Sequentially, individuals need to choose a school without observing which school is good. However, each individual student receives a *signal*, $s \in \{N, C\}$. The signal is *accurate*, so that if Northwestern is the good school, then the signal realization will be N with probability $p > \frac{1}{2}$. Meanwhile, if UChicago is the good school, the signal realization will be C with probability $q > \frac{1}{2}$. Each student can observe both the actions of past students, and (sometimes) the signals those students received as well.

Part A *(4 points)* Assume there are $\{t\}_{t=0}^{\infty}$ many students. Thus, there will be $\{s_t\}$ many signals.

(1) *(1 point)* Suppose Northwestern is the good school and $p = \frac{3}{4}$. In the first one million signals, how many N signals do we expect to see?

(2) *(2 points)* Suppose student 1 million sees exactly as many N signals as predicted before. Which school will they choose? What about students after that?

(3) *(1 point)* Give an intuitive argument for why, after sufficiently many previous students receive signals, *every* student will choose the good school.

Part B *(4 points)* Let $p = q$. Consider the individual who first observes a signal, before anyone else.

(1) *(1 point)* Suppose the individual sees signal N. What is the probability of seeing N when the good school is Northwestern? What about the probability of seeing N when the good school is UChicago?

(2) *(1 point)* Suppose Northwestern and UChicago are equally likely to be the good school. Then what is the total probability of seeing N?

(3) *(1 point)* Based on your answers in (1) and (2), what is the probability that Northwestern is the good school, conditional on seeing signal N?

(4) *(1 point)* Using your answers above, show that if the individual sees signal N, they will pick Northwestern, and if they see C, they will pick UChicago.

Part C *(6 points)* Consider now the point of view of the second individual who can see their own signal and the action (but not the signal) of the first individual.

(1) *(1 point)* Suppose the second individual sees the first individual pick UChicago. Does the second person know what signal this person got, and if so, which one?

(2) *(2 points)* Suppose this individual saw signal C. Which school will they choose? Why?

(3) *(2 points)* Suppose instead they see signal N after the first person picked UChicago. What is the probability of this happening if the good school is Northwestern? UChicago?

(4) *(1 point)* Explain why the second person will "follow their signal" regardless of the first person's action: they will choose Northwestern if they see N and UChicago if they see C.

Part D *(7 points)* Consider now the third individual, and suppose both individuals beforehand had picked Northwestern.

(1) *(2 points)* What is the probability that both individuals picked Northwestern if it was the good school? What about if it was the bad school?

(2) *(1 point)* Thus, conditional on seeing the two other individuals seeing Northwestern, what is the probability it is the good school?

(3) *(1 point)* If the third person sees signal N, what school will they pick?

(4) *(2 points)* Suppose now the third person sees the signal C. What school will they pick? Justify your answer probabilistically. *(Hint: this sub-question is computationally longer than the others. Compute the conditional probabilities of seeing C and the other two individuals picking Northwestern conditional on Northwestern and UChicago being the good school, respectively.)*

(5) *(1 point)* If both individuals one and two picked the same school (say Northwestern), how many individuals after them will also pick Northwestern?

Part E *(2 points)* Suppose Northwestern is the good school. Is it possible that students will go to UChicago (the equivalent of jumping off a bridge[1]) even though it is the bad school if (i) they can see the *signals* past students received, or (ii) they can only see whether past students chose Northwestern or UChicago? Interpret your answers using economic intuition. *(Hint: Use Parts (A) - (D))*

Part F *(2 points)* This behavior is often known as *herding*. Can you think of another example other than the one given in this problem where herding phenomena occurs?

[1] This is a joke.

Problem 4: Stimmies and the Permanent Income Hypothesis (25 points)

(In this problem, we rigorously microfound the permanent income hypothesis of Milton Friedman and consider how it interacts with the Keynesian hypothesis.)

Part A *(2 points)* Suppose we are in a Keynesian economy with a marginal propensity to consume of 0.8 and no taxes. What is the expenditure multiplier? If the government exogenously increases spending by 100 million dollars, by how much will GDP increase?

Part B *(2 points)* Now consider a microfounded version of this problem. Assume there is a (single) representative agent with the following utility function from consumption, c_t:

$$u(c_t) = b_1 c_t - \frac{b_2}{2} c_t^2$$

The consumer is infinitely lived, and discounts the future at rate β. Suppose that the consumer consumes the sequence $\{c_t\}_{t=0}^{\infty}$. What is the lifetime utility of this stream to the consumer?

Part C *(2 points)* Assume now that, instead of facing a consumption stream, the consumer faces a stream of *random* income[2], $y_t \in \mathcal{Y}$, which is independent across time. In period t, they can choose to split their income in two ways: they can either invest it in an asset, or they can choose to consume it. If they invest it in an asset, they will get a return of $(1+r) = \frac{1}{\beta}$ in the next period. Suppose a_t is the amount of savings that an individual has at time t. Write out a_{t+1}, the savings they will have at time $(t+1)$, as a function of (r, y_t, a_t, c_t). *(Hint: your function will require only addition and multiplication. Think economically!)*

Part D *(1 point)* Explain why, using your answers from Parts (B) and (C), along with the fact that the income is *random*, that the optimal sequence of consumption and savings $\{c_t, a_{t+1}\}$ can be found by solving the maximum expressed below:

$$\max_{\{a_{t+1}, c_t\}} \left\{ \sum_{t=0}^{\infty} \mathbb{E}\left(\beta^t \left(b_1 c_t - \frac{b_2}{2} c_t^2\right)\right)\right\} \text{ s.t. } a_{t+1} = (1+r)(y_t + a_t - c_t)$$

We will also add the assumption the individual cannot gain utility by saving money forever:

$$\mathbb{E}\left[\lim_{T\to\infty} \beta^T a_T^2\right]$$

This is to ensure the problem is well-defined; if this confuses you, feel free to ignore it. We are now ready to do some calculus to explicitly solve the problem.

Part E *(7 points)* It is a fact of mathematical optimization that the problem (with constraints) defined above will have the same maximizers as the following unconstrained problem (called the *Lagrangian*):

$$\mathcal{L}(\{c_t\}, \{a_{t+1}\}) = \mathbb{E}_{\{y_t\}\sim\mathcal{Y}}\left[\beta^t \left(\left(b_1 c_t - \frac{b_2}{2} c_t^2\right) + \lambda_t((1+r)(y_t + a_t - c_t) - a_{t+1})\right)\right]$$

for some weakly positive sequence $\{\lambda_t\}$, which are called *Lagrange multipliers.* The second term follows by subtracting the constraint from itself, so it is negative if the constraint doesn't bind and zero otherwise.

(1) *(2 points)* Differentiate in c_t as a variable (noting this is different from c_{t+1}, c_{t-1}, and the other indices) and set it equal to 0. You can use without proof in this section that

$$\frac{d}{dx}\mathbb{E}[f(x)] = \mathbb{E}\left[\frac{d}{dx} f(x)\right] \text{ and } \mathbb{E}[0] = 0$$

(Hint: your answer in c_t should not have an expectation operator).

[2]Note: If it helps, you may assume $y_t = 0$ with probability $\frac{1}{2}$ and 1 with probability $\frac{1}{2}$ and manipulate this distribution without loss of credit from here on out. For full credit, though, we ask you to clearly demarcate when this substitution is used.

(2) *(1 point)* Rearrange the derivative so that you obtain c_t as a function of the parameters (λ_t, r, b_1, b_2).

(3) *(2 points)* Differentiate now instead in a_{t+1}, noting that in this case, a_{t+1} shows up twice: in the time t equation and the time $t+1$ equation. *(Hint: in the time $t+1$ portion of the equation, e.g. the part with β^{t+1}, there will be an expectation. Your answer should be of form $\beta^t \lambda_t = \mathbb{E}[something]$).*

(4) *(2 points)* Recall $\beta = (1+r)^{-1}$. Use the equations from the above computations to prove

$$c_t = \mathbb{E}[c_{t+1}]$$

(Hint: Using the result from Part (III), get an expression for λ_{t+1}. Relate this to consumption using the first order conditions in c_t and c_{t+1}).

Part F *(2 points)* In Part (E), you showed that $c_t = \mathbb{E}[c_{t+1}]$. Interpret this condition economically, and explain how the model predicts consumers predict their consumption today.

Part G *(2 points)* Using your answer in Parts (E) and (F), write c_t as a function of r and c_{t+j} for some $j \in \mathbb{N}$ periods in the future.

Part H *(3 points)* Define future expected wealth at time t to be

$$W_t = a_t + \sum_{j=0}^{\infty} \beta^j \mathbb{E}_t(y_{t+j})$$

Using your answer in Part (G), show that

$$c_t = \frac{rW_t}{1+r}$$

(Hint: sum a geometric series and use the fact $\beta = (1+r)^{-1}$)

Part I *(2 points)* Using your answers above, explain how you have just proved *Friedman's permanent income hypothesis*:

$$\Delta c_t = c_t - c_{t-1} = \frac{r}{1+r}(W_t - W_{t-1})$$

that is, the change in consumption today is a weighted fraction of the total change in lifetime total wealth induced by an exogenous shock.

Part J *(2 points)* Compare and contrast your answers from Parts (A) and (I). Use this contrast to explain why, even if the marginal propensity to consume is very high, stimulus may not be as effective as the baseline Keynesian model predicts.

Problem 5: An Orchestra of Tools (25 points)

(In this problem, we consider how to estimate a dependent effect when the independent variable is not easily measurable but correlated with another measurable event.)

Part A *(2 points)* Suppose we would like to study the effect of education on wages, and we have a large sample of individuals' education and wages. We can run the regression

$$w_i = \beta e_i + \varepsilon_i$$

where e_i is individual i's education (measured in years of schooling), w_i is individual i's wage, and ε_i is everything else that we cannot observe. Suppose that both $\mathbb{E}[\varepsilon_i] = 0$ and $\mathbb{E}[\varepsilon_i | w_i, e_i] = 0$; that is, our unobservables are independent of the data we can observe, and the unobservables will eventually "wash out." How should we interpret β?

Part B *(3 points)* Suppose now that $\mathbb{E}[\varepsilon_i] = c \neq 0$, so that there is (constant) systematic bias in the errors. Show that this problem can be fixed by subtracting a constant term from all of your observed data, and will recover the same β as if $\mathbb{E}[\varepsilon_i] = 0$.

Part C *(3 points)* Suppose now, however, that one unobservable is ability, and higher-ability individuals are likely to get higher wages even with the same amount of education (though we do not know by how much). Will this violate the assumption $\mathbb{E}[\varepsilon_i | w_i, e_i] = 0$? How will this affect our estimate of β?

Part D *(2 points)* Explain intuitively why this problem is much harder to fix, and cannot be done in the same way as we did for the problem in Part (B).

Part E *(3 points)* Suppose now that we can have information about each individual's SAT scores as well. For each of the following, explain if you think that the SATs will be very correlated or only a little correlated with the listed variable, and (briefly) explain why.

(1) Wages at a company.

(2) The education a student receives.

(3) Other unobservables (e.g. gender, location, height, etc.)

Part F *(2 points)* Let c_i denote individual i's college exam score and suppose that c_i is uncorrelated with the error term ε_i (so that $\mathbb{E}[\varepsilon_i | w_i, c_i] = 0$). Consider the estimate

$$w_i = \delta c_i + \varepsilon_i$$

Will our estimate of δ be accurate? What will δ measure?

Part G *(6 points)* The variable c_i is called an *instrument*. Ultimately, we do not care too much about δ, but want to estimate β. How can we do that?

(1) *(2 points)* Suppose we can estimate the effect of college test scores on education. Using the noise term η_i and coefficient α, write a regression with e_i as the dependent variable and c_i as the independent variable, under the assumption $\mathbb{E}[\eta_i | e_i, c_i] = 0$.

(2) *(2 points)* Substitute your equation for w_i into the regression from Part (A). Interpreting the new error term as $\delta\eta_i + \varepsilon_i$, and under the assumptions we have made, is it true that $\mathbb{E}[\delta\eta_i + \varepsilon_i | w_i, c_i] = 0$?

(3) *(2 points)* Write β as a function of δ and α. Will this be an accurate estimator?

Part H *(4 points)* This regression method in Part (G) is called *two-stage least squares* regression. Can you think of situations where the estimation strategy may fail? Give two different, concrete examples or situations where the instrument may not be "valid."

Problem 6: Supply Chains and Con-shoe-mer Welfare (25 points)

(This problem uses a game theoretic approach to derive conclusions about economic welfare in a simple model of a vertically integrated supply chain.)

Preliminaries Suppose there are only two firms in the market for shoes. Firm 1 produces shoelaces. Firm 2 then buys shoelaces from Firm 1 and uses them to produce shoes. Firm 2 sells these finished shoes to consumers. Assume Firm 1 and Firm 2 are two independent firms. That is, they do not collude in order to maximize their collective profits. Suppose that the following two-stage game is played:

1. Firm 1 chooses to announce a positive quantity of shoelaces (denoted q_1) that they will produce. This automatically determines the price of each shoelace, which is, for positive constants a and b, given by the equation:

$$p = a - bq_1$$

2. Firm 2 announces the quantity of shoes (denoted q_2) they will produce. They then buy enough shoelaces (one shoelace for each shoe) to produce that many finished shoes. For simplicity, we assume that consumers will buy all shoes produced by Firm 2.

We express the market quantity of shoes sold, Q, as $Q = \min(q_1, q_2)$. The market price, P, of a shoe is expressed by the following linear inverse demand function:

$$P = \alpha - \beta Q$$

where α and β are positive constants.

Part A *(1 point)* Let π_1 represent the profit for Firm 1. Suppose that the marginal cost for the production of each shoelace is c_1. Determine an expression for π_1.

Part B *(2 points)* Let π_2 represent the profit for Firm 2. Suppose that the only costs incurred for Firm 2 is the cost of buying shoelaces. Determine an expression for π_2.

Part C *(4 points)* Suppose that we fix the quantity of shoelaces produced by Firm 1 (that is, treat p_1 and q_1 as constants). Determine the optimal quantity q_2^* of shoes that Firm 2 ought to produce. Your answer should be a function of q_1.

Part D *(4 points)* Using the previous part, Firm 1 can reasonably predict the number of shoes Firm 2 will produce in response to what Firm 1 announces as their q_1. Knowing this, determine the optimal quantity of shoelaces that Firm 1 should announce.

Part E *(4 points)* Based on the optimal q_1 and q_2 announced by Firms 1 and 2, determine the resulting profits for each firm, π_1 and π_2, the market price, P, and market quantity, Q.

We now compare the results of this situation to that of a monopoly. Suppose Firms 1 and 2 combine into a single firm which has a monopoly over the shoe market. Now, the combined firm announces a single market quantity, q_M, of shoes which it will produce. Price is determined by the same market inverse demand function.

Part F *(2 points)* Determine an expression for π_M, the total profit of the monopoly.

Part G *(2 points)* Assuming consumers will buy every shoe produced by the monopoly, what is the optimal quantity, q_M^*, that the monopoly will choose to announce?

Part H *(3 points)* Suppose the two combined firms split monopoly profits equally. In which market structure is each firm better off? Use your previous work to justify your answer.

Part I *(4 points)* Calculate the monopoly price and monopoly quantity for this market. Are consumers better off under a monopoly?