Industrial Organization Field Exam, 2025

This exam is comprised of three sections.

The first section is for material covered in ECON 220A taught by Nano Barahona. The second section is for material covered in ECON 220B taught by Matthew Backus and Quitze Valenzuela-Stookey. The third section is for material covered in ECON 220C taught by Ben Handel and Carolyn Stein. Each section has 100 points.

You have to answer only two sections. If you attempt to answer more than two sections, be very clear about which two sections you want us to grade. We will only assign points on two sections. For Sections 2 and 3, please be sure to answer both parts, i.e. for both instructors. You have 3 hours to complete the exam.

1 Section #1 (220A) - 100 points

Nevo (2001) study markups and conduct in the ready-to-eat cereal industry. They use scanner data from the IRI Infoscan Data Base to estimate a model of supply and demand for cereal.

They use a random coefficient logit (RCL) model to estimate consumer demand. The indirect utility that consumer i receives from inside good j in market (city-quarter) t is

$$u_{ijt} = x_j\beta_i^* + \alpha_i^* p_{jt} + \xi_j + \Delta\xi_{jt} + \epsilon_{ijt}, \tag{1}$$

where x_j is a K-dimensional vector of observable characteristics, p_{jt} is the retail price, ξ_j is the national mean valuation of the unobserved (by the econometrician) product characteristics, $\Delta\xi_{jt}$ is a market specific deviation from this mean, and ϵ_{ijrt} is a stochastic extreme value 1 shock.

The observable product characteristics include a constant (i.e., an indicator that equals 1 for an inside good), sugar, mushy (i.e., whether the product gets gets soggy in milk), fiber, and dummies on whether the product belongs to the all-family, kids, or adults segments. They control for ξ_j using product dummy variables, and specify the consumer-specific coefficients as $[\alpha_i^*, \beta_i^*]' = [\alpha, \beta]' + \Pi D_i + \Sigma v_i$, where D_i is a vector of consumer demographics (income, age, and child) and v_i is an unobserved consumer characteristics assumed to follow a standard-normal distribution.

Answer the following questions about the paper. Be brief in your answers.

1. (13 points) Explain what is the main hypothesis that the author is trying to test in the paper. What's the overall strategy?

2. (15 points) Nevo (2001) estimate the demand model following BLP (1995). Write down a pseudo-code that can estimate the model. Be explicit about every step.

Nevo (2001) estimate the demand using the steps you provided above, which requires a set of instruments Z. An important limitation of their setting, is that there is no variation in each brand's observed characteristics and choice set composition over time and across cities. To identify the model, they use an approach similar to that used by Hausman (1996) and exploit the panel structure of the data.

3. (12 points) Explain what are moment conditions used to estimate the model. How many moments are there? How many parameters? Discuss the validity of the instruments in terms of relevance and exclusion. Do you see any potential identification issues?

4. (12 points) Suppose that Nevo (2001) data had variation in product characteristics or choice sets. Explain what moments you would have added to the estimator to better estimate Π and Σ. What are the benefits of adding these additional moments relative to what Nevo (2001) did?

5. (12 points) Suppose Nevo (2001) also had access to micro-data with information linking purchases to individual family characteristics. Explain how you would have used this information to identify the model parameters. Which model parameters could you estimate more precisely with these data?

6. (12 points) Nevo (2001) presents an alternative identification strategy where the brand dummy variables are not included as regressors but are used as IV's instead. According to the paper, using the brand dummy variables as IV's is a nonparametric way to use all the information contained in product characteristics. In Nevo's words in Appendix B, "this is the equivalent of using brand characteristics as IV's." Discuss.

7. (12 points) According to Nevo (2001), Quaker CapN Crunch (a kids cereal) and Post Grape Nuts (a wholesome simple nutrition cereal) have similar market shares. Suppose we are interested in learning what happens to the market shares of these two brands when the price of another kids cereal goes up using Nevo's estimated model. Explain what are the potential problems with this exercise if Nevo (2001) had used a plain logit instead of a random coefficients logit to estimate demand.

8. (12 points) Describe the main finding of the paper. What does it say about conduct in this industry? What explains high markups?

2 Section #2 (220B) - 100 points

Part 1 of 2, Backus:

Please note: all of these questions can be answered in just a few sentences or lines of algebra. If you get carried away writing long answers, you will risk running out of time.

Conduct Testing: In Backus, Conlon, and Sinkinson (2021) you learned about a procedure for testing models of conduct in firm pricing based on a comparison of the implied marginal costs. Suppose that you have already estimated demand $s_j(\mathbf{p})$ and the associated matrix of demand derivatives, $\Omega_{jk} = \partial s_j(\mathbf{p})/\partial p_k$.

BCS use a set of variables z to test the moment restriction $A(z_{jt}) \cdot \omega_{jt}^m = 0$ for competing models.

1) First, using the demand system, show how to solve for marginal costs under the behavioral assumption of "single-product pricing," i.e. that all products are priced as if they were owned by an independent firm.

2) For each of the following classes of variables, explain why it is either useful or not useful for discriminating between models of firm conduct in pricing:

1. Demand shifters.
2. Cost shifters.
3. BLP instruments.

3) Give an example of a pair of models of firm conduct in the pricing game that are *not* testable in this framework.

Production Function Estimation: We learned about production function estimation in the context of Ackerberg, Caves, and Frazer (2015). For the purposes of this question, assume we are interested in a Cobb-Doubglas production technology with two inputs, k and ℓ, with k a dynamic input and ℓ perfectly variable:

$$y_{jt} = \beta^\ell \ell_{jt} + \beta^k k_{kt} + \omega_{jt} + \varepsilon_{jt}$$

Recall that ω_{jt} in this setting stands in for the part of productivity that is known to the firm but not the econometrician, while ε_{jt} is the part of productivity that is unknown to either.

4) ACF show that the labor coefficient β^ℓ is not identified in the first-stage regression of the Olley and Pakes (1995) framework, which used investment i_{jt} to construct a proxy for ω_{jt}

$$OP: \quad y_{jt} = \beta^\ell \ell_{jt} + \Phi(k_{jt}, i_{jt}) + \varepsilon_{jt}$$

Briefly explain why β^ℓ is not identified.

5) Despite this, they still run their own version of first-stage regression, with ℓ_{jt} inside of $h(\cdot)$.

$$ACF: \quad y_{jt} = \Phi(k_{jt}, \ell_{jt}, i_{jt}) + \varepsilon_{jt}$$

If this regression doesn't identify β^ℓ, what does it accomplish?

* In case it is confusing, let me acknowledge that ACF use m_{jt} instead of i_{jt} as their proxy variable, following Levinsohn and Petrin (2003), but that is not important here.

6) Some have criticized ACF, along with OP and LP, because the estimator does not permit generic models of market power in the output market. Give an example of how such market power could lead to a violation of the assumptions of the model.

Part 2 of 2, Valenzuela-Stookey:

This question considers the version of the model from Ortner et al. (2024) "Mediated Collusion" without mediation. The model is as follows. There are $n \geq 2$ firms (bidders) competing in a first-price procurement auction with reserve price 1. All firms have a production cost of 0. Firms simultaneously place bids $\mathbf{p} = (p_i)_{i=1}^n$ where $p_i \in [0, 1]$ for all i. If a unique firm i places the lowest bid p_i, they win the auction and receive payment p_i. Losing firms receive 0. In case of a tie, the winner is chose uniformly at random.

The firms can jointly commit to impose a penalty of size x on some firms, as a function of the information revealed by the auction. Assume for this question that the penalty is only ever imposed on the winning firm (this is in fact without loss of generality). Additionally, we restrict attention to the case where the only information revealed by the auction is the

identity of the winning bidder (but not their bid). A *strategy profile* in this case consists of a bidding strategy for each firm (a distribution over $[0, 1]$) and a punishment strategy (a map from the identity of the winning bidder to a probability of punishment for that bidder). Let the CDF of F_i be firm i's bids, and let γ_i be the probability that i is punished conditional on winning. A strategy profile is a (Bayes Nash) equilibrium if each firm's bidding strategy maximizes its expected payoff, given the punishment strategy and the strategies of the other firms.

The firms commit to the punishment strategy up front to maximize "cartel profit", which is defined as the sum of the firms' expected payoffs or, equivalently, the expected winning bid (taking for granted that the penalty is not imposed on path).

1. Show that the winning bid cannot exceed x with positive probability in any equilibrium. (*Hint:* As an intermediate step, prove that in equilibrium it cannot be that every firm j gets a strictly positive expected payoff by bidding at the supremum of the support of F_j.)
2. Construct an equilibrium such that cartel profits equal x. (*Hint:* there is one such equilibrium in which there is a single bidder who wins with probability 1.)

3 Section #2 (220C) - 100 points

Part 1 of 2, Stein:

This question primarily focuses on Budish, Williams, and Roin (2015) "Do Firms Underinvest in Long-Term Research? Evidence from Cancer Clinical Trials." To remind you of the setup of this paper, here is some notation:

- t_{invent} is the date that idea is invented, and is normalized to 0. t_{comm} is the time that it takes to commercialize the idea and bring it to market
- It costs c to try to commercialize an idea, which must be paid at t_{invent}. If this cost is paid, the idea is successfully commercialized with probability p
- If the product is successfully commercialized, it is useful until superseded. This risk of obsolescence is $(1 - \gamma)$ every year
- The firm makes profit π for every year that their successfully commercialized idea is under patent and non-obsolete. The entry of a generic once patent protection ends or obsolescence drives profits to 0.

- For every year that the invention is under patent, it generates social value of v^{monop}. If the product were instead priced at marginal cost by a social planner, it would generate social value of $v > v^{monop}$ per year.
- The discount rate is r. Thus, the social planner's effective discount factor (including obsolescence) is $\delta = \frac{\gamma}{(1+r)} < 1$. Private firms discount by an additional factor $\eta < 1$ due to corporate short-termism, leading to a discount factor of $\eta\delta$
- Patents last for a term of t_{patent} from the filing date. Firms can file at either t_{invent} or t_{comm}. If they file at t_{invent}, they get the patent for sure. If they file the patent at t_{comm}, they get the patent with probability $q \leq 1$ reflecting the risk of being beaten by another firm to file for the patent.

Please answer the following:

1. [2 points] The authors focus on the case where $q = 0$. What does this imply, and why do the authors do this?
2. [6 points] Under what condition does a profit-maximizing firm attempt to commercialize an invention? When writing out this condition, highlight the "effective monopoly life" or EML (this is the expected number of years, discounted by the firm, that the firm will have monopoly rights).
3. [6 points] Now suppose that a social planner owned the firm, and would price at marginal cost if the invention is successfully commercialized. When would the social planner attempt to commercialize an invention? When writing out this condition, highlight the "effective total life" or ETL (this is the expected number of years, discounted by society, that the invention will be commercialized and non-obsolete).
4. [6 points] When does the decision to invest differ between a firm and the social planner? Are we more likely to have this scenario when EML/ETL is large or small?
5. [6 points] If t_{comm} increases, what does this do to EML/ETL? What does this imply for firm investment?
6. [5 points] Figures 2 and 3 in the paper (reproduced below) try to test the comparative static above. Explain how the graphs relate to the comparative static, and what concerns you might have about these empirical tests.

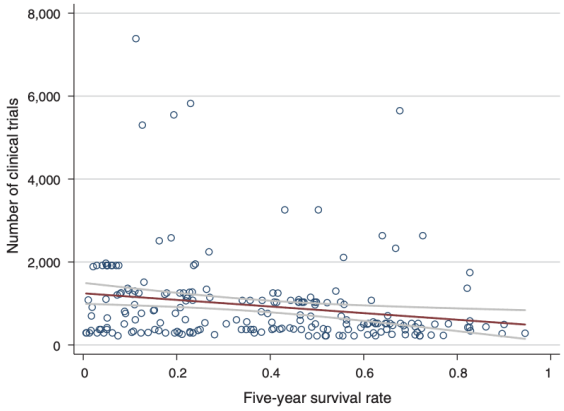

FIGURE 2. SURVIVAL TIME AND R&D INVESTMENTS: CANCER-STAGE DATA

Notes: This figure shows the relationship between the five-year survival rate among patients diagnosed with each cancer-stage between 1973–2004 (the cohorts for which five-year survival is uncensored), and the number of clinical trials enrolling patients of each cancer-stage from 1973–2011. Note that because we here count the number of clinical trials patients of each cancer-stage are eligible to enroll in, a higher count of trials appears here than in Figure 1 because many trials enroll patients of more than one cancer-stage type. The level of observation is the cancer-stage. For details on the sample, see the text and online Data Appendix.

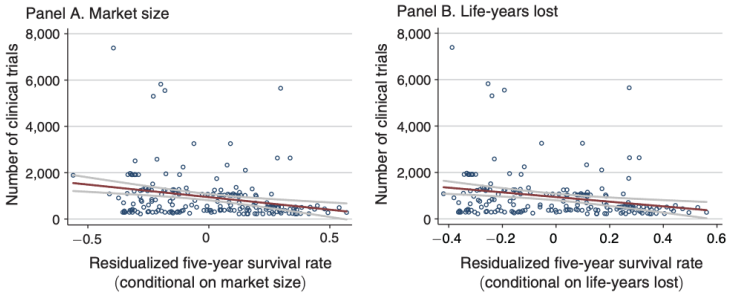

FIGURE 3. SURVIVAL TIME AND R&D INVESTMENTS: RESIDUALIZED CANCER-STAGE DATA

Notes: This figure shows the relationship between residualized versions of the five-year survival rate among patients diagnosed with that cancer-stage between 1973–2004 (the cohorts for which five-year survival is uncensored), and the number of clinical trials enrolling patients of each cancer-stage from 1973–2011. The level of observation is the cancer-stage. Panel A residualizes market size; panel B residualizes life-years lost. Market size denotes the inclusion of a covariate measuring the number of patients diagnosed with that cancer-stage between 1973–2009. Life-years lost is measured as age-gender-year specific life expectancy (in the absence of cancer) in the year of diagnosis, less observed survival time in years, averaged over patients diagnosed with that cancer-stage between 1973–1983 (to minimize censoring) multiplied by market size. For details on the sample, see the text and online Data Appendix.

7. [5 points] Figure 5 (reproduced below) looks at publicly-funded and privately-funded clinical trials separately. Interpret both panels of this figure, and explain how this helps address your concerns in 6.

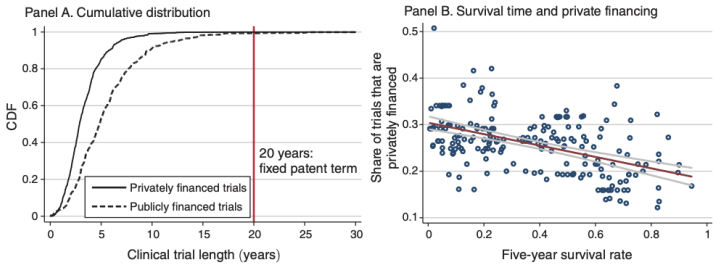

FIGURE 5. SURVIVAL TIME AND FINANCING OF CLINICAL TRIALS

Notes: This figure shows two analyses of how public and private financing of clinical trials differ. Panel A plots the cumulative distribution function of clinical trial length in years, omitting the handful of observations with length greater than 30 years for improved readability. The level of observation is the clinical trial. The vertical line at 20 years denotes the length of the fixed patent term. Panel B plots the relationship between the five-year survival rate among patients diagnosed with each cancer-stage between 1973–2004 (the cohorts for which five-year survival is uncensored), and the share of clinical trials enrolling patients of that cancer-stage from 1973–2011 that were privately financed. The level of observation is the cancer-stage. For details on the sample, see the text and online Data Appendix.

8. [6 points] In Hill and Stein (2025) "Race to the Bottom: Competition and Quality in Science" the authors employ a test of a similar spirit. Explain the test they run, and how it helps to ameliorate a similar concern in their setting.

9. [3 points] Are Budish, Roin, and Williams able to distinguish which distortion matters matters more between corporate short-termism and the fixed patent term distortion?

10. [5 points] Discuss two policy changes that might help increase welfare in the face of these two distortions.

Part 2 of 2, Handel:

Question 1

A. (20 points) Describe the central tradeoff studied in HHW and what the authors find empirically regarding this tradeoff. Describe the tradeoff in the context of the degree of risk-rating the regulator allows in a competitive market.

B. (20 points) How does the regulatory / contract structure in Ghili et al. (2023) differ from that in HHW (2015)? How does the assumption relaxed in HHW (2015) for the Ghili et al. (2023) paper impact the tradeoff you discussed in A.?

C. (20 points) For the contracts focused on in Ghili et al. (2023), how does steepness of lifetime income path impact welfare for those contracts vs. standard ACA exchange contracts with no risk rating allowed? Explain why using marginal utilities at different points in time as well as long run risk protection.

D. (20 points) The authors study model extensions that relate to (i) consumer myopia and (ii) switching costs. For each of these foundations describe (i) how contract structure

changes and (ii) how welfare of using these contracts changes. Discuss as comparative static (i.e., what happens as switching costs or myopia increase / decrease).

Question 2

A. (20 points) In the product market traps social media paper, what is the key feature of the setting that causes product market traps to be relevant relative to a typical market?

B. (20 points) Draw a figure clearly highlighting why and to what extent (hypothertically) individuals in the social media setting can have high desire to use social media but low utility from the existence of social media?

C. (20 points) Describe the three key economic steps in the experimental elicitation of preferences, designed to estimate key aspects of preferences in the market.

Question 3

A. (20 points) In the paper on inertia and adverse selection (Handel 2013), describe in detail the evidence on dominated plan choice. Then, please describe why this evidence is suggestive of inertia, but still leaves room for other explanations?

(20 points) How are risk preferences identified in the choice environment? What model for risk preferences are assumed? Describe one shortcoming of this model.