Labor Economics Field Exam

June 2017

Instructions

You have 4 hours to complete this exam.

This is a closed book examination. No written materials are allowed. You can use a calculator.

THE EXAM IS COMPOSED OF THREE QUESTIONS. EACH QUESTION IS WORTH 100 POINTS. YOU MUST OBTAIN AT LEAST 75 POINTS IN AT LEAST TWO QUESTIONS TO PASS THE LABOR FIELD EXAM.

Please answer each question in separate booklets.

First Question. 100 Points

Consider a couple composed of a wife and husband who derive utility from consumption $c_{i,t}$ and being in good health $H_{i,t}$, for $i = w, m$. The variable $H_{i,t}$, which describes the health status of spouse i in period t, evolves according to the following function:

$$\ln H_{i,t} = \gamma_1 \ln H_{i,t-1} + \gamma_2 \ln h_{i,t},$$

where $h_{i,t}$ is the investment in health, $\gamma_1 > 0$, and $\gamma_2 > 0$. The utility of spouse i takes the following form:

$$u\left(c_{i,t}, H_{i,t}\right) = \alpha \ln c_{i,t} + \ln H_{i,t}.$$

The wife and husband have identical discount factor $\beta = 1$. The husband lives for 1 period, whereas the wife lives for 2 periods and in the second period she is a widow. The wife and husband enter period 1 with the same amount of health status H_1. The couple's income is equal to Y_1 in the first period and Y_2 in the second.

The couple chooses how much income each spouse consumes $c_{i,t}$, for $i = w, m$, how much to invest in the health of the two spouses $h_{w,t}$ and $h_{m,t}$, and how much to save in a risk-free asset b_t by making decisions that are ex-ante efficient. The price of consumption is normalized to be equal to 1, the price of one unit of investment in health is p_t, and the gross return on the risk-free asset is R_t, for $t = 1, 2$.

1. (10 points) Write down the couple's problem (objective function and budget constraints).

2. (10 points) Solve the problem of the widowed wife in the second period and find the corresponding value function.

3. (10 points) Solve the problem of the couple in the first period.

4. (10 points) How are the choices of consumption, savings, and health investment affected by the fact that the wife will be a widow in the second period?

   For the rest of the question, assume that the two spouses have the same decision power.

5. (10 points) Does the household invest more in the health of the wife or husband? Prove your answer and provide the intuition.

Now suppose that both spouses live for only one period and the coefficient γ_2 varies between spouses with $\gamma_{2,w} < \gamma_{2,m}$.

6. (10 points) Solve the problem of the couple.

7. (10 points) Now, does the household invest more in the health of the wife or husband? Prove your answer and provide the intuition.

8. (10 points) Using your answers to the previous parts, provide one testable implication for the first model (the wife lives longer) and a testable implication for the second model (the wife lives for as long as the husband, but $\gamma_{2,w} < \gamma_{2,m}$).

9. (10 points) Indicate a data set and estimation method that can be used to estimate the model.

10. (10 points) Provide an informal argument explaining which variables (variation in the data) you need to identify the parameters of the model.

Labor Field Exam 2017

Graduate Labor Economics II – 261B, Prof. Till von Wachter

June 16, 2017

- Please write your Student ID on your bluebook.
- There are two questions, worth a total of 75 points. There are 7 subquestions in Question 1, each worth 5 points. There are 4 subquestion in Question 2, each worth 10 points.
- You should allocate half of your time to each question.
- Please make an effort to write down formulas where they help to clarify what you are talking about.
- Also, please be CONCISE and write LEGIBLY! *In most cases a couple sentences will suffice!*

Choose an empirical paper from the reading list of the course or the problem sets. Answer the following questions referring to that paper.

- If you do not remember details, *do not worry* - just concentrate on what you learned in class.
- Begin by writing down the paper that you will be talking out!

1 Briefly summarize the causal question, the identification strategy, and the implementation of the paper. [35 Points]

You should aim on writing no more than 1 page, at most 2 pages, on this question. To guide you, here are a few specific aspects.

1. Causal Question
   (a) What is the causal question of interest of the paper? Be specific about what the main outcome, the main treatment, and about what the counterfactual is.
   (b) Explain how the paper motivates why this causal question is an economically interesting or important question. Be brief but specific.
2. Identification
   (a) What is the main identification problem that the paper faces?
   (b) How does the paper attempt to solve the identification problem? I.e., what is the source of manipulation used, and what is the key identifying assumption (or assumptions)?
   (c) Do you think the identifying assumption(s) made in the paper are reasonable? Discuss the effect of potential failures of the assumption(s).
3. Implementation
   (a) Given the identifying assumption(s), what is the main chosen statistical method used for inference?
   (b) Summarize the implementation of the estimation procedure (e.g., What is the data they use and why? What is the sample used and why? What other choices have to be made and why? What are the choices of the paper?)

2 Answer 2 of the following 4 questions (for a total of 4 subquestions). Be as explicit as you can (equations help!). [40 Points]

1. Group-Level Error Terms

   (a) Briefly explain the problem of group-level error terms. Discuss when it is likely to have the biggest effect, and what potential solutions are (mention at least two).

   (b) Do you think the problem of group-level error terms is relevant for the paper you are discussing (or may be potentially relevant)? What does the paper actually do (if anything)?

2. Heterogeneous Treatment Effects

   (a) Briefly exlpain how heterogeneous coefficients can affect the interpretation of the coefficients in an OLS and an IV regression.

   (b) Does this paper deal with potentially heterogeneous treatment effects? How is the paper's inference affected (or how do you think it would be affected)?

3. Measurement Error

   (a) Briefly explain the effect of classical measurement error in the dependent and the independent variables on the bias of OLS for continous variables. Discuss at least two approaches to address classical measurement error in the independent variable

   (b) What are the potential measurement error problems of the paper you are discussing? Is the error likely to be classical, mean reverting, or else? How do you think the paper could address the question

4. Sample-Selection Bias

   (a) Briefly discuss the problem of sample-selection bias (it may be helpful to write down one or two equations for a specific example, say a wage regression). Suggest two solutions of sample-seletion bias that have been proposed in the literature.

   (b) Does the paper you are dealing with have a problem of sample-selection bias? If so, how does it address it? If there is a problem and it is not addressed, how could sample-selection bias affect the estimates of the paper?

Third Question. 100 Points

This question focuses on the assumptions needed to identify causal effects of the treatment choice on the outcome (treatment effects) in standard models of policy evaluation.

The general Roy Model is defined by six random variables X, Z, V, Y, U, T in probability space $(\Omega, \mathscr{F}, \mathbf{Prob})$. The causal relation among those variables is determined by unobserved autonomous (a.k.a. structural) equations described below:

	Variable Description	**Model Equations**
1	Pre-program Variables:	$X = f_X(\epsilon_X)$
2	Instrumental Variable:	$Z = f_Z(X, \epsilon_Z)$
3	Unobserved Pre-treatment Counfounder:	$V = f_V(X, \epsilon_V)$
4	Observed Outcome:	$Y = f_Y(X, T, U, \epsilon_Y)$
5	Unobserved Post-treatment Counfounder:	$U = f_U(X, V, T, \epsilon_U)$
6	Treatment Choice:	$T = f_T(X, Z, V)$

Regularity conditions commonly evoked in this model are: (1) error terms $(\epsilon_X, \epsilon_Z, \epsilon_V, \epsilon_Y, \epsilon_U)$ are statistically independent; (2) $E(|Y|) < \infty$; and (3) $\mathbf{Prob}(T = t|Z = z, X = x) > 0$ for all $t \in \text{supp}(T)$, $x \in \text{supp}(X)$ and $z \in \text{supp}(Z)$.

1. (5 points) Represent the generalized model as a Directed Acyclic Graph (DAG).

2. (5 points) Suppose a researcher is interested in the identification of treatment effects, that is, $Y(t) - Y(t')$ for $t, t' \in \text{supp}(T)$, where $Y(t)$ stands for the counterfactual outcome when T is fixed at $t \in \text{supp}(T)$. In this case the Roy model can be simplified to key random variables that suffice to investigate the identification of treatment effects. This simpler IV model is often called the *Marginalised* Roy Model. Write the equations of the Marginalized Roy Model and draw its DAG representation.

3. (10 points) Consider the Marginalized Roy model of previous item. Treatment status T and outcome Y are observed, thus $E(Y|T = t)$ can be identified through observed data. However it is well-known that the Marginalised Roy Model does not render the identification of counterfactual outcome expectation $E(Y(t))$ without further assumptions.

(a) Use the structural equations of the marginalized Roy Model to define counterfactual outcome $Y(t)$.

(b) Use the law of iterated expectations to express $E(Y(t))$ and $E(Y|T=t)$ as a weighted average of $E(Y|T=t, V=v)$.

(c) Use the equations of previous item to explain the difference between the causal operation of fixing and statistical conditioning.

4. (5 points) Key statistical properties of instrumental variable Z are: (a) exclusion restriction and (b) IV Relevance. State these two properties in the notation used for the Marginalized Roy Model.

5. (5 points) The Matching Assumption enables the identification of treatment effects. Which of the statistical relations that hold among the variables $V, Y, Y(t), Z, T, T(z)$ in the marginalized Roy Model is most related to the matching assumption?

6. (5 points) What is the main critics of the Matching Assumption?

7. (10 points) Consider the a binary choice where the matching assumption holds for a set of some observed variables X. Let the propensity score be $P(x) = \mathbf{Prob}(T=1|X=x)$ and let the density of P be $f_P(p); p \in [0,1]$. Describe a procedure the utilizes the estimated propensity scores to evaluate the Average Treatment effect $\mathbf{E}(Y(1)-Y(0))$.

8. (10 points) Consider the binary choice model where $\text{supp}(T) = \{0,1\}, T=1$ for treated and $T=0$ for control. Explain the equivalence result of Vytlacil (2002).

9. (15 points) Consider the Marginalized Roy Model with binary choice $\text{supp}(T) = \{0,1\}$ where Monotonicity holds. Let the propensity score be $P(z) = \mathbf{Prob}(T=1|Z=z)$. Suppose you estimate the liner-in-parameters equation:

$$Y = \sum_{k=0}^{3} \kappa_k P^k + \epsilon_Y.$$

How would you evaluate the marginal treatment effect using the equation estimates? How would you use the estimated marginal treatment effect to evaluate the Average Treatment effect $\mathbf{E}(Y(1)-Y(0))$?

For the remaining questions, consider your Marginalized Roy Model where:

- T is categorical and takes values in $\text{supp}(T) \in \{1, 2, \ldots, t_{N_T}\}$.
- Z is categorical and takes values in $\text{supp}(Z) = \{z_1, z_2, \ldots, z_{N_Z}\}$.

10. (10 points) What are the monotonicity conditions regarding counterfactual choice $T_\omega(z); z \in \text{supp}(Z)$ of agent ω would generate:

    (a) The Unordered Choice Model.

    (b) The Ordered Choice Model with random thresholds.

11. (10 points) What are the separability conditions governing the choice T, instrument Z and unobserved confounder V that would generate:

    (a) The Unordered Choice Model.

    (b) The Ordered Choice Model with random thresholds.

12. (10 points) Let **In** be the incentive matrix of dimension $N_Z \times N_T$ such that $\mathbf{In}(z, t)$ denotes the incentive of agent ω to choose choice $t \in \text{supp}(T)$ when the instrument Z takes value $z \in \text{supp}(Z)$. Assume that the Weak Axiom of Revealed Preferences and Normal Choices hold. State the conditions on **In** that generate:

    (a) The Unordered Choice Model.

    (b) The Ordered Choice Model with random thresholds.