Labor Economics Field Exam

June 2016

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 an economy in which each individual lives from period $t = 1$ to period $t = T$ and has preferences over consumption and leisure that can be represented using the following utility function:

$$u(c_t, l_t) = \alpha (c_t)^{\delta} (l_t)^{\gamma}$$

with α, δ and $\gamma > 0$ and $\delta + \gamma < 1$.

Each individual is endowed at the beginning of his/her life with assets b_1 and, in each period, with 1 unit of time that can be divided between leisure l_t and hours of work h_t. In each period, she or he has to decide how much to consume, how much time to spend on leisure and hours of work, and how much to save on a risk-free asset with gross return R_t $(= 1 + r_t)$. There is no uncertainty, the risk-free asset is the only asset in the economy, and, if an individual choose to supply labor, she or he receives a wage $w_{i,t}$ for each hour worked.

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

2. (10 points) Derive the Marshallian leisure demand functions.

3. (10 points) Derive the Frish (λ-constant) leisure demand functions.

4. (10 points) Derive the labor supply elasticities for the two leisure demand functions.

5. (10 points) Provide an intuition for why the Marshallian elasticity takes that form. Do the same for the Frish elasticity.

   Add error terms to the Marshallian and Frish leisure demand functions in the form of measurement errors.

6. (10 points) Indicate a data set and estimation method that can be used to estimate the Marshallian leisure demand functions.

7. (10 points) Do the same for the Frish leisure demand functions.

   Now suppose there is an aggregate shock that affects the wage rate of all individuals in the same way. Specifically, the wage paid to the workers $w_{i,t}$ is the hourly wage $\bar{w}_{i,t}$ plus a per-hour bonus θ_t that is identical for all workers, i.e. $w_{i,t} = \bar{w}_{i,t} + \theta_t$, with $E[\bar{w}_{i,t}] = E[\bar{w}_{i,t+1}]$ for every t, with $\theta_1 = 0$, where the expectation is taken with respect to the cross-section of individuals in the economy in a given period. In each period, you observe the hourly wage $\bar{w}_{i,t}$, but not the bonus θ_t. Consider the Frish leisure demand functions you derived earlier and simply replace the new definition of the wage in it, i.e. you do not have to re-derive the Frish leisure demand functions with the aggregate shock.

8. (10 points) Suppose you want to estimate the leisure Frish elasticity with respect to the hourly wage $\bar{w}_{i,t}$ and you only observe data for $t = \tau, \ldots, T$, with $\tau > 1$. Does the data set and estimation method you have proposed earlier allow you to recover the leisure Frish elasticity when aggregate shocks are present? If yes, explain why. If not, propose an alternative data set and estimation method.

9. (10 points) How is the Frish elasticity affected by the aggregate shock? Provide an intuition for your answer.

   Now suppose that the aggregate shock evolves according to the following equation: $\theta_t = \rho\theta_{t-1} + \epsilon_t$, where ϵ_t is identically and independently distributed over time with mean 0.

10. (10 points) Propose a method to estimate ρ.

Second Question. 100 Points

There are two questions with roughly equal weight. Please write legibly. Please make an effort to write down a formula where it helps to clarify what you are talking about.

1. Instrumental Variable Estimates of the Returns to Schooling

   Consider the following cross-sectional model for individual earnings:

$$log\ y_i = \alpha_i + \beta_i S_i + \epsilon_i \tag{1}$$

   where y_i and S_i are log earnings and years of schooling of individual i, respectively. α_i is an individual constant that may be correlated with schooling, and β_i is the return to schooling, which is allowed to vary across individuals. Let $B = E(\beta_i)$ be the average return to education in the population.

   (a) Under what conditions will the least squares estimate of the coefficient on schooling in model (1), β_{OLS}, be a consistent estimate of the population average return to education ? What are the sources of omitted variable bias (a.k.a., ability bias) and selectivity bias (a.k.a., self-selection bias)? For 2 individuals, i and j, with different abilities (a_i, a_j), marginal returns to education (b_i, b_j), and marginal costs of education (r_i, r_j), graphically depict the education selection process and how it relates to the least squares coefficient, β_{OLS} (hint: put (log y) on the y-axis and S on the x-axis and exploit the explicit functions for returns and costs we had assumed in class).

   (b) Suppose S_i is measured with error that is classically distributed (consider the case in which $\alpha_i = \alpha$ and $\beta_i = \beta$ for all i). How will this measurement error bias the least squares estimate of the returns to education, and how is this bias related to the noise-to-total variance ratio corresponding to S_i? Explain how throwing in additional control variables (X_i) into the regression may exacerbate the measurement error bias.

   (c) Suppose there exists a variable, z_i, that differentially affects the costs of schooling across individuals for exogenous reasons (does not have an independent effect on the benefits of schooling):

$$S_i = \theta z_i + v_i \tag{2}$$

Under what conditions will two-stage least squares (2SLS) estimation of equations (2) and (1) yield a consistent estimate of the average return to education ? Under these conditions, explain how 2SLS accounts for both the ability and the self-selection biases (one way to show this is to use a control function formulation). Suppose the assumptions do not hold; what parameter(s) might 2SLS estimate and under what conditions? Give the parameter an economic interpretation.

(d) Now explain the Garen control function approach to accounting for ability and self-selection biases. Under what conditions does it work? Show how it parametrizes the 2 sources of bias.

(e) Show that IV can have a worse omitted variable bias problem than OLS when the instrument has a weak relationship to schooling. Even without a remaining bias, IV suffers from small sample bias problems even in large sample if the first stage relationship is weak. What solution has been proposed in the literature? Does that solve the first problem as well?

2. Twins and the Return to Computer Use

An analyst estimates the following model for the log of wages (w) for workers using data from a recent Current Population Survey:

$$log\ w = \alpha + \beta S + \gamma Exp + \delta Exp2 + \theta Male + \lambda UseComputer + \epsilon \tag{3}$$

where (α, β, γ, δ, θ, λ) are coefficients, S is years of schooling, Exp represents potential labor market experience, Male is a dummy variable equal to 1 for male workers and 0 for females, and $UseComputer$ is a dummy variable equal to 1 for workers who state that they use a computer on the job.

(a) Another approach to address omitted variable bias has been the use of data on twins. There are two approaches to use twin data differencing and correlated random effects. Briefly explain them both.

(b) Show how you would use twin data to estimate the returns to computer use. Are you worried about measurement error in this context? How does first differencing affect the influence of classical measurement error?

(c) What does a twin-estimator implicitly assume about the choice of computer use within the family? Given this assumption, when is a within family estimator indeed better than a between family estimators? How could you try to assess this assumption using data on observable characteristics (you can follow the suggestion by Ashenfelter and Rouse)?

(d) One way to use twin information is to include the average propensity of computer use among both twins as separate regressor in the individual model for earnings. How is this similar to the control function approach?

(e) In class, we have shown how the twin estimator is a special case of a matching estimator. Write down the general matching estimator of the effect of computer use on wages. Define the propensity score. What is its role in matching estimators and what justifies its use?

Third Question. 100 Points

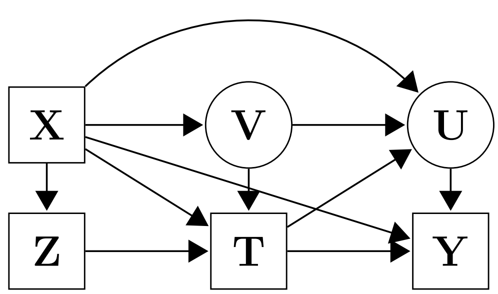
This question focuses on the assumptions needed to identify causal effects of the treatment choice on the outcome (treatment effects) in standard models of policy evaluation.

Let a General Roy Model be defined by the following six random variables defined in probability space $(\Omega, \mathscr{F}, \mathbf{Prob})$:

	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:	$T = f_T(X, Z, V)$

Some regularity conditions 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 $z \in \text{supp}(Z)$. The General Roy Model can be conveniently represented by a Directed Acyclic Graph (DAG):

Figure 1: **DAG Representation of the General Roy Model**



Where arrows denote causal relations, circles denote unobserved variables and squares denote observed variables.

1. (10 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 $Y(t) = f_Y(X, t, U, \epsilon_Y)$. 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.

2. (10 points) The Marginalised Roy Model does not render the identification of treatment effects without further assumptions. Consider the binary choice model where $\text{supp}(T) = \{0, 1\}, T = 1$ for treated and $T = 0$ for control. A key assumption that allows for the identification of binary treatment effect is the separability condition. State the separability condition using the variables of the marginalized Roy Model.

3. (10 points) In the binary choice model, where $\text{supp}(T) = \{0, 1\}$, let T_ω such that $\omega \in \Omega$ denotes a measurement ω of random variable T. And let $T_\omega(z)$ be the counterfactual choice of agent ω when the instrument Z is set at $z \in \text{supp}(Z)$. Use this notation to *equivalently* state the separability condition in terms of counterfactual choices of the agents.

4. (5 points) Consider your Marginalized Roy Model. State the key statistical property of instrumental variables that (in addition to separability) allows to identify treatment effects.

5. (5 points) State the Matching Assumption, that is, the statistical property of observed pre-program variables that is used to identify of treatment effects in choice models where matching holds.

6. (10 points) Consider your Marginalised Roy Model. Which of the statistical relations that hold among the variables $V, Y, Y(t), Z, T, T(z)$ is most related to the matching assumption?

7. (10 points) Consider the General Roy Model of figure 1. Which causal links would need to be suppressed for the matching assumption to hold? In the marginalised Roy Model, which causal link should be *modified* (not suppressed) for the matching assumption to hold?

8. (15 points) Consider the Marginalized Roy Model with binary choice. Let the propensity score be $P = \mathbf{Prob}(T = 1|Z)$ and let the density of P be $f_P(p); p \in [0, 1]$. 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 $\mathbf{E}(Y(1) - Y(0))$?

9. (15 points) Consider the a binary choice where the matching assumption holds. Let the propensity score be $P = \mathbf{Prob}(T = 1|X)$ and let the density of P be $f_P(p); p \in [0, 1]$. Suppose you estimate the liner-in-parameters equation:

$$Y = \sum_{k=0}^{3} \cdot\zeta_k^1 T \cdot P^k + \sum_{k=0}^{3} \zeta_k^0 \cdot (1 - T) \cdot P^k + \epsilon_Y.$$

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

10. (10 points) 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 $\text{supp}(Z) = \{z_1, z_2, \ldots, z_{N_Z}\}$.

What condition on the counterfactual choice $T_\omega(z); z \in \text{supp}(Z)$ of agent ω would generate a ordered discrete-choice model with random thresholds?