

PRELIMINARY EXAMINATION FOR THE Ph.D. DEGREE

---

---

- Please answer all three (3) of the following questions.
- Put your identifying number on your exam.

1) Public preschools provide a 100% price subsidy for child care for the first  $h$  hours of work per day, where  $h$  is number of hours in the school day. In theory, access to public preschool therefore makes it easier for mothers to enter the labor force. To see whether this prediction is borne out in practice, a researcher wants to estimate the relationship between public preschool availability and maternal labor force participation. The model of interest is

$$(1) \quad y_i = \theta T_i + f(X_i) + \varepsilon_i$$

where  $y_i$  is an indicator variable equal to 1 if mother  $i$  is in the labor force, 0 otherwise;  $T_i$  is an indicator variable equal to 1 if there are public preschools in mother  $i$ 's area of residence, 0 otherwise; and  $f(\cdot)$  is some potentially non-linear function of mother  $i$ 's observable characteristics,  $X_i$ .  $\varepsilon_i$  is an error term which is potentially correlated with  $T_i$ . For simplicity, assume that the sample is limited to women with children of preschool age. Also assume that each mother is observed only once.

- a) One approach to identifying  $\theta$  is matching on observables. Which assumption(s) would need to be satisfied for the matching estimator of  $\theta$  to be identified? Which of the following – multivariate matching or propensity score matching – is more appropriate for this application? Justify your answer.
- b) An alternative identification strategy is to exploit variation over time in state initiatives that subsidize preschool provision by public schools. The researcher re-expresses model (1) and writes down a “first-stage” equation for  $T$ :

$$(2) \quad y_{ist} = \theta T_{ist} + X'_{ist} \beta + \lambda_t + \gamma_s + \varepsilon_{ist}$$

$$(3) \quad T_{ist} = \pi_1 P_{st} + X'_{ist} \Pi_2 + \lambda'_t + \gamma'_s + v_{ist}$$

where  $s$  denotes state and  $t$  denotes year;  $\lambda_t$ ,  $\gamma_s$ ,  $\lambda'_t$ , and  $\gamma'_s$  are fixed effects; and  $P_{st}$  is equal to 1 if state  $s$  subsidizes preschool provision in year  $t$ , 0 otherwise. For the two-stage least squares (TSLS) estimator of  $\theta$  to be identified, it must be the case that  $\pi_1 \neq 0$ . What must also be true for the TSLS estimator of  $\theta$  to be identified? Describe two informal ways in which the researcher might get some sense of whether this second assumption is satisfied.

- c) Bertrand, Duflo, and Mullainathan (2004) criticize the standard implementation of models such as (3). Describe their critique. What are several suggested solutions to the problem that they discuss?
- d) Suppose that there are heterogeneous treatment effects, i.e., the true model is

$$y_{ist} = \theta_i T_{ist} + X'_{ist} \beta + \lambda_t + \gamma_s + \varepsilon_{ist}$$

What is the appropriate interpretation of the TSLS estimator? Explain.

- e) Assume once again that there is a homogeneous treatment effect. For simplicity, also assume that there is no selection on observable or unobservable characteristics. Will misclassification in  $y_i$  bias estimates of  $\theta$ ? Formally justify your answer.

- 2) *Richland* is a first world country that provides public assistance for needy families through a program called CASHAID. To be eligible for aid, the family must have income below a maximum income threshold. Further, the transfer program is only available to families headed by single women. The benefits are structured like a traditional income support program: the family receives a guaranteed income of \$B per month and for every dollar earned the benefit is reduced by \$0.50.
- a) A vocal critic of the program claims that the program is inefficient because it discourages work. Present an economic model that substantiates this claim. Summarize the best evidence from the literature covered in class that speaks to this claim. (When discussing the literature, make sure you provide a concise discussion of the research design and identification strategy employed in the paper.)
  - b) This same critic claims that there a second source of inefficiency in the program in that it discourages marriage. Present an economic model that substantiates this claim. Summarize the best evidence from the literature covered in class that speaks to this claim. (When discussing the literature, make sure you provide a concise discussion of the research design and identification strategy employed in the paper.)
  - c) In response to these concerns, a reform of CASHAID is implemented in *Richland*. The categorical eligibility is removed and now the program is extended to all (income eligible) families with children (married couples and single parent families). Use your economic model above to discuss the impacts of this reform on labor supply. Discuss impacts on labor force participation and hours worked.
  - d) You are hired to do the evaluation of the reform. Suppose the reform is implemented at the same time across all areas. How would you evaluate the reform? Present the research design and discuss the assumptions under which this is a valid estimator. How would you explore/test the validity of these assumptions?
  - e) Suppose you find out that the reform was implemented at difference times across the states. How would you modify your evaluation of the impacts of reform? Present the research design and discuss the assumptions under which this is a valid estimator. How would you explore/test the validity of these assumptions? Contrast this estimator to the one used in (d) above.

**Question 3. Optimal Taxation and Minimizing Alternative Measures of Excess Burden**

In a one-consumer economy with fixed producer prices, consumer prices  $\{q_i\}$  differ from producer prices  $\{p_i\}$  by the amount of the tax on commodities  $2, \dots, n$ :

$$t_i = q_i - p_i \quad i=2, \dots, n.$$

Income and cross-price effects are *not* assumed to be zero.

- a. Set up, but do not solve, the optimal tax problem of maximizing utility subject to raising a required amount of revenue  $R_0$ .

We will now reinterpret this problem using the concept of *equivalent variation*.

- b. Use the equivalent variation evaluation of the loss in consumer welfare from commodity taxation and the amount of tax revenue actually raised to define the deadweight loss resulting from the tax vector  $\{t_i\}$ .
- c. Show that the problem you set up in part (a) is equivalent to, and therefore has the same solution as, the problem of minimizing deadweight loss subject to raising the required revenue.
- d. Using  $\mu$  as the Lagrange multiplier for the problem in *part (b)*, find the first order condition for the problem of minimizing deadweight loss subject to a revenue constraint. Give an intuitive explanation of the FOC, including a clear discussion of the meaning of  $\mu$  which tells why this multiplier is different from the multiplier  $\lambda$  you would have derived from the more familiar formulation in part (a).

Finally, let us use the concept of *compensating variation* to approach this problem.

- e. Now use instead the compensating variation evaluation of the loss in consumer welfare from commodity taxation and the amount of tax revenue actually raised to define the deadweight loss resulting from the tax vector  $\{t_i\}$ .
- f. Is the problem you set up in part (a) equivalent to the problem of minimizing this new concept of deadweight loss subject to raising the required revenue? If it is, show that it is. If it is not, give an intuitive explanation of why it is not.