Problem Set 3: Targeting, IV, DD

ECON 24450: Inequality and the Social Safety Net (Deshpande)

Due Thursday, May 13, 9am

1. **Screening theory.** Consider an economy with two types of individuals: those who are able (a) and those who are not able (d). Let u_w^a (u_w^d) denote the average utility of working for the able (not-able) type, and u_b^a (u_b^d) denote the utility of receiving benefits for the able (not-able) type. Let ϵ represent the idiosyncratic preference for working, with ϵ following a uniform distribution with range $[\underline{\epsilon}, \bar{\epsilon}]$ and density function $f(\epsilon) = \frac{1}{\bar{\epsilon}-\underline{\epsilon}}$. The utility of working for type $i \in \{a, d\}$ is given by $u_w^i + \epsilon$ and the utility of receiving benefits is $u_b^i - c^i$, where c^i denotes the cost to type i of applying for benefits.

   (a) Write down the condition for type i to apply for benefits.

   (b) Derive an expression for the mass of type i who apply for benefits.

   (c) Consider an ordeal that increases the application cost for type i from c^i to $c^i + \Delta c^i$. Derive an expression of the change in the mass of type i who apply for benefits.

   (d) Consider the following two cases. For each case, does the increase in ordeals improve or worsen targeting? What types of application costs are relevant to each case and why?

      i. $\Delta c^d < \Delta c^a$

      ii. $\Delta c^d > \Delta c^a$

   (e) What does this exercise tell us about the Nichols and Zeckhauser (1982) hypothesis? **opposite effect**

2. **Screening; instrumental variables.** Enrollment in social safety net programs often involves "ordeals" like waiting in line and completing complicated paperwork. Consider the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) program, which provides nutrition assistance to pregnant women, infants, and children.

   (a) Suppose a state would like WIC benefits to go to type L mothers, but not type H mothers. The state cannot tell a mother's type so it cannot tag. The state wants to distinguish between the two types of mothers by imposing documentation requirements for the program. A mother of type i receives utility $v_i > 0$ from enrolling in WIC. Time spent gathering and providing documentation is wasted, so a mother of type i who spends t hours providing documentation loses utility $c_i t$, where $c_i > 0$. Thus the total utility of a person of type i who spends t hours on documentation and receives WIC is $v_i - c_i t$. A mother who spends zero hours on documentation and receives no WIC gets a utility of zero. How should the state choose documentation requirements t^* such that *only* type L mothers choose to enroll in WIC? This is called a "separating equilibrium". What is the condition on v_L, v_H, c_L, and c_H such that a separating equilibrium could possibly exist? Interpret the economic meaning of this condition.

   (b) Suppose the government now wants to use the information from part (a) to target other social programs. That is, the state will treat any mother who has provided documentation and enrolled in WIC as type L with respect to future social programs. Denote the utility gain from these future social programs as $x > 0$. What will be the impact on the range of waiting times that can separate type L and type H citizens if the future programs benefit only mothers who are believed to be type L (i.e., mothers who received WIC). Explain the economic intuition behind your answer.

   (c) Suppose now that the state wants WIC to go to type L mothers, but it will allow type H mothers to buy into WIC at a price p (which you can assume is also the utility cost). It offers mothers a choice: she can provide documentation, which takes time $t > 0$, or buy into the program at price

p. Mothers who do neither will not get WIC, and no mother can enroll in WIC twice. The state wants to set p^* and t^* such that type L mothers choose to provide documentation and enroll, and type H mothers choose to buy in. This requires four conditions. Write each one:

i. Type L mothers choose documentation over buy-in (this is type L's "incentive compatibility" constraint).

ii. Type L mothers choose to enroll (this is type L's "individual rationality" constraint).

iii. Type H mothers choose buy-in over documentation (this is type H's "incentive compatibility" constraint).

iv. Type H mothers choose to enroll (this is type H's "individual rationality" constraint).

v. Show that your conditions produce the following result: $c_L t^* \leq p^* \leq min\{v_H, c_H t^*\}$, $t^* \leq \frac{v_L}{c_L}$

vi. Explain more generally what the incentive compatibility constraint is and why it is important in tagging and screening.

(d) Explain the role of the envelope theorem in the theoretical justification behind ordeals, in-kind transfers, and restrictions on earned income. Draw a picture to help explain. What empirical estimates would we need in order to decide whether the documentation ordeal improves social welfare? What is a situation in which the documentation ordeal could decrease social welfare?

(e) Figlio, Hamersma, and Roth (2009) use an IV strategy to evaluate the effect of prenatal WIC participation on birth outcomes in Florida. Their instrument is income reporting (documentation) requirements that affect one group of women but not another group, since documentation requirements in other settings have been shown to reduce take-up.

i. Write down the structural equation of interest, first stage equation, and reduced form equation.

ii. What are the assumptions required for the exclusion restriction and monotonicity to hold in this setting? How could you test these assumptions?

iii. Characterize the four groups created by IV (always takers, never takers, compliers, and defiers) in terms of how the instrument affects their treatment status. What are the likely characteristics of each group (e.g., in the MMS paper, we said never takers were probably healthy)?

iv. Figlio et al. (2009) estimate that WIC participation reduces the likelihood of adverse birth outcomes (like birth weights below 2500g). Describe your confidence in using this estimate to extrapolate the effects of the following policy changes on birth weight:

A. Reducing documentation requirements for WIC

B. Eliminating cognitively challenging paperwork from the WIC application process

C. Automatically enrolling all income-eligible pregnant women in WIC

v. Figlio et al. (2009) study only effects on take-up of WIC. What would you advise them to do to study effects on the targeting of WIC?

3. **Disability as tag; difference-in-differences.** In most developed countries, disability is used as a tag to target resources.

(a) Describe advantages and disadvantages of categorical welfare (i.e., tagging) relative to traditional cash welfare programs. What rules should guide the government as it decides what the appropriate categories are in designing these programs; that is, which kind of tags should it use (e.g., single motherhood, homelessness, disability)?

(b) Work through the Akerlof (1978) model from class and explain how tagging helps address the "iron triangle" problem.

(c) Gruber (2000) estimates the effect of disability insurance benefits on labor supply using a difference-in-differences model. In particular, he takes advantage of the fact that Quebec has a distinct disability insurance program from the rest of Canada. In January 1987, the rest of Canada raised its benefits by 36% while benefits in Quebec remained the same, as shown in this figure:

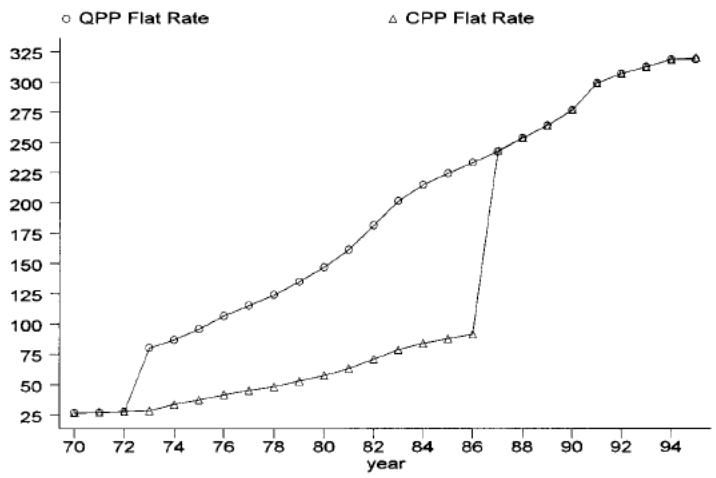

FIG. 1.—Flat-rate portion in Quebec and the rest of Canada

He estimates an elasticity of labor force nonparticipation with respect to benefit levels of 0.28-0.36 (i.e., an increase of 10% DI benefit levels increases labor force nonparticipation by 2.8-3.6%).

i. Write the key identifying assumption for Gruber's DD strategy to identify the effect of disability insurance on labor force participation. What are two possible scenarios in which this assumption would be violated?

same as 4 treatment * after + controls is pre post

ii. Write down the pre-post version of the DD estimating equation. Does this equation allow you to test the key identifying assumption? If yes, state exactly how you would use the regression coefficients to test this assumption. If no, explain why not.

iii part e of problem 4

iii. Write down the year-by-year version of the DD estimating equation. Does this equation allow you to test the key identify assumption? If yes, state exactly how you would use the regression coefficients to test this assumption. If no, explain why not.

iv. What does this paper (and other papers like it) tell us about the quality of disability as a tag? What other information do we need to know to make this assessment?

4. **Data exercise: difference-in-differences.** We studied the earned income tax credit in our section on labor supply, but I also mentioned several papers that estimate the effects of the EITC on other outcomes, like children's educational achivement and health. This data exercise is based on the paper "Income, the Earned Income Tax Credit, and Infant Health" (Hoynes, Miller, and Simon 2015), which you should read before starting this question.

   (a) What is the causal relationship of interest in this paper? Write down the structural equation. Note: this equation does not appear in the paper.

   (b) What is the problem with estimating this equation in cross-sectional data? In what direction is the estimate likely to be biased?

   (c) Explain the authors' proposed difference-in-differences strategy.

      i. What are the two differences?

      ii. What is the key identifying assumption?

      iii. Describe a plausible story that could violate the DD identifying assumption in this context.

   (d) Write down the "reduced form" difference-in-differences equation. (This does appear in the paper.) Explain each term of the equation, how it relates to the DD strategy, and the coefficient of interest.

   (e) How would you alter the equation above to test the parallel trends assumption? Write down that alternative specification.

(f) Download the data set from Canvas. Import the data set to your preferred statistical analysis program (e.g., Stata, R). For each of the regressions below, follow these guidelines:

- Restrict the sample to single women with a high school education or less, and to effective tax years 1991 through 1998
- Weight your regressions using cellnum (the number of births in the state-year-parity-demographic cell)
- Cluster standard errors by state.
- Let outcome variable Y_{pjst} be the fraction low birth weight multiplied by 100.

(g) Estimate pre/post DD regressions (replicate even columns of Table 2 in paper):

i. Consider model (1) from the paper:

$$Y_{pjst} = \alpha + \delta After_t \times Parity2plus_p + \beta X_{st} + \gamma_p + \eta_s + \delta_t + \phi_j + \epsilon_{pjst}$$

where Y_{pjst} is the fraction of low birth weight infants multiplied by 100 for the cell defined by parity p, demographic group j, state s and effective tax year t. γ_p is a set of dummy variables for birth order, η_s is a set of dummy variables for state of residence, and δ_t is a set of dummy variables for effective tax year (we do not provide these dummy variables in the data set – you will need to construct them). In the data set that we provide, the terms ϕ_j and X_{st} correspond to the following variables: *hispanic*, *black*, *other*, *hispanicmiss*, *racemiss*, *age2*, *age3*, *high*, *reform*, *a_urate_st*, *threshpreg* and all variables beginning with "*fe_I*".

A. Construct a dummy variable $After_t$ that equals one for effective tax years 1994 through 1998, and 0 otherwise.

B. Construct a dummy variable $Parity2plus_p$ that equals one for second- or higher birth orders, and 0 otherwise.

C. Estimate model (1) by OLS.

ii. Consider model (1'):

$$Y_{pjst} = \alpha + \delta_1 After_t \times Parity2_p + \delta_2 After_t \times Parity3plus_p + \beta X_{st} + \gamma_p + \eta_s + \delta_t + \phi_j + \epsilon_{pjst}$$

Model (1') decomposes the policy impact into second births and third- or higher order births.

A. Estimate model (1') by OLS.

B. Explain why it makes intuitive sense that the policy impact is larger for third- or higher order births than for second order births.

iii. Limit the sample to the set of second- or higher order births. Consider model (1"):

$$Y_{pjst} = \alpha + \delta After_t \times Parity3plus_p + \beta X_{st} + \gamma_p + \eta_s + \delta_t + \phi_j + \epsilon_{pjst}$$

A. Estimate model (1") by OLS.

B. In model (1"), who is the treatment group? Who is the control group?

iv. Construct a table that contains the estimated coefficients on the interaction terms in models (1), (1') and (1"). Report standard errors in parentheses below each estimate. In your table, indicate whether each estimated coefficient is significant at a 1%, 5% or 10% level.

(h) Estimate year-by-year DD regressions (replicate Figure 3-A in paper):

i. Estimate the "year-by-year" version of model (1) by replacing the interaction term $After_t \times Parity2plus_p$ with a full set of effective tax year dummies interacted with $Parity2plus_p$. Normalize the coefficient on the interaction term to 0 in 1993 (by omitting this term from your regression equation). Plot the estimated coefficients on the interaction terms by effective tax year.

ii. Does your plot support the parallel trends assumption for model (1)? Explain.