Problem Set 2, Economics 249, Professor Bitler

Due date: Wednesday, June 4

Problem 1. Means and heterogeneity across the distribution

In this problem, you will examine data from two waiver experiments, Connecticut's Jobs First (CJF) and Minnesota's Family Investment Program (MFIP). We discussed Bitler, Gelbach, and Hoynes (AER, 2006), which used the CJF data to look at heterogeneity in the effects of Jobs First across the distribution and whether it was consistent with predictions of a simple static labor supply model. To learn more about the two experiments, you can look at the final reports for Minnesota (Turning Welfare into a Work Support), and Connecticut (Jobs First Final Report on Connecticut's Welfare Reform Initiative). Here you will replicate some of the findings in our paper, and use some of the same methods to look at the distributional effects of reform in Minnesota.

These data are restricted use. I have obtained permission from MDRC for you to use them solely for the purposes of this assignment. You will need to first sign a form agreeing to follow the conditions set by MDRC, and then I will email you a PGP encrypted version of the CJF and MFIP data sets. Please delete the raw data files when you are finished with the homework, and send me an email confirming that you have done so. If you wish to use this data for other purposes, you must submit an application to MDRC.

First you will explore the effects of the CJF in parts a–e. Then you will look at heterogeneity in the MFIP program in parts f–h.

a. Describe the main differences between Jobs First and AFDC. How would these affect labor supply and welfare participation?

b. Create a table of means for all of the variables. It should contain means for the treatment group, means for the control group, and their difference (treatment minus control). The variable "treatmnt" is 1 if assigned to Jobs First and zero otherwise. Note where the individual differences are statistically

significant at the 5 percent level. (Hint: You can do this as a regression or with a t-test.) What does this comparison suggest for the mean impacts of the program on average earnings and transfers for the first 7 quarters?

c. Do a joint test that the following pre-RA means are jointly significant: recpient white black hisp marnvr marapt nohsged hsged mthsgrad kidctgt2 agelt25 age2534 agegt34 ernpqa18 adcpqa17 fstpqa17 anyernpqa18 anyadcpqa17 anyfstpqa17 missnohsged misshsged misskidctgt2 missmarnvr missmarapt. (Hint: You can do this with seemingly unrelated regression, but you will need to leave out one of each set of perfectly collinear variables. So, for example, agelt25, age2545, and agegt34 are perfectly collinear, so you would need to leave one of them out.)

d. Now do the same joint test but weight it using the inverse propensity score weight (variable name: pscorewt). Does the propensity score weight balance the sample across groups in the pre-RA period? What are the weighted mean impacts on earnings and transfers?

e. Compute the deciles of earnings and transfers for the treatment and control groups without using the propensity score weights for quarters 1–7 (variable names: ernqa1t7 and trans1t7). Use these to compute the QTEs for quantiles 10, 20, 30, 40, 50, 60, 70, 80, and 90. Do the same, but use the propensity score weights (variable name: pscorewt). Create two graphs, one with both sets of QTEs for earnings and one with both sets of QTEs for transfers. Do the weighted and unweighted estimates tell the same story? How are they different from the quarterly findings in the published paper? Why might the amplitude of the QTEs be smaller (in magnitude) using average earnings than using quarterly earnings?

Now you will look for heterogeneity in MFIP.

f. Describe the main differences between the Family Investment Program and AFDC. There are two programs tested in MFIP, full-MFIP, and incentives-only MFIP (which is carrots but no sticks). How

would these affect labor supply and welfare participation?

g. Do a joint test for whether the pre-RA variables (listed below) are significantly different across the two programs being evaluated. These are full MFIP (treatment variable name: prog1, 1 if in full-MFIP, 0 if in control group, missing if in incentives-only MFIP) and incentives-only MFIP (treatment variable name: prog2, 1 if in incentives-only, 0 if in control group, missing if in full-MFIP).

Variables to test: black white asian age gender marital degree urbanrur pearn8 afcpm23 fspm23 prog1 prog2.

h. First calculate the mean treatment effects for both programs. Compute the QTEs of earnings and transfers for both programs (variable names: earnuse and transavg, note that these are annual, and must be converted to quarterly to be comparable to CJF). Create two graphs, one with QTEs for earnings and one for transfers. Do you see evidence of heterogeneity consistent with a simple static model? What do the differences between full MFIP and incentives-only MFIP tell you about carrots and sticks in welfare reform?

Problem 2.

Here we will replicate some of the findings in Card, Dobkin, and Maestas (forthcoming, AER). I will email you a compressed version of an exceprt of the NHIS data they use as well.

First, rescale the age in quarters variable to be 0 at age 65, and also create the square of this rescaled variable. Create a dummy for being at least 65. Interact this dummy with both the new age in quarters variable and its square. Finally, create 3 new variables: the average insurance coverage for each age in quarters, and the same average for males and females separately for each age in quarters. (Hint: You can use egen, but don't forget to only use one observation per age or age and gender group in the regressions.)

a. Now we will replicate part of NBER WP 10365, Table 3 (rows 1, 9, and 10 of columns 1–3). For each specification, include the "I am 65" dummy, and the age in quarters quadratic and their interactions with the "I am 65" dummy, and cluster the SEs by age in quarters. We will explore the effect of turning 65 on insurance coverage (variable: insured). Column 1 uses the age averages. Column 2 uses the micro data. Column 3 uses the micro data and also controls for gender, race/ethnicity, year dummies, and education dummies. Can you get close to Table 3?

b. Suppose instead you run a model comparing those over 65 with those under 65 (a differences model), with the gender/education/race/ethnicity/education/year of survey controls. How different would your conclusion have been? What does this suggest about the effect at 65?

c. Run the specification of column 3 with the micro data, but with the outcomes hospitalized in the last 12 months (hosp_12m) and saw a doctor in the last 2 weeks (md_2wks). What do you conclude about whether the change in insurance status has effects on health?

c. Choose a counterfactual age at which to look for discontinuous changes in coverage. Recreate the RD estimates with controls for insurance coverage, hospitalization, and seeing a doctor in the last 2 weeks. What age did you pick and why? What do your results tell you about the Card, Dobkin, and Maestas findings?