Alex Miller

Problem Set 3

1.

In “Minimum Wages and Employment” Card and Kreuger were attempting to quantify the effect that an increase in a state minimum wage would have on low-income industries. The specific policy selected for the study was New Jersey’s relatively rapid increase of their state minimum wage from $4.25 an hour to $5.05, making it the highest in the nation at the time. To measure the economic impact of the policy, Card and Kreuger interviewed restaurants and recorded their employment numbers, starting wages, meal prices, and perks such as free or reduced meals.

The increase to New Jersey’s minimum wage was criticized as harmful to employers and employees based on theoretical foundations, but studying the real-world effects of the policy would be more difficult because the policy change did not occur in a vacuum; in fact, a small recession hit in the middle of Card and Kreuger’s study, making direct comparisons pre- and post-policy inherently biased. In order to compensate for any economic changes to the US as a whole, Card and Kreuger also interviewed restaurants across state lines in Pennsylvania, where no minimum wage shock occurred. To control for economic changes solely within New Jersey, they made sure to interview restaurants that were already paying above minimum wage and would not be affected by the policy shock.

2.

In defining a new model based on this paper’s data, Card and Kreuger’s “full-time equivalence” or FTE measure (full time employees counted as full with part-time employees counted as half) was chosen as the dependent variable. The FTE was defined as function of both the supply and demand for labor. For the supply of labor, the most relevant variable was the starting wage. In order to quantify the effect of the minimum wage policy, the starting wage was split into two parts: the portion of the starting wage required by state law, and the portion of the starting wage paid above and beyond the minimum. For the demand of labor, several proxy variables for the size of the restaurant were used along with dummy variables for each chain to account for exogenous variation in demand for labor by certain types of cooking (e.g. KFC, with just a fryer, might require fewer workers than a Burger King, with a fryer and a grill). The final model specification was as follows:

*(FTE)= β0 + β1\*(minimum wage) + β2\*(starting wages above minimum) + β3\*(hours open) + β4\*(number of registers) + β5-7\*(restaurant chain dummy variables)*

If our estimation of β1 results in a large, positive, significant coefficient, then we can infer that the minimum wage has a big positive effect on employment and any policy that increases minimum wage will also increase employment. On the other hand, a negative significant value would tell us that increases in the minimum wage would restrict employment as predicted by the theoretical models.

Our relevant variables were calculated with the following commands:

. quietly tab chain, gen(chaindum)

. gen minwage = 4.25

. replace minwage = 5.05 if state==1 & year==2

. gen abovemin = wage\_st-minwage

. gen fte = empft+(emppt\*0.5)

A pooled OLS regression was estimated and the elasticities of the coefficients were calculated with the following commands. The results also follow:

. reg fte minwage abovemin hrsopen nregs chaindum2-chaindum4, robust

. mfx, dyex

**Linear regression Number of obs = 752**

**F( 7, 744) = 62.16**

**Prob > F = 0.0000**

**R-squared = 0.3451**

**Root MSE = 7.3007**

**----------------------------------------------------------------**

**| Robust**

**fte | Coef. Std. Err. t P>|t|**

**-------------+--------------------------------------------------**

**minwage | 1.19534 .7485032 1.60 0.111**

**abovemin | 1.919192 1.005132 1.91 0.057**

**hrsopen | 1.819716 .222026 8.20 0.000**

**nregs | 1.020154 .3310285 3.08 0.002**

**chaindum2 | -1.533088 1.210456 -1.27 0.206**

**chaindum3 | -3.103495 .7580529 -4.09 0.000**

**chaindum4 | 7.045779 1.390257 5.07 0.000**

**\_cons | -18.03822 5.113993 -3.53 0.000**

**----------------------------------------------------------------**

**Elasticities after regress**

**y = Fitted values (predict)**

**= 17.786237**

**----------------------------------------------------------------**

**variable | dy/ex Std. Err. z P>|z|**

**---------+------------------------------------------------------**

**minwage | 5.462959 3.42082 1.60 0.110**

**abovemin | .4427413 .23188 1.91 0.056**

**hrsopen | 26.31874 3.21118 8.20 0.000**

**nregs | 3.698057 1.19998 3.08 0.002**

**chaind~2 | -.3078408 .24306 -1.27 0.205**

**chaind~3 | -.7552389 .18447 -4.09 0.000**

**chaind~4 | .9650468 .19042 5.07 0.000**

Despite the fact that the minimum wage change has a positive coefficient and a positive elasticity, it lacks statistical significance even at the 0.10 level in the pooled OLS model. We are forced to fail to reject the null hypothesis and conclude that the minimum wage has no discernable impact on employment. The starting wage paid above the minimum wage is more significant, but misses the 0.05 mark. In this model, the strongest determinates of employment are found to be on the demand side of labor, with the open hours, registers, and type of restaurant.

Interpreting the model elasticities as if they were significant, we can see that labor is much more elastic with respect to the minimum wage than any wages paid above it. Where we would expect a small change in the minimum wage to create a large change in employment, we can also expect a large change in wages above minimum wage to create a relatively smaller change in employment.

3.

A random effects model and its elasticities were estimated on the same data with the same specification using the following commands. The results follow as well:

. xtset sheet year

. xtreg fte minwage abovemin hrsopen nregs chaindum2-chaindum4, re

. mfx, dyex

**Random-effects GLS regression Number of obs = 752**

**Group variable: sheet Number of groups = 407**

**R-sq: within = 0.0316 Obs per group: min = 1**

**between = 0.4176 avg = 1.8**

**overall = 0.3450 max = 2**

**Wald chi2(7) = 306.88**

**corr(u\_i, X) = 0 (assumed) Prob > chi2 = 0.0000**

**----------------------------------------------------------------**

**fte | Coef. Std. Err. z P>|z|**

**-------------+--------------------------------------------------**

**minwage | 1.279639 .709426 1.80 0.071**

**abovemin | 1.998773 .9579281 2.09 0.037**

**hrsopen | 1.752826 .1670521 10.49 0.000**

**nregs | 1.03148 .3111307 3.32 0.001**

**chaindum2 | -1.84191 1.204849 -1.53 0.126**

**chaindum3 | -3.052681 .8968108 -3.40 0.001**

**chaindum4 | 6.871489 1.185942 5.79 0.000**

**\_cons | -17.4391 4.293264 -4.06 0.000**

**-------------+--------------------------------------------------**

**sigma\_u | 4.1014308**

**sigma\_e | 6.0267774**

**rho | .31653252 (fraction of variance due to u\_i)**

**----------------------------------------------------------------**

**Elasticities after xtreg**

**y = Linear prediction (predict)**

**= 17.789093**

**----------------------------------------------------------------**

**variable | dy/ex Std. Err. z P>|z|**

**---------+------------------------------------------------------**

**minwage | 5.848225 3.24223 1.80 0.071**

**abovemin | .4610998 .22099 2.09 0.037**

**hrsopen | 25.3513 2.41609 10.49 0.000**

**nregs | 3.739115 1.12785 3.32 0.001**

**chaind~2 | -.3698517 .24193 -1.53 0.126**

**chaind~3 | -.7428732 .21824 -3.40 0.001**

**chaind~4 | .9411747 .16244 5.79 0.000**

**----------------------------------------------------------------**

Acknowledging the fact that there are two observations for every one restaurant in our dataset has allowed us to recognize the unobserved heterogeneity inherent in our data and simply ignored in our pooled OLS estimation. In order for our random effects estimation to be unbiased, we must assume that this unobserved heterogeneity is uncorrelated with our independent variables.

Turning to the interpretation of the results, the coefficients on both the minimum wage and the starting wage paid above the minimum have strengthened positively. Now we can reject the null hypothesis that there is no effect of minimum wage on employment at the 0.10 level. Just as with the OLS model, the minimum wage policy is much more elastic than wages paid above the minimum.

We can test if the use of a random effects model is more appropriate than a pooled OLS model by comparing the variance matrices and testing to see if the variance of the unobserved heterogeneity term is equal to zero. The hypothesis test for this analysis is as follows:

H0: Var(σ2c) = 0

H1: Var(σ2c)≠ 0

If we fail to reject the null hypothesis, there is no unobserved heterogeneity and we can use the pooled OLS. If we reject the null hypothesis, there is unobserved heterogeneity and random effects will be the more efficient model. This test was conducted in STATA with the following command and had the following results:

. xttest0

**Breusch and Pagan Lagrangian multiplier test for random effects**

**fte[sheet,t] = Xb + u[sheet] + e[sheet,t]**

**Estimated results:**

**| Var sd = sqrt(Var)**

**---------+-----------------------------**

**fte | 80.62743 8.979278**

**e | 36.32205 6.026777**

**u | 16.82173 4.101431**

**Test: Var(u) = 0**

**chibar2(01) = 33.54**

**Prob > chibar2 = 0.0000**

We can reject the null hypothesis and conclude that random effects is more efficient than pooled OLS. This does not necessarily mean that random effects is the best test. In reality, it is hard to find situations in which the unobserved heterogeneity is uncorrelated with the independent variables, so random effects will produce biased estimates.

4.

Next, a fixed effects model was estimated three ways. First, the following commands were used to estimate a fixed effects model using the “dummy variable” approach:

. quietly tab sheet, gen(sheetdum)

. gen yeardum = 0

. replace yeardum = 1 if year==2

. set matsize 425

. quietly reg fte minwage abovemin hrsopen nregs chaindum2-chaindum4 sheetdum2-sheetdum410 yeardum, robust

These commands generate a new dummy variable for each different restaurant, and include that information in the model. The results are in the table at the end of this question. Second, the following command was used to estimate a fixed effects model using the STATA demeaning approach:

. xtreg fte minwage abovemin hrsopen nregs chaindum2-chaindum4 yeardum, fe

This method works by calculating a mean for each variable, and subtracting each observation from that mean. The results are in the table at the end of this question. Lastly, a fixed effects model was estimated using the first-differences approach with the following commands:

. webuse Card\_Kreuger\_wide

. quietly tab chain, gen(chaindum)

. gen minwage1 = 4.25

. gen minwage2 = 4.25

. replace minwage2 = 5.05 if state==1

. gen abovemin1 = wage\_st1-minwage1

. gen abovemin2 = wage\_st2-minwage2

. gen fdfte = (empft2+(0.5\*emppt2))-(empft1+(0.5\*emppt1))

. gen fdminwage = minwage2-minwage1

. gen fdabovemin = abovemin2-abovemin1

. gen fdhrsopen = hrsopen2-hrsopen1

. gen fdnregs = nregs2-nregs1

\*Chain dummies are time invariant

. reg fdfte fdminwage fdabovemin fdhrsopen fdnregs, robust

I chose not to generate the first-differenced chain dummies because they do not vary across time (theoretically), and would difference out to zero. The results from these three models have been added into the table below:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Model 1** | **Model 2** | **Model 3** |
| *DV: fte* | FE Dummy | FE Demean | FE First-Diff |
|  |  |  |  |
| minwage | 4.140\* | 4.140\*\* | 4.140\* |
|  | (1.961) | (1.529) | (1.878) |
|  |  |  |  |
| abovemin | 2.440 | 2.440 | 2.440 |
|  | (1.401) | (1.323) | (1.342) |
|  |  |  |  |
| hrsopen | 0.965\* | 0.965\* | 0.965\* |
|  | (0.391) | (0.443) | (0.374) |
|  |  |  |  |
| nregs | 1.199\* | 1.199\* | 1.199\* |
|  | (0.556) | (0.589) | (0.532) |
|  |  |  |  |
| chaindum2 | -24.44\*\*\* | . | . |
|  | (5.288) | . | . |
|  |  |  |  |
| chaindum3 | -15.27\*\* | . | . |
|  | (5.003) | . | . |
|  |  |  |  |
| chaindum4 | 0.162 | . | . |
|  | (4.665) | . | . |
|  |  |  |  |
| Constant | -3.152 | -18.88\* | -2.274 |
|  | (11.75) | (9.583) | (1.370) |
|  |  |  |  |
| R2 | 0.7980 | 0.0495 | 0.0488 |
| N | 752 | 752 | 345 |
| Standard errors in parentheses. Sheet and year dummies omitted for brevity sake. Significance is denoted as follows: \* p<0.05 \*\* p<0.01 \*\*\* p<0.001 | | | |

Although the dummy approach did create coefficients for the chain dummies, we know that these variables are time invariant and cannot be measured in a fixed effects context. Looking across all three methods, the coefficients are exactly the same. Unlike random effects, we now get very significant results for the minimum wage. The elasticities (mfx, dyex) for the demeaned fixed effects estimates are as follows:

**Elasticities after xtreg**

**y = Linear prediction (predict)**

**= 17.786237**

**----------------------------------------------------------------**

**variable | dy/ex Std. Err. z P>|z|**

**---------+------------------------------------------------------**

**minwage | 18.91853 6.98751 2.71 0.007**

**abovemin | .5627766 .3052 1.84 0.065**

**hrsopen | 13.95401 6.40222 2.18 0.029**

**nregs | 4.348027 2.13336 2.04 0.042**

**----------------------------------------------------------------**

With our fixed effects estimator, the elasticity of employment with respect to the minimum wage has shot through the roof. This means that we can expect very large increases in employment with moderate increases to the minimum wage.

5.

Regardless of the model used, each and every one reported a positive coefficient for the minimum wage. Arguably, the most valid test given the data is the fixed effects model, and this model in particular estimated very strong, positive correlation between the minimum wage and employment. Fixed effects is the best test for this data because we know there is unobserved heterogeneity with the pooled OLS model and we can very easily assume that there is correlation between store-specific effects and our chosen independent variables. With a coefficient of 4.14, we can expect at least 4 new full-time equivalent workers for every $1 increase in the minimum wage.

6. (a)

In a simple panel context, the strict exogeneity assumption told us that the data for an individual could not be correlated with the error for any other period. The unique twist to the model in this problem is that there are not only unobserved factors for each individual, but also unobserved factors for each family. Since both the individual and family effects are time invariant, they should both “cancel out” in a fixed effects context; so the strict exogeneity assumption will remain as just E[x1’ϵ2]=0 and E[x2’ϵ1]=0. For random effects, however, since both the individual and family specific effects are allowed to remain in the error structure of the model, the strict exogeneity assumption must be expanded. In a two time period case, since both c and f are time invariant, we just need to add four more assumptions: E[x1’c]=0, E[x2’c]=0, E[x1’f]=0, and E[x2’f]=0.

(b)

If Σ for the individual looks like…

|  |  |  |
| --- | --- | --- |
| σ2ϵ +σ2c | … | σ2c |
| … | … | … |
| σ2c | … | σ2ϵ +σ2c |

Then we can imagine one more matrix for the family on the way up to Ω (let’s call it Φ) that looks like…

|  |  |  |
| --- | --- | --- |
| Σ+σ2f | … | σ2f |
| … | … | … |
| σ2f | … | Σ+σ2f |

Where every σ2f represents a matrix with the same dimensions of Σ, filled with just σ2f.

Then finally Ω would look like…

|  |  |  |
| --- | --- | --- |
| Φ | … | 0 |
| … | … | … |
| 0 | … | Φ |

Where every cell filled with a 0 represents a matrix with the same dimensions of Φ, filled with zeroes.

7.

Given a dataset that looks like:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Person | Year | Wage | Education | Experience | Training |
| 1 | 2000 | 18 | 12 | 0 | 1 |
| 1 | 2001 | 19 | 12 | 1 | 1 |
| 1 | 2002 | 20 | 12 | 2 | 1 |
| 2 | 2000 | 21 | 14 | 1 | 0 |
| 2 | 2001 | 23 | 14 | 2 | 0 |
| 2 | 2002 | 25 | 14 | 3 | 0 |

Taking the first differences of the data looks like this:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Person | Year | Wage | Education | Experience | Training |
| 1 | 2001 | 1 | 0 | 1 | 0 |
| 1 | 2002 | 2 | 0 | 2 | 0 |
| 2 | 2001 | 2 | 0 | 1 | 0 |
| 2 | 2002 | 4 | 0 | 2 | 0 |

In this scenario, both education and training did not vary with time. As a result, subtracting the first value from the subsequent values gives us a column of zeroes. These need to be omitted from the regression because they are perfectly collinear. The same thing happens if we were to demean the data.