Preliminary

Irvin Rojas and Jisang Yu

5/03/2021

Data summary

- (I don't really know how to make a summary table... so here is the ugly one)

```
#libraries
library(tidyverse)
library(foreign)
library(plm)
library(lmtest)

#read data
data<-read.dta("cleaned_v1.dta")
data_ind<-read.dta("cleaned_ind_v1.dta")

#mutate individual data
data_ind %>%
  mutate(
    cd_a=as_factor(cd_a),
    sector=as_factor(sector),
    occ=as_factor(occ),
  )
```

```
data %>%
  summarise(
          count = n(),
          mean_d_wage = mean(d_mhrwageactive, na.rm = TRUE),
          sd_d_wage = sd(d_mhrwageactive, na.rm = TRUE),
          mean_w_covid_apr = mean(w_covid_apr, na.rm = TRUE),
          sd_w_covid_apr = sd(w_covid_apr, na.rm = TRUE))
```

```
##   count mean_d_wage sd_d_wage mean_w_covid_apr sd_w_covid_apr
## 1    37   0.6018542  11.58796         2181.928       838.4911
```

```
data_ind %>%
  summarise(
          count = n(),
          mean_d_wage = mean(d_hrwage, na.rm = TRUE),
          sd_d_wage = sd(d_hrwage, na.rm = TRUE))
```

```
##   count mean_d_wage sd_d_wage
## 1 49139   -3.276601  65.91979
```

Preliminary regressions: City-level

- First, let's assess whether the Covid exposure of the Mexican migrants in US has any impacts on the local wages (W) in the Mexican cities

$$\Delta W_c = \alpha + \beta \sum_s \theta_{sc} Covid_s + \varepsilon_c$$

where

$$\Delta W_c = W_{c,2020q4} - W_{c,2019q4}$$

$$\theta_{sc} = \frac{\text{Number of travelers from city } c \text{ to state } s}{\text{Total number of travels from city } c}$$

and $Covid_s$ is the number of cases in state s (per million people).

```
#regression
reg1<-lm(d_mhrwageactive~w_covid_apr, data=data)
coeftest(reg1)
```

```
##
## t test of coefficients:
##
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) 16.3771725  5.5299734  2.9615 0.005936 **
## w_covid_apr -0.0069460  0.0023705 -2.9301 0.006421 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Second, let's assess whether there are differential effects between agricultural and non-agricultural sectors:

$$\Delta W_c = \alpha + \beta_1 \sum_s \theta_{sc} Covid_s + \beta_2 \sum_s \theta_{sc} Ag_s Covid_s + \varepsilon_c$$

where Ag_s is the share of Mexican-born agricultural workers over total agricultural workers in state s.

```
#regression2
reg2<-lm(d_mhrwageactive~w_covid_apr+w_covid_apr_ag, data=data)
coeftest(reg2)
```

```
##
## t test of coefficients:
##
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)    14.1762750  6.4987427  2.1814 0.037407 *
## w_covid_apr    -0.0074778  0.0025245 -2.9621 0.006044 **
## w_covid_apr_ag  0.0303088  0.0458163  0.6615 0.513495
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Preliminary regressions: Individual-level

- Similarly, let's assess whether the Covid exposure of the Mexican migrants in US has any impacts on the wages (W) earned by the Mexican individuals

$$\Delta W_{ci} = \alpha + \beta \sum_s \theta_{sc} Covid_s + \varepsilon_{ci}$$

where

$$\Delta W_{ci} = W_{ci,2020q4} - W_{ci,2019q4}$$

$$\theta_{sc} = \frac{\text{Number of travelers from city } c \text{ to state } s}{\text{Total number of travels from city } c}$$

and $Covid_s$ is the number of cases in state s (per million people).

```
#regression
regind1<-lm(d_hrwage~w_covid_apr, data=data_ind, weights=weight)
coeftest(regind1)
```

```
##
## t test of coefficients:
##
##                Estimate  Std. Error t value  Pr(>|t|)
## (Intercept) -4.00584947  0.84870160 -4.7200 2.368e-06 ***
## w_covid_apr  0.00069108  0.00030155  2.2918   0.02192 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Second, let's assess whether there are differential effects between agricultural and non-agricultural sectors:

$$\Delta W_{ci} = \alpha + \beta_1 \sum_s \theta_{sc} Covid_s + \beta_2 \sum_s \theta_{sc} Ag_s Covid_s + \varepsilon_c$$

where Ag_s is the share of Mexican-born agricultural workers over total agricultural workers in state s.

```
#regression2
regind2<-lm(d_hrwage~w_covid_apr+w_covid_apr_ag, data=data_ind, weights=weight)
coeftest(regind2)
```

```
##
## t test of coefficients:
##
##                   Estimate  Std. Error t value  Pr(>|t|)
## (Intercept)    -4.67113622  1.14078223 -4.0947 4.237e-05 ***
## w_covid_apr     0.00056851  0.00033265  1.7090   0.08745 .
## w_covid_apr_ag  0.00895338  0.01025884  0.8727   0.38281
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Funny results from the full sample. Let's look blue collar and agriculture separately:
  - Agriculture

```
#regression
regind1<-lm(d_hrwage~w_covid_apr, data=data_ind, weights=weight, subset=(occ=="agriculture"))
coeftest(regind1)
```

```
##
## t test of coefficients:
##
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.7705323  9.0961794  0.8543   0.3938
## w_covid_apr -0.0031676  0.0024107 -1.3140   0.1902
```

```
#regression2
regind2<-lm(d_hrwage~w_covid_apr+w_covid_apr_ag, data=data_ind, weights=weight, subset=(occ=="agricultu
coeftest(regind2)
```

```
##
## t test of coefficients:
##
```

```
##                   Estimate  Std. Error t value  Pr(>|t|)
## (Intercept)    -77.7453492  12.2384622 -6.3525 1.122e-09 ***
## w_covid_apr     -0.0135999   0.0023695 -5.7395 2.980e-08 ***
## w_covid_apr_ag   1.0015696   0.1102566  9.0840 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Blue collar

```
#regression
regind1<-lm(d_hrwage~w_covid_apr, data=data_ind, weights=weight, subset=(occ=="blue collar"))
coeftest(regind1)
```

```
##
## t test of coefficients:
##
##                Estimate  Std. Error t value Pr(>|t|)
## (Intercept) -3.07068109  1.25386412 -2.4490  0.01435 *
## w_covid_apr -0.00010998  0.00043163 -0.2548  0.79888
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#regression2
regind2<-lm(d_hrwage~w_covid_apr+w_covid_apr_ag, data=data_ind, weights=weight, subset=(occ=="blue colla
coeftest(regind2)
```

```
##
## t test of coefficients:
##
##                   Estimate  Std. Error t value Pr(>|t|)
## (Intercept)    -3.44374102  1.76033547 -1.9563  0.05047 .
## w_covid_apr    -0.00016641  0.00047037 -0.3538  0.72351
## w_covid_apr_ag  0.00485209  0.01606898  0.3020  0.76270
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Considerations

- Small sample: Number of cities=37
- Other sectors?
- Mexico local Covid exposure?
- Control for the importance of the migration in city c?