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")
data_ind_ag<-read.dta("cleaned_ind_ag_v1.dta")

#mutate individual data
data_ind %>%
  mutate(
    mx_mun=as_factor(mx_mun),
  )
data_ind_ag %>%
  mutate(
    mx_mun=as_factor(mx_mun),
  )
```

```
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),
          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 49139   -3.276601  65.91979         2490.867       1836.172
```

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:

$$\Delta W_{mi} = \alpha + \beta \sum_s \theta_{sm} Covid_s + \varepsilon_{mi}$$

where

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

$$\theta_{sm} = \frac{\text{Number of travelers from Mexican municipality } m \text{ to state } s}{\text{Total number of travels from Mexican municipality } m}$$

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) -2.83715408  0.44935927 -6.3138 2.747e-10 ***
## w_covid_apr  0.00018132  0.00014127  1.2835    0.1993
## ---
## 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:

$$\Delta W_{mi} = \alpha + \beta_1 \sum_s \theta_{sm} Covid_s + \beta_2 \sum_s \theta_{sm} Ag_s Covid_s + \varepsilon_{mi}$$

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)    -2.40796670  0.61825553 -3.8948 9.842e-05 ***
## w_covid_apr     0.00024959  0.00015659  1.5939    0.1110
## w_covid_apr_ag -0.00534816  0.00529145 -1.0107    0.3122
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

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

```
##
## t test of coefficients:
##
##                Estimate  Std. Error t value Pr(>|t|)
## (Intercept) -1.6940e+00  1.6230e+00 -1.0437   0.2969
## w_covid_apr -2.2726e-05  4.1505e-04 -0.0548   0.9563
```

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

```
##
## t test of coefficients:
##
##                   Estimate  Std. Error t value Pr(>|t|)
## (Intercept)    -1.0420e+00  2.3508e+00 -0.4432   0.6577
## w_covid_apr     3.7408e-05  4.4384e-04  0.0843   0.9328
## w_covid_apr_ag -6.5223e-03  1.7007e-02 -0.3835   0.7014
```

Considerations

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