Replication of: Feng et al (2010), Linkages among climate change, crop yields, and Mexico-US cross-border migration

Rafael Grillo Avila, Sol Kim, Catherine Ledna, & Samira Siddique

October 8, 2018

Executive Summary / Write-Up

Objective

Our objective was to reproduce, as much as possible and given the supplementary material, the results of the Feng et al paper, which can be found summarized in Table 1 (14258), with robustness checks in Table 2 (14259). Table 1 is displayed below:

Progress and Stumbling Blocks

First, given all the papers we've read with data and scripts that were, however flawed, available online, it was surprising to see that this was not the case for the paper. Although the supplementary material was much appreciated, as it makes it easier to understand the context of the data, it is impossible to know whether one was taking the same steps as the authors without having access to the code the authors used itself.

The second stumbling block came when trying to download the emigration data from the website of Integrated Public Use Microdata Series - International (IPUMS). Unfortunately, downloading data from the website required creating an account that is still pending approval. We therefore proceed using two methods:

1. First, we manually type in the data presented in "Table S1. Data for Each Mexican State," presented below:
2. Second, we download data from Mexico's National Institute of Statistics and Geography, abbreviated INEGI based on the Spanish spelling. We obtain data from each of the "Censo General de Población y Vivienda" (General Population and Housing Census) 1995, 2000, and 2005 (XI, XII, and XIII), and navigate to "Tabulados" (Tabulations) and filter the results by "Interactivos (cubos)" (cubed interactive groups) and "Area Geográfica" "Nacional y Entidad Federativa" (Geography: National and State). Then we run a consultation on "Poblacion total con estimación" (Total population with estimation) by "Entidad municipio y loc" (Municipal entity and locale) and "Edad quinquenal" (Age by five-year groupings). We then export the results to .csv, clean the data by eliminating non-data rows and columns, and import into R. We are able to obtain total population data this way. However, we realize that obtaining the data for projections of death rates and for migration turns out to be too time-consuming and death rates too unclear that we resort to using the first method described above to test the regression part of the analysis first.

We finish by trying to run the regressions described in the paper on the manually inputted data.

Results

It is interesting to note that the log yield of corn plus wheat for Veracruz was lower than the log yield of just corn for the years 2000 to 2005. This calls into question the data presented in the charts, and emphasizes the need for current or future replication attempts to really get to the source.

Table 1. Estimated effect of crop yields on emigration (β)

Yield statistic	Pooled OLS	RE	FE	FE-TSLS	FE-LIML
Corn					
Log of crop yields	0.005 (0.008)	0.003 (0.009)	−0.117† (0.026)	−0.211† (0.036)	−0.225† (0.042)
State dummies	No	No	Yes	Yes	Yes
F-statistic*	—	—	—	5.5	5.5
P value	—	—	—	0.0009	0.0009
Adjusted R^2	0.0115	0.0115	0.6775	0.5897	0.5619
Corn plus wheat					
Log of crop yields	0.006 (0.008)	0.005 (0.009)	−0.113† (0.031)	−0.183† (0.036)	−0.214† (0.051)
State dummies	No	No	Yes	Yes	Yes
F-statistic*	—	—	—	9.2	9.2
P value	—	—	—	0.0000	0.0000
Adjusted R^2	0.0139	0.0139	0.6490	0.6076	0.5622

Sample size is 64 for all regressions. Numbers reported in parentheses are robust SEs that allow for possible heteroscedasticity.
**F*-statistic for instruments in first-stage equation.
†Significant at 1% level.

Figure 1:

We first test for a reduced form relationship between climate and migration based on Part 3 of the supplemental data, creating quadratic equations as necessary. Using the numbers provided and the plm package, we actually get results for the pooled OLS that are different from the reported results in Table S2. The results are closer when we run Fixed Effects, even though the estimates are still slightly off here.

Turning our attention to testing the two-stage least squares, with both fixed effects and without, we get results that are generally an order of magnitude lower and statistically insignificant, though in the same general direction as the results of the main study.

Code

Setup

Begin by loading the required packages and data, and defining necessary project settings.

```
# Set seed
knitr::opts_chunk$set(set.seed(123456789))

# Install and load useful libraries
# install.packages("pacman")
library(pacman)

p_load(readr, lubridate, ggplot2, stargazer, fastDummies, pdftools,
       foreign, pander, knitr, purrr, tidyverse, tidyselect, stringr,
       xlsx, plm)
```

Next, we construct the emigration data using the residual approach mentioned in the Supplementary Material.

We try to obtain population and emigration data from the Integrated Public Use Microdata Series - International (IPUMS), but, failing that, obtain it from .csv files downloaded from Mexico's National Institute of Statistics and Geography (INEGI).

```
# Download 1995 - 2005 total population by state and age group
popMX1995 <- read_csv("INEGI_Exporta_20181009014955_1995.csv")
```

```
## Warning: Missing column names filled in: 'X1' [1]
```

Table S1. Data for Each Mexican State

A. For the period of 1995-2000

Mexican State	Log Corn Yield	Log Yield of Corn Plus Wheat	Ratio of Emigrants	Annual Precipitation (m)	Annual Mean Temperature (°C)	Summer Mean Temperature (°C)
Aguascalientes	-0.020	-0.018	8.4%	0.37	17.52	21.12
Baja California	0.893	1.549	6.4%	0.21	18.84	22.07
Baja California Sur	1.465	1.472	4.0%	0.18	22.46	25.56
Campeche	0.182	0.182	5.0%	1.41	26.13	27.91
Chiapas	0.581	0.581	6.0%	1.77	23.84	25.01
Chihuahua	0.685	0.757	7.3%	0.35	17.53	23.77
Coahuila	-0.297	0.010	6.2%	0.35	19.95	25.87
Colima	0.837	0.837	4.4%	0.85	25.12	26.76
Distrito Federal	0.505	0.505	2.3%	0.76	15.85	17.80
Durango	0.115	0.179	10.1%	0.34	17.36	22.32
Guanajuato	0.612	0.928	10.1%	0.56	18.41	21.38
Guerrero	0.762	0.762	7.6%	0.91	25.12	26.43
Hidalgo	0.504	0.514	8.4%	0.68	16.41	18.69
Jalisco	1.191	1.207	8.0%	0.62	20.54	23.49
México	1.247	1.223	3.2%	0.57	14.47	16.45
Michoacán	0.796	0.891	11.5%	0.63	19.68	22.13
Morelos	0.750	0.755	8.1%	0.98	21.75	23.52
Nayarit	1.034	1.034	9.3%	1.21	25.11	27.81
Nuevo León	-0.452	-0.297	5.3%	0.48	20.31	25.28
Oaxaca	0.178	0.174	5.9%	1.66	21.83	23.37
Puebla	0.409	0.418	4.3%	1.84	17.45	19.47
Querétaro	0.424	0.457	7.1%	0.40	18.32	21.28
Quintana Roo	-0.764	-0.764	4.1%	1.42	25.72	27.54
San Luis Potosí	-0.503	-0.503	8.2%	0.76	21.07	24.76
Sinaloa	1.719	1.686	6.2%	0.60	25.13	29.10
Sonora	1.545	1.663	6.9%	0.33	22.39	27.81
Tabasco	0.173	0.173	3.6%	2.37	27.20	29.37
Tamaulipas	0.572	0.499	7.9%	0.68	23.62	28.14
Tlaxcala	0.514	0.586	6.8%	0.84	14.46	16.33
Veracruz	0.492	0.491	5.9%	1.60	23.01	25.82
Yucatán	-0.421	-0.421	4.4%	0.96	26.36	28.30
Zacatecas	-0.228	-0.213	12.0%	0.46	17.22	21.35

Figure 2:

B. For the period of 2000-2005

Mexican State	Log Corn Yield	Log Yield of Corn Plus Wheat	Ratio of Emigrants	Annual Precipitation (m)	Annual Mean Temperature (°C)	Summer Mean Temperature (°C)
Aguascalientes	-0.204	-0.205	1.5%	0.54	17.62	20.78
Baja California	0.987	1.509	3.2%	0.19	18.76	22.74
Baja California Sur	1.631	1.647	-0.6%	0.18	22.20	25.28
Campeche	0.218	0.218	2.5%	1.48	25.96	28.05
Chiapas	0.627	0.627	2.5%	1.99	23.84	24.80
Chihuahua	0.944	0.974	5.5%	0.42	18.24	24.42
Coahuila	-0.209	-0.075	0.7%	0.44	20.80	26.49
Colima	0.910	0.910	5.0%	0.86	25.71	27.10
Distrito Federal	0.428	0.428	1.3%	0.77	15.95	17.31
Durango	0.324	0.333	4.7%	0.41	17.53	22.05
Guanajuato	1.044	1.125	5.1%	0.76	18.19	20.66
Guerrero	0.804	0.804	7.9%	1.00	25.03	26.10
Hidalgo	0.794	0.792	4.5%	0.62	16.62	18.59
Jalisco	1.474	1.474	3.5%	0.78	20.53	23.10
México	1.181	1.164	3.7%	0.71	14.37	15.90
Michoacán	0.960	1.011	9.1%	0.88	18.28	20.33
Morelos	0.855	0.859	6.0%	1.04	21.24	22.76
Nayarit	1.274	1.274	6.8%	1.14	24.93	27.61
Nuevo León	-0.187	0.118	0.9%	0.67	18.25	21.95
Oaxaca	0.198	0.190	6.1%	1.54	22.57	23.81
Puebla	0.448	0.448	3.6%	1.52	17.73	19.66
Querétaro	0.862	0.867	-0.03%	0.59	18.11	20.16
Quintana Roo	-0.951	-0.951	-4.9%	1.43	25.59	27.18
San Luis Potosí	-0.485	-0.484	4.5%	0.77	21.25	24.45
Sinaloa	1.972	1.923	3.6%	0.62	24.91	28.91
Sonora	1.436	1.590	1.0%	0.36	22.43	28.50
Tabasco	0.394	0.394	1.8%	2.28	26.59	28.57
Tamaulipas	0.665	0.663	1.6%	0.81	23.66	27.56
Tlaxcala	0.795	0.779	0.4%	0.65	14.45	16.01
Veracruz	0.587	0.586	3.2%	1.72	23.26	25.78
Yucatán	-0.426	-0.426	-1.4%	0.92	26.17	28.31
Zacatecas	0.113	0.113	7.2%	0.56	16.90	20.36

Figure 3:

```
## Parsed with column specification:
## cols(
##   .default = col_number(),
##   X1 = col_character()
## )
## See spec(...) for full column specifications.
```

```
popMX2000 <- read_csv("INEGI_Exporta_20181009020938_2000.csv")
```

```
## Warning: Missing column names filled in: 'X1' [1]
```

```
## Parsed with column specification:
## cols(
##   X1 = col_character(),
##   Total = col_number(),
##   `De 0 a 4 a<f1>os` = col_number(),
##   `De 5 a 9 a<f1>os` = col_number(),
##   `De 10 a 14 a<f1>os` = col_number(),
##   `De 15 a 19 a<f1>os` = col_number(),
##   `De 20 a 24 a<f1>os` = col_number(),
##   `De 25 a 29 a<f1>os` = col_number(),
##   `De 30 a 34 a<f1>os` = col_number(),
##   `De 35 a 39 a<f1>os` = col_number(),
##   `De 40 a 44 a<f1>os` = col_number(),
##   `De 45 a 49 a<f1>os` = col_number(),
##   `De 50 a 54 a<f1>os` = col_number(),
##   `De 55 a 59 a<f1>os` = col_number(),
##   `De 60 a 64 a<f1>os` = col_number(),
##   `De 65 a 69 a<f1>os` = col_number(),
##   `De 70 a 74 a<f1>os` = col_number(),
##   `De 75 y m<e1>s a<f1>os` = col_number(),
##   `No especificado` = col_number()
## )
```

```
popMX2005 <- read_csv("INEGI_Exporta_20181009021100_2005.csv")
```

```
## Warning: Missing column names filled in: 'X1' [1]
```

```
## Parsed with column specification:
## cols(
##   .default = col_number(),
##   X1 = col_character()
## )
## See spec(...) for full column specifications.
```

Since we only obtain population data in this manner, we resort to manually copying the data from the PDF file. We copy both the emigration data and the crop data in this way to test the regressions, and then return to trying to construct the dataset from the sources provided.

```
# Hand-writing Table S1 data for 1995-2000 and 2000-2005
mexicanState <- c("Aguascalientes", "Baja California", "Baja California Sur", "Campeche", "Chiapas", "Ch

timePeriod1995to2000 <- rep("1995-2000", length(mexicanState))
timePeriod2000to2005 <- rep("2000-2005", length(mexicanState))

logCornYield1995to2000 <-
  c(-0.020, 0.893, 1.465, 0.182, 0.581, 0.685, -0.297, 0.837, 0.505,
```

```
    0.115, 0.612, 0.762, 0.504, 1.191, 1.247, 0.796, 0.750, 1.034,
    -0.452, 0.178, 0.409, 0.424, -0.764, -0.503, 1.719, 1.545, 0.173,
    0.572, 0.514, 0.492, -0.421, -0.228)

logYieldOfCornPlusWheat1995to2000 <-
  c(-0.018, 1.549, 1.472, 0.182, 0.581, 0.757, 0.010, 0.837, 0.505,
    0.179, 0.928, 0.762, 0.514, 1.207, 1.223, 0.891, 0.755, 1.034,
    -0.297, 0.174, 0.418, 0.457, -0.764, -0.503, 1.686, 1.663, 0.173,
    0.499, 0.586, 0.491, -0.421, -0.213)

ratioOfEmigrants1995to2000 <-
  c(0.084, 0.064, 0.040, 0.050, 0.060, 0.073, 0.062, 0.044, 0.023,
    0.101, 0.101, 0.076, 0.084, 0.080, 0.032, 0.115, 0.081, 0.093,
    0.053, 0.059, 0.043, 0.071, 0.041, 0.082, 0.062, 0.069, 0.036,
    0.079, 0.068, 0.059, 0.044, 0.120)

annualPrecipitation1995to2000 <-
  c(0.37, 0.21, 0.18, 1.41, 1.77, 0.35, 0.35, 0.85, 0.76, 0.34, 0.56,
    0.91, 0.68, 0.62, 0.57, 0.63, 0.98, 1.21, 0.48, 1.66, 1.84, 0.40,
    1.42, 0.76, 0.60, 0.33, 2.37, 0.68, 0.84, 1.60, 0.96, 0.46)

annualMeanTemp1995to2000 <-
  c(17.52, 18.84, 22.46, 26.13, 23.84, 17.53, 19.95, 25.12, 15.85,
    17.36, 18.41, 25.12, 16.41, 20.54, 14.47, 19.68, 21.75, 25.11,
    20.31, 21.83, 17.45, 18.32, 25.72, 21.07, 25.13, 22.39, 27.20,
    23.62, 14.46, 23.01, 26.36, 17.22)

summerMeanTemp1995to2000 <-
  c(21.12, 22.07, 25.56, 27.91, 25.01, 23.77, 25.87, 26.76, 17.80,
    22.32, 21.38, 26.43, 18.69, 23.49, 16.45, 22.13, 23.52, 27.81,
    25.28, 23.37, 19.47, 21.28, 27.54, 24.76, 29.10, 27.81, 29.37,
    28.14, 16.33, 25.82, 28.30, 21.35)

logCornYield2000to2005 <-
  c(-0.204, 0.987, 1.631, 0.218, 0.627, 0.944, -0.209, 0.910, 0.428,
    0.324, 1.044, 0.804, 0.794, 1.474, 1.181, 0.960, 0.855, 1.274,
    -0.187, 0.198, 0.448, 0.862, -0.951, -0.485, 1.972, 1.436, 0.394,
    0.665, 0.795, 0.587, -0.426, 0.113)

# Why is the log yield of corn plus wheat lower than the log yield of corn?
logYieldOfCornPlusWheat2000to2005 <-
  c(-0.205, 1.509, 1.647, 0.218, 0.627, 0.974, -0.075, 0.910, 0.428,
    0.333, 1.125, 0.804, 0.792, 1.474, 1.164, 1.011, 0.859, 1.274,
    0.118, 0.190, 0.448, 0.867, -0.951, -0.484, 1.923, 1.590, 0.394,
    0.663, 0.779, 0.586, -0.426, 0.113)

ratioOfEmigrants2000to2005 <-
  c(0.015, 0.032, -0.006, 0.025, 0.025, 0.055, 0.007, 0.050, 0.013,
    0.047, 0.051, 0.079, 0.045, 0.035, 0.037, 0.091, 0.060, 0.068,
    0.009, 0.061, 0.036, -0.003, -0.049, 0.045, 0.036, 0.010, 0.018,
    0.016, 0.004, 0.032, -0.014, 0.072)

annualPrecipitation2000to2005 <-
```

```
  c(0.54, 0.19, 0.18, 1.48, 1.99, 0.42, 0.44, 0.86, 0.77, 0.41, 0.76,
    1.00, 0.62, 0.78, 0.71, 0.88, 1.04, 1.14, 0.67, 1.54, 1.52, 0.59,
    1.43, 0.77, 0.62, 0.36, 2.28, 0.81, 0.65, 1.72, 0.92, 0.56)

annualMeanTemp2000to2005 <-
  c(17.62, 18.76, 22.20, 25.96, 23.84, 18.24, 20.80, 25.71, 15.95,
    17.53, 18.19, 25.03, 16.62, 20.53, 14.37, 18.28, 21.24, 24.93,
    18.25, 22.57, 17.73, 18.11, 25.59, 21.25, 24.91, 22.43, 26.59,
    23.66, 14.45, 23.26, 26.17, 16.90)

summerMeanTemp2000to2005 <-
  c(20.78, 22.74, 25.28, 28.05, 24.80, 24.42, 26.49, 27.10, 17.31,
    22.05, 20.66, 26.10, 18.59, 23.10, 15.90, 20.33, 22.76, 27.61,
    21.95, 23.81, 19.66, 20.16, 27.18, 24.45, 28.91, 28.50, 28.57,
    27.56, 16.01, 25.78, 28.31, 20.36)

tbl1995to2000 <-
  tibble(mexicanState = mexicanState,
         timePeriod = timePeriod1995to2000,
         logCornYield = logCornYield1995to2000,
         logYieldOfCornPlusWheat = logYieldOfCornPlusWheat1995to2000,
         ratioOfEmigrants = ratioOfEmigrants1995to2000,
         annualPrecipitation = annualPrecipitation1995to2000,
         annualMeanTemp = annualMeanTemp1995to2000,
         summerMeanTemp = summerMeanTemp1995to2000)

tbl2000to2005 <-
  tibble(mexicanState = mexicanState,
         timePeriod = timePeriod2000to2005,
         logCornYield = logCornYield2000to2005,
         logYieldOfCornPlusWheat = logYieldOfCornPlusWheat2000to2005,
         ratioOfEmigrants = ratioOfEmigrants2000to2005,
         annualPrecipitation = annualPrecipitation2000to2005,
         annualMeanTemp = annualMeanTemp2000to2005,
         summerMeanTemp = summerMeanTemp2000to2005)

dfMXTot <- bind_rows(tbl1995to2000, tbl2000to2005)
```

Replication

Now we test for a reduced form relationship between climate and migration based on Part 3 of the supplemental data, creating quadratic equations as necessary.

```
# Create additional variables
dfMXTotEdit <- dfMXTot %>%
  mutate(annualPrecipSq = (annualPrecipitation^2),
         annualMeanTempSq = (annualMeanTemp^2) / 100,
         summerMeanTempSq = (summerMeanTemp^2) / 100)

pooledReducedForm <-
  plm(ratioOfEmigrants ~ annualPrecipitation + annualPrecipSq +
        annualMeanTemp + annualMeanTempSq + summerMeanTemp +
        summerMeanTempSq, data = dfMXTotEdit)
```

```r
summary(pooledReducedForm)
```

```
## Oneway (individual) effect Within Model
##
## Call:
## plm(formula = ratioOfEmigrants ~ annualPrecipitation + annualPrecipSq +
##     annualMeanTemp + annualMeanTempSq + summerMeanTemp + summerMeanTempSq,
##     data = dfMXTotEdit)
##
## Balanced Panel: n = 32, T = 2, N = 64
##
## Residuals:
##        Min.     1st Qu.      Median     3rd Qu.        Max.
## -3.8936e-02 -9.1622e-03  2.4395e-16  9.1622e-03  3.8936e-02
##
## Coefficients:
##                       Estimate Std. Error t-value Pr(>|t|)
## annualPrecipitation -0.122384   0.145284 -0.8424  0.40726
## annualPrecipSq       0.019011   0.054135  0.3512  0.72829
## annualMeanTemp      -0.307426   0.134262 -2.2898  0.03040 *
## annualMeanTempSq     0.655142   0.318224  2.0587  0.04967 *
## summerMeanTemp       0.145483   0.101145  1.4384  0.16226
## summerMeanTempSq    -0.226720   0.210113 -1.0790  0.29048
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    0.029876
## Residual Sum of Squares: 0.017631
## R-Squared:      0.40985
## Adj. R-Squared: -0.42998
## F-statistic: 3.00942 on 6 and 26 DF, p-value: 0.022764
```

```r
feReducedForm <-
  lm(ratioOfEmigrants ~ annualPrecipitation + annualPrecipSq +
        annualMeanTemp + annualMeanTempSq + summerMeanTemp +
        summerMeanTempSq + factor(mexicanState), data = dfMXTotEdit)

summary(feReducedForm)
```

```
##
## Call:
## lm(formula = ratioOfEmigrants ~ annualPrecipitation + annualPrecipSq +
##     annualMeanTemp + annualMeanTempSq + summerMeanTemp + summerMeanTempSq +
##     factor(mexicanState), data = dfMXTotEdit)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.038936 -0.009162  0.000000  0.009162  0.038936
##
## Coefficients:
##                                       Estimate Std. Error t value
## (Intercept)                           1.427404   1.342331   1.063
## annualPrecipitation                  -0.122384   0.145284  -0.842
## annualPrecipSq                        0.019011   0.054135   0.351
```

```
## annualMeanTemp                           -0.307426   0.134262  -2.290
## annualMeanTempSq                          0.655142   0.318224   2.059
## summerMeanTemp                            0.145483   0.101145   1.438
## summerMeanTempSq                         -0.226720   0.210113  -1.079
## factor(mexicanState)Baja California      -0.012805   0.039814  -0.322
## factor(mexicanState)Baja California Sur -0.024112   0.071476  -0.337
## factor(mexicanState)Campeche              0.014344   0.160395   0.089
## factor(mexicanState)Chiapas               0.166361   0.143471   1.160
## factor(mexicanState)Chihuahua            -0.106156   0.056517  -1.878
## factor(mexicanState)Coahuila             -0.059296   0.058698  -1.010
## factor(mexicanState)Colima                0.017747   0.135503   0.131
## factor(mexicanState)Distrito Federal      0.049781   0.061263   0.813
## factor(mexicanState)Durango              -0.052617   0.036047  -1.460
## factor(mexicanState)Guanajuato            0.096705   0.035706   2.708
## factor(mexicanState)Guerrero              0.082045   0.133371   0.615
## factor(mexicanState)Hidalgo               0.074748   0.042609   1.754
## factor(mexicanState)Jalisco               0.098064   0.062202   1.577
## factor(mexicanState)Mexico               -0.011507   0.107582  -0.107
## factor(mexicanState)Michoacan             0.163614   0.047308   3.458
## factor(mexicanState)Morelos               0.175865   0.086741   2.027
## factor(mexicanState)Nayarit               0.070744   0.130877   0.541
## factor(mexicanState)Nuevo Leon           -0.012487   0.045647  -0.274
## factor(mexicanState)Oaxaca                0.204685   0.115683   1.769
## factor(mexicanState)Puebla                0.165509   0.073649   2.247
## factor(mexicanState)Queretaro             0.048035   0.034476   1.393
## factor(mexicanState)Quintana Roo         -0.003797   0.152887  -0.025
## factor(mexicanState)San Luis Potosi       0.083343   0.070515   1.182
## factor(mexicanState)Sinaloa              -0.032621   0.113947  -0.286
## factor(mexicanState)Sonora               -0.047279   0.076877  -0.615
## factor(mexicanState)Tabasco              -0.001450   0.202445  -0.007
## factor(mexicanState)Tamaulipas            0.013928   0.095177   0.146
## factor(mexicanState)Tlaxcala              0.004403   0.104449   0.042
## factor(mexicanState)Veracruz              0.130188   0.123297   1.056
## factor(mexicanState)Yucatan              -0.060640   0.154364  -0.393
## factor(mexicanState)Zacatecas             0.016386   0.031349   0.523
##                                          Pr(>|t|)
## (Intercept)                               0.29739
## annualPrecipitation                       0.40726
## annualPrecipSq                            0.72829
## annualMeanTemp                            0.03040 *
## annualMeanTempSq                          0.04967 *
## summerMeanTemp                            0.16226
## summerMeanTempSq                          0.29048
## factor(mexicanState)Baja California       0.75031
## factor(mexicanState)Baja California Sur   0.73856
## factor(mexicanState)Campeche              0.92942
## factor(mexicanState)Chiapas               0.25677
## factor(mexicanState)Chihuahua             0.07160 .
## factor(mexicanState)Coahuila              0.32171
## factor(mexicanState)Colima                0.89681
## factor(mexicanState)Distrito Federal      0.42384
## factor(mexicanState)Durango               0.15636
## factor(mexicanState)Guanajuato            0.01180 *
## factor(mexicanState)Guerrero              0.54380
```

```
## factor(mexicanState)Hidalgo              0.09116 .
## factor(mexicanState)Jalisco              0.12699
## factor(mexicanState)Mexico               0.91564
## factor(mexicanState)Michoacan            0.00188 **
## factor(mexicanState)Morelos              0.05298 .
## factor(mexicanState)Nayarit              0.59342
## factor(mexicanState)Nuevo Leon           0.78659
## factor(mexicanState)Oaxaca               0.08856 .
## factor(mexicanState)Puebla               0.03333 *
## factor(mexicanState)Queretaro            0.17533
## factor(mexicanState)Quintana Roo         0.98037
## factor(mexicanState)San Luis Potosi      0.24794
## factor(mexicanState)Sinaloa              0.77693
## factor(mexicanState)Sonora               0.54390
## factor(mexicanState)Tabasco              0.99434
## factor(mexicanState)Tamaulipas           0.88478
## factor(mexicanState)Tlaxcala             0.96670
## factor(mexicanState)Veracruz             0.30073
## factor(mexicanState)Yucatan              0.69764
## factor(mexicanState)Zacatecas            0.60560
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02604 on 26 degrees of freedom
## Multiple R-squared:  0.7287, Adjusted R-squared:  0.3427
## F-statistic: 1.888 on 37 and 26 DF,  p-value: 0.04686
```

Using the numbers provided and the plm package, we actually get results for the pooled OLS that are different from the reported results in Table S2. The results are closer when we run Fixed Effects, even though the estimates are still slightly off here.

Manually running a two-stage least squares fixed effects estimate, we obtain the following:

```
# For Corn only
firstStageCorn <-
  lm(logCornYield ~ annualPrecipitation + annualPrecipSq +
        annualMeanTemp + annualMeanTempSq + summerMeanTemp +
        summerMeanTempSq, data = dfMXTotEdit)

# For Corn + Wheat
firstStageCornPlusWheat <-
  lm(logYieldOfCornPlusWheat ~ annualPrecipitation + annualPrecipSq +
        annualMeanTemp + annualMeanTempSq + summerMeanTemp +
        summerMeanTempSq, data = dfMXTotEdit)

dfMXTotEdit2 <- dfMXTotEdit %>%
  mutate(logCornYieldIV = firstStageCorn$fitted.values,
         logYieldOfCornPlusWheatIV = firstStageCornPlusWheat$fitted.values)

#Corn
secondStageCornFE <-
  lm(ratioOfEmigrants ~ logCornYieldIV + factor(mexicanState),
     data = dfMXTotEdit2)

secondStageCornPLM <-
```

```
  plm(ratioOfEmigrants ~ logCornYieldIV,
     data = dfMXTotEdit2)

#Corn + Wheat
secondStageCornPlusWheatFE <-
  lm(ratioOfEmigrants ~ logYieldOfCornPlusWheatIV + factor(mexicanState),
     data = dfMXTotEdit2)

secondStageCornPlusWheatPLM <-
  plm(ratioOfEmigrants ~ logYieldOfCornPlusWheatIV,
     data = dfMXTotEdit2)

results <- list(secondStageCornFE, secondStageCornPLM,
                secondStageCornPlusWheatFE, secondStageCornPlusWheatPLM)

map(results, summary)
```

```
## [[1]]
##
## Call:
## lm(formula = ratioOfEmigrants ~ logCornYieldIV + factor(mexicanState),
##     data = dfMXTotEdit2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.04524 -0.01609  0.00000  0.01609  0.04524
##
## Coefficients:
##                                      Estimate Std. Error t value
## (Intercept)                         0.0666476  0.0300121   2.221
## logCornYieldIV                     -0.0392924  0.0474762  -0.828
## factor(mexicanState)Baja California 0.0177661  0.0385331   0.461
## factor(mexicanState)Baja California Sur 0.0060099  0.0557494   0.108
## factor(mexicanState)Campeche       -0.0237673  0.0338386  -0.702
## factor(mexicanState)Chiapas        -0.0100402  0.0309255  -0.325
## factor(mexicanState)Chihuahua      -0.0103522  0.0429487  -0.241
## factor(mexicanState)Coahuila       -0.0068756  0.0322375  -0.213
## factor(mexicanState)Colima         -0.0010543  0.0307562  -0.034
## factor(mexicanState)Distrito Federal -0.0153639  0.0363733  -0.422
## factor(mexicanState)Durango         0.0086391  0.0361962   0.239
## factor(mexicanState)Guanajuato      0.0303668  0.0310600   0.978
## factor(mexicanState)Guerrero        0.0273693  0.0307160   0.891
## factor(mexicanState)Hidalgo         0.0254005  0.0331785   0.766
## factor(mexicanState)Jalisco         0.0157282  0.0320950   0.490
## factor(mexicanState)Mexico          0.0011334  0.0363716   0.031
## factor(mexicanState)Michoacan       0.0625850  0.0326097   1.919
## factor(mexicanState)Morelos         0.0280864  0.0318780   0.881
## factor(mexicanState)Nayarit         0.0285728  0.0308463   0.926
## factor(mexicanState)Nuevo Leon     -0.0221515  0.0310219  -0.714
## factor(mexicanState)Oaxaca          0.0094827  0.0307311   0.309
## factor(mexicanState)Puebla         -0.0154448  0.0314034  -0.492
## factor(mexicanState)Queretaro      -0.0001328  0.0358840  -0.004
## factor(mexicanState)Quintana Roo   -0.0648877  0.0336484  -1.928
## factor(mexicanState)San Luis Potosi 0.0162151  0.0308230   0.526
```

```
## factor(mexicanState)Sinaloa              0.0248368  0.0433602   0.573
## factor(mexicanState)Sonora               0.0238624  0.0511561   0.466
## factor(mexicanState)Tabasco             -0.0246507  0.0308163  -0.800
## factor(mexicanState)Tamaulipas           0.0129548  0.0356286   0.364
## factor(mexicanState)Tlaxcala            -0.0005640  0.0344557  -0.016
## factor(mexicanState)Veracruz            -0.0115185  0.0320222  -0.360
## factor(mexicanState)Yucatan             -0.0362995  0.0307834  -1.179
## factor(mexicanState)Zacatecas            0.0337628  0.0343475   0.983
##                                        Pr(>|t|)
## (Intercept)                              0.0338 *
## logCornYieldIV                           0.4142
## factor(mexicanState)Baja California      0.6480
## factor(mexicanState)Baja California Sur  0.9148
## factor(mexicanState)Campeche             0.4877
## factor(mexicanState)Chiapas              0.7476
## factor(mexicanState)Chihuahua            0.8111
## factor(mexicanState)Coahuila             0.8325
## factor(mexicanState)Colima               0.9729
## factor(mexicanState)Distrito Federal     0.6757
## factor(mexicanState)Durango              0.8129
## factor(mexicanState)Guanajuato           0.3358
## factor(mexicanState)Guerrero             0.3798
## factor(mexicanState)Hidalgo              0.4497
## factor(mexicanState)Jalisco              0.6275
## factor(mexicanState)Mexico               0.9753
## factor(mexicanState)Michoacan            0.0642 .
## factor(mexicanState)Morelos              0.3851
## factor(mexicanState)Nayarit              0.3614
## factor(mexicanState)Nuevo Leon           0.4805
## factor(mexicanState)Oaxaca               0.7597
## factor(mexicanState)Puebla               0.6263
## factor(mexicanState)Queretaro            0.9971
## factor(mexicanState)Quintana Roo         0.0630 .
## factor(mexicanState)San Luis Potosi      0.6026
## factor(mexicanState)Sinaloa              0.5709
## factor(mexicanState)Sonora               0.6441
## factor(mexicanState)Tabasco              0.4298
## factor(mexicanState)Tamaulipas           0.7186
## factor(mexicanState)Tlaxcala             0.9870
## factor(mexicanState)Veracruz             0.7215
## factor(mexicanState)Yucatan              0.2473
## factor(mexicanState)Zacatecas            0.3332
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03071 on 31 degrees of freedom
## Multiple R-squared:  0.5503, Adjusted R-squared:  0.08606
## F-statistic: 1.185 on 32 and 31 DF,  p-value: 0.3188
##
##
## [[2]]
## Oneway (individual) effect Within Model
##
## Call:
```

```
## plm(formula = ratioOfEmigrants ~ logCornYieldIV, data = dfMXTotEdit2)
##
## Balanced Panel: n = 32, T = 2, N = 64
##
## Residuals:
##        Min.     1st Qu.      Median     3rd Qu.        Max.
## -5.1443e-02 -9.9350e-03 -6.9389e-18  9.9350e-03  5.1443e-02
##
## Coefficients:
##                 Estimate Std. Error t-value Pr(>|t|)
## logCornYieldIV -0.022985   0.014109 -1.6291   0.1134
##
## Total Sum of Squares:    0.027974
## Residual Sum of Squares: 0.025767
## R-Squared:      0.078862
## Adj. R-Squared: -0.87199
## F-statistic: 2.65402 on 1 and 31 DF, p-value: 0.11341
##
## [[3]]
##
## Call:
## lm(formula = ratioOfEmigrants ~ logYieldOfCornPlusWheatIV + factor(mexicanState),
##     data = dfMXTotEdit2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.04506 -0.01696  0.00000  0.01696  0.04506
##
## Coefficients:
##                                        Estimate Std. Error t value
## (Intercept)                             0.060726   0.033583   1.808
## logYieldOfCornPlusWheatIV              -0.020344   0.046164  -0.441
## factor(mexicanState)Baja California     0.010111   0.040645   0.249
## factor(mexicanState)Baja California Sur -0.010986   0.057803  -0.190
## factor(mexicanState)Campeche            -0.021610   0.037858  -0.571
## factor(mexicanState)Chiapas             -0.010508   0.031954  -0.329
## factor(mexicanState)Chihuahua           0.001934   0.042082   0.046
## factor(mexicanState)Coahuila            -0.010712   0.032441  -0.330
## factor(mexicanState)Colima              -0.004004   0.031135  -0.129
## factor(mexicanState)Distrito Federal    -0.025312   0.033984  -0.745
## factor(mexicanState)Durango             0.016716   0.035633   0.469
## factor(mexicanState)Guanajuato          0.027652   0.031057   0.890
## factor(mexicanState)Guerrero            0.025363   0.031520   0.805
## factor(mexicanState)Hidalgo             0.019013   0.032259   0.589
## factor(mexicanState)Jalisco             0.011133   0.031753   0.351
## factor(mexicanState)Mexico              -0.009035   0.033778  -0.267
## factor(mexicanState)Michoacan           0.057057   0.031982   1.784
## factor(mexicanState)Morelos             0.022875   0.031238   0.732
## factor(mexicanState)Nayarit             0.026603   0.032516   0.818
## factor(mexicanState)Nuevo Leon          -0.020924   0.031432  -0.666
## factor(mexicanState)Oaxaca              0.007776   0.031558   0.246
## factor(mexicanState)Puebla              -0.015235   0.033149  -0.460
## factor(mexicanState)Queretaro           -0.007685   0.035668  -0.215
## factor(mexicanState)Quintana Roo        -0.062667   0.037288  -1.681
```

```
## factor(mexicanState)San Luis Potosi      0.013894   0.030948   0.449
## factor(mexicanState)Sinaloa              0.010605   0.039909   0.266
## factor(mexicanState)Sonora               0.007333   0.050049   0.147
## factor(mexicanState)Tabasco             -0.025467   0.031671  -0.804
## factor(mexicanState)Tamaulipas           0.003737   0.033574   0.111
## factor(mexicanState)Tlaxcala            -0.009582   0.032199  -0.298
## factor(mexicanState)Veracruz            -0.010388   0.034174  -0.304
## factor(mexicanState)Yucatan             -0.038471   0.032233  -1.194
## factor(mexicanState)Zacatecas            0.039532   0.034753   1.138
##                                        Pr(>|t|)
## (Intercept)                              0.0803 .
## logYieldOfCornPlusWheatIV                0.6625
## factor(mexicanState)Baja California      0.8052
## factor(mexicanState)Baja California Sur  0.8505
## factor(mexicanState)Campeche             0.5722
## factor(mexicanState)Chiapas              0.7445
## factor(mexicanState)Chihuahua            0.9636
## factor(mexicanState)Coahuila             0.7435
## factor(mexicanState)Colima               0.8985
## factor(mexicanState)Distrito Federal     0.4620
## factor(mexicanState)Durango              0.6423
## factor(mexicanState)Guanajuato           0.3801
## factor(mexicanState)Guerrero             0.4272
## factor(mexicanState)Hidalgo              0.5599
## factor(mexicanState)Jalisco              0.7283
## factor(mexicanState)Mexico               0.7909
## factor(mexicanState)Michoacan            0.0842 .
## factor(mexicanState)Morelos              0.4695
## factor(mexicanState)Nayarit              0.4195
## factor(mexicanState)Nuevo Leon           0.5105
## factor(mexicanState)Oaxaca               0.8070
## factor(mexicanState)Puebla               0.6490
## factor(mexicanState)Queretaro            0.8308
## factor(mexicanState)Quintana Roo         0.1029
## factor(mexicanState)San Luis Potosi      0.6566
## factor(mexicanState)Sinaloa              0.7922
## factor(mexicanState)Sonora               0.8845
## factor(mexicanState)Tabasco              0.4275
## factor(mexicanState)Tamaulipas           0.9121
## factor(mexicanState)Tlaxcala             0.7680
## factor(mexicanState)Veracruz             0.7632
## factor(mexicanState)Yucatan              0.2417
## factor(mexicanState)Zacatecas            0.2640
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03095 on 31 degrees of freedom
## Multiple R-squared:  0.5432, Adjusted R-squared:  0.07168
## F-statistic: 1.152 on 32 and 31 DF,  p-value: 0.3476
##
##
## [[4]]
## Oneway (individual) effect Within Model
##
```

```
## Call:
## plm(formula = ratioOfEmigrants ~ logYieldOfCornPlusWheatIV, data = dfMXTotEdit2)
##
## Balanced Panel: n = 32, T = 2, N = 64
##
## Residuals:
##        Min.     1st Qu.      Median     3rd Qu.        Max.
## -5.0632e-02 -1.0901e-02 -9.7578e-19  1.0901e-02  5.0632e-02
##
## Coefficients:
##                            Estimate Std. Error t-value Pr(>|t|)
## logYieldOfCornPlusWheatIV -0.01504    0.01324  -1.136   0.2647
##
## Total Sum of Squares:    0.027974
## Residual Sum of Squares: 0.026856
## R-Squared:      0.039963
## Adj. R-Squared: -0.95104
## F-statistic: 1.29041 on 1 and 31 DF, p-value: 0.26468
```

With both fixed effects and without, we get results that are generally an order of magnitude lower and statistically insignificant, though in the same general direction as the results of the main study.