Assignment 3

ESM 229 - Winter 2021

Minnie Ringland

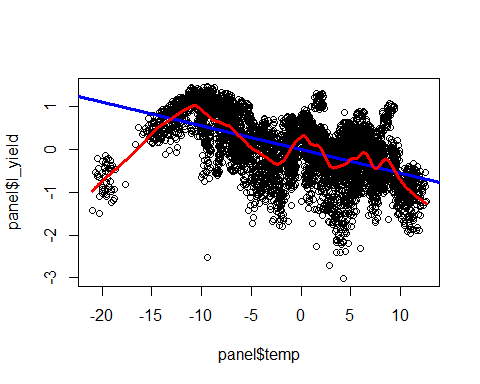
2/1/2021

## Part 1. Exploring the data

For different countries, we have a time series of temperature (deg C) and agricultural yield (kg/hectare).

What is the relationship between temperature and (log) yield?

panel <- read.csv("country\_year\_panel.csv", header=TRUE)  
  
# Overlapping plots  
plot(panel$temp, panel$l\_yield)  
abline(lm(panel$l\_yield~panel$temp), lwd=3, col="blue")  
lines(lowess(panel$temp, panel$l\_yield, f=.1, iter=0), lwd=3, col="red")

 Black circles = observations  
Blue line = linear regression  
Red line = non-parametric local linear regression

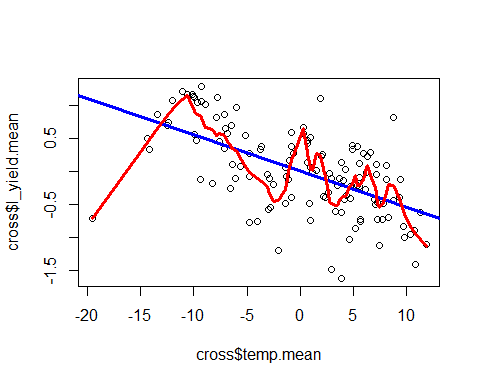
The spread of the data points and the variation in the red line, in addition to the low temperature cluster, show that there may be patterns in the data that a linear approximation will leave out.

## Part 2. Cross-sectional vs. within models: all countries

### Average temperature (~ climate)

We’ll average all observations for a country across all years, and generate a single mean set of values for each country.

# "Collapse" into country means  
cross <- summaryBy(l\_yield+temp ~ cID, FUN=c(mean), data=panel)  
  
# Then plot again  
plot(cross$temp.mean, cross$l\_yield.mean)  
abline(lm(cross$l\_yield.mean~cross$temp.mean), lwd=3, col="blue")  
lines(lowess(cross$temp.mean, cross$l\_yield.mean, f=.1, iter=0), lwd=3, col="red")



Now, the spread of data is reduced and the linear approximation seems more reasonable - the single low temperature observation appears to be an outlier. If we move forward with the assumption of a linear relationship, we can calculate the approximate effect on temperature on yield:

# Calculate coefficient on temperature using the linear model  
linear\_temp\_coeff <- lm(cross$l\_yield.mean~cross$temp.mean)$coefficients[2]

Temperature coefficient using the linear model = *-0.0549831*  
This means we’d expect a 5% reduction (negative relationship) in yield from a 1 degree increase in temperature.

### Cross-sectional estimate with the “within” or fixed effects estimate

Now we will try a model that enables a different type comparison by grouping (by country) and then measuring within-group deviation from the country mean:

# Running a fixed effect regression  
fixed\_effects\_coeff <- plm(l\_yield ~ temp, data=panel, index=c("cID", "year"), model="within")$coefficients

Temperature coefficient using the fixed effect regression = *-0.0546861*  
The expected impact is similar to our previous model - about 5% reduction in yield per degree temperature increase.

These two models have different advantages - the cross-sectional model should capture adaptation to temperature change because we average across time, and the fixed effects model should account for variation in temperature between countries. Since these models agree, we can surmise that most countries have been able to adapt their agricultural techniques to make up for any temperature variations they experienced, and that the limit of the climate effect may be ~5% yield loss.

## Part 3. Cross-sectional vs. within models: continents

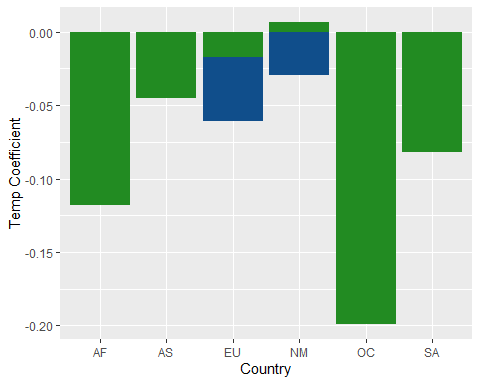
We have reason to believe that climate and weather effects are heterogeneous - they may be different for different parts of the world. So we’ll redo our regressions for each continent to try to isolate geographic patterns.

continents <- c("AF","AS","EU", "OC","SA","NM")  
  
part3 <- data.frame("CrossSectional" = 1:6, "FixedEffects" = 1:6, row.names = continents)  
  
for (i in continents) {  
 continent\_panel <- subset(panel, panel$continent==i, select=c(cID, year, l\_yield, temp, continent))  
 continent\_cross <- summaryBy(l\_yield+temp ~ cID, FUN=c(mean), data=continent\_panel)  
 part3[i,1] <- lm(continent\_cross$l\_yield.mean~continent\_cross$temp.mean)$coefficients[2]  
 part3[i,2] <- plm(l\_yield ~ temp, data=continent\_panel, index=c("cID", "year"), model="within")$coefficients  
}  
  
kable(part3, caption = "Coefficient Comparison Across Continents")

Coefficient Comparison Across Continents

|  |  |  |
| --- | --- | --- |
|  | CrossSectional | FixedEffects |
| AF | -0.0261016 | -0.1181200 |
| AS | -0.0090463 | -0.0451354 |
| EU | -0.0610025 | -0.0168930 |
| OC | -0.0844075 | -0.1989222 |
| SA | -0.0091302 | -0.0816297 |
| NM | -0.0297575 | 0.0067167 |

ggplot(data=part3)+  
 geom\_col(aes(x=row.names(part3), y=CrossSectional), fill="dodgerblue4") +  
 geom\_col(aes(x=row.names(part3), y=FixedEffects), fill="forestgreen") +  
 labs(x = "Country", y = "Temp Coefficient")



Now we see that the coefficients produced by the two models do not agree when we subset by continent. For Africa, Asia, Oceana, and South America, the fixed effects model shows a stronger negative effect on yields. This indicates that adaptation has not been as prevalent or effective, and this makes sense because many of the countries in these continents are already quite hot, and would be unable to adapt to even higher temperatures based on physical or financial limitations.

For North America, the fixed effects model actually gives a slightly positive coefficient, indicating that some countries may actually benefit from higher temperatures in terms of yield. This makes sense because cold parts of Canada and the northern US will actually become more productive as they warm.

We would expect Europe to show a similar pattern to North America, but we still see a negative trend in the fixed effects model. This could be attributable to the high variation in wealth and historical climate across European countries - some northern countries may benefit but the net effect is negative due to southern countries. Additionally, the cross-sectional model for Europe shows higher impact, showing that adaptation may be underestimated by the model.

## Bonus

# Manual fixed effects regression  
# First, de-mean:  
  
manual <- panel %>%   
 group\_by(cID) %>%   
 mutate(l\_yield.mean = mean(l\_yield),  
 temp.mean = mean(temp)) %>%   
 mutate(l\_yield.demean = l\_yield - l\_yield.mean,  
 temp.demean = temp - temp.mean)  
  
manual\_coeff <- lm(manual$l\_yield.demean~manual$temp.demean)$coefficients[2]

Temperature coefficient using R’s fixed effect regression = *-0.0546861* Temperature coefficient using manually coded fixed effects regression = *-0.0546861*

# Then plot again  
plot(manual$temp.demean, manual$l\_yield.demean)  
abline(lm(manual$l\_yield.demean~manual$temp.demean), lwd=3, col="blue")  
lines(lowess(manual$temp.demean, manual$l\_yield.demean, f=.1, iter=0), lwd=3, col="red")

