Regression Discontunuity from Tariffication

Gabriel Domingo

May 2021

Contents

1 Introduction

This paper studies the effects of Tariffication through a Regression Discontinuity in Time Framework. The regression discontinuity concept identifies the effect of a treatment around a cut-off value. Identification is achieved by observations around the "running variable" cut-off value to be similar, except for the exogenous change brought about by the treatment. Traditionally, RD is cross-sectional in nature, with a policy change affecting some agents, consumers and not others due to a characteristic measured by the running variable. Crucially, the affected agents may not be able to influence where they are relative to the cut-off.

A form of Regression Discontinuity is RD in Time. RD in Time is when the running variable is time (Hausman and Rapson 2018). Another factor which differentiates RD in Time is that identification is not cross-sectional, the treatment typically affects the population, and enters into effect at a specific point in time. An example of RD in Time is a paper by Davis (2008) who looked at the effect on car bans on Mexico City's pollution level.

The event we are studying here is the Rice Tariffication law, which removed quantitative restrictions on Rice importation and replaced them with a tariff of 35%. The Month of effectivity is March 2019.[1]

[1]https://www.dof.gov.ph/rice-tariffication-a-gamechanger-in-2019/

"When the RTL or Republic Act (RA) No. 11203 was signed into law by President Duterte last Feb. >14, Dominguez asked the DA four days later to convene the National Food Authority Council (NFAC) > to immediately begin drafting the implementing rules and regulations (IRR) of the law and ensure > that a tariffied regime for rice imports is in place by March 5."

2 Data and Exploration

We begin with writing down the model for the RD in Time. We model the dependent variable, which is at the provincial-month level, as a function of province fixed effects, the dummy for tariffication, which we set at March 2019, and a vector of controls, most importantly variables for a flexible polynomial trend specification. All data used come from PSA online data portal.

$$y_{i,t} = \gamma_i + \gamma_1 1\{\text{Tariffication}_t\} + \gamma_2 x_{i,t} + u_{i,t} \tag{1}$$

First, we measure the missingness of the data. In a RD in time, we would like to run the province regressions with provinces with few missing data on either side of the treatment date.

The table below indicates which ones have more than 8 missing months out of 12 months in the year after tariffication. For the months prior to March 2019, we would also like to identify those with the least missing values for two years prior to tariffication. Of the 24 months, we would like to not include provinces with more than 18 missing.

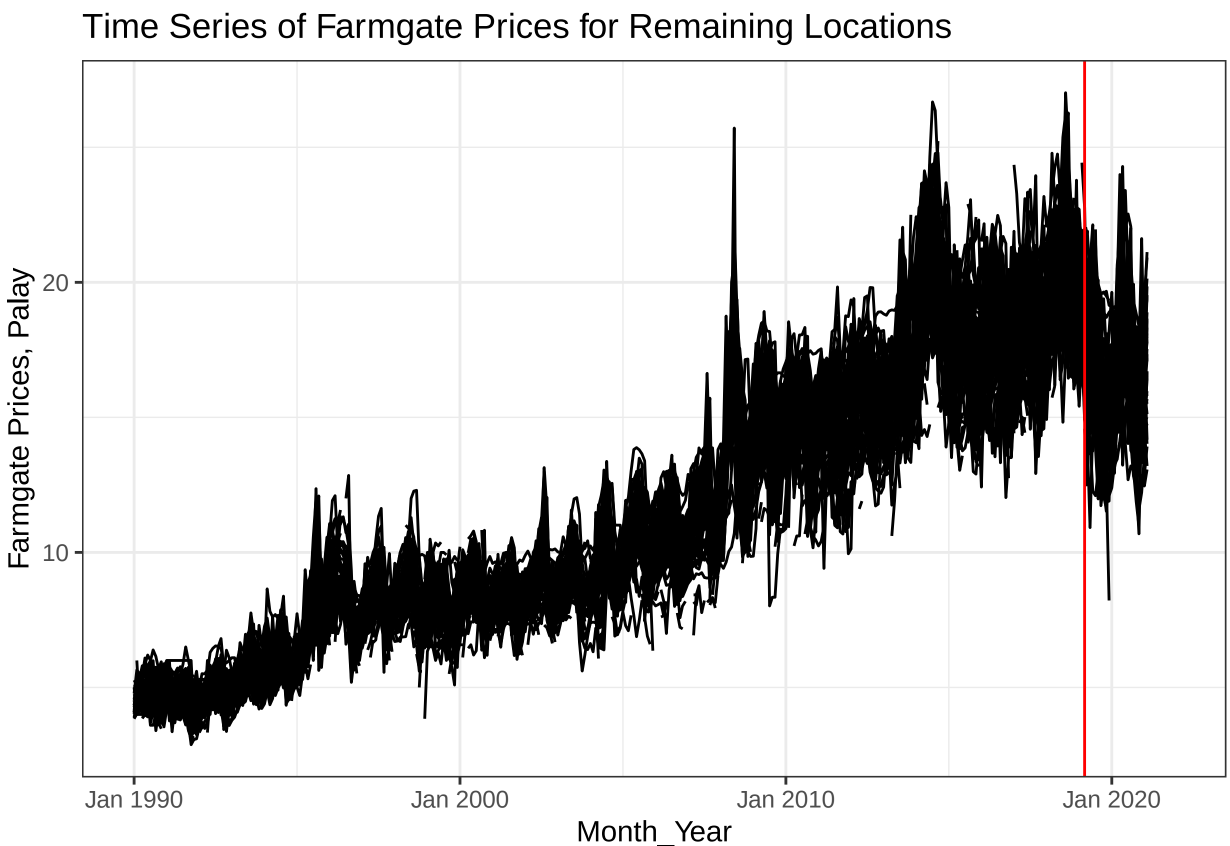
```
##  [1] "Benguet"            "Mountain Province" "Batanes"
##  [4] "Catanduanes"        "Cebu"              "Siquijor"
##  [7] "Camiguin"           "Davao Occidental"  "Dinagat Islands"
## [10] "Basilan"            "Lanao del Sur"     "Sulu"
## [13] "Tawi-Tawi"          "Marinduque"
```

The union of the two lists is excluded from the regression. We list down the provinces with too many missing values.

```
##  [1] "Benguet"            "Mountain Province" "Batanes"
##  [4] "Catanduanes"        "Cebu"              "Siquijor"
##  [7] "Camiguin"           "Davao Occidental"  "Dinagat Islands"
## [10] "Basilan"            "Lanao del Sur"     "Sulu"
## [13] "Tawi-Tawi"          "Marinduque"
```

We continue the examination of the farmgate prices by plotting the data. The red line is the date of the tariffication.

```
require(ggplot2)
ggplot(provp2)+geom_line(aes(x=Month_Year,
                         y=value,group=Geolocation))+
  geom_vline(xintercept = as.yearmon("Mar 2019"),color="red")+
  ylab("Farmgate Prices, Palay")+
  ggtitle("Time Series of Farmgate Prices for Remaining Locations")+
  theme_bw()
```



2.1 Initial Regressions

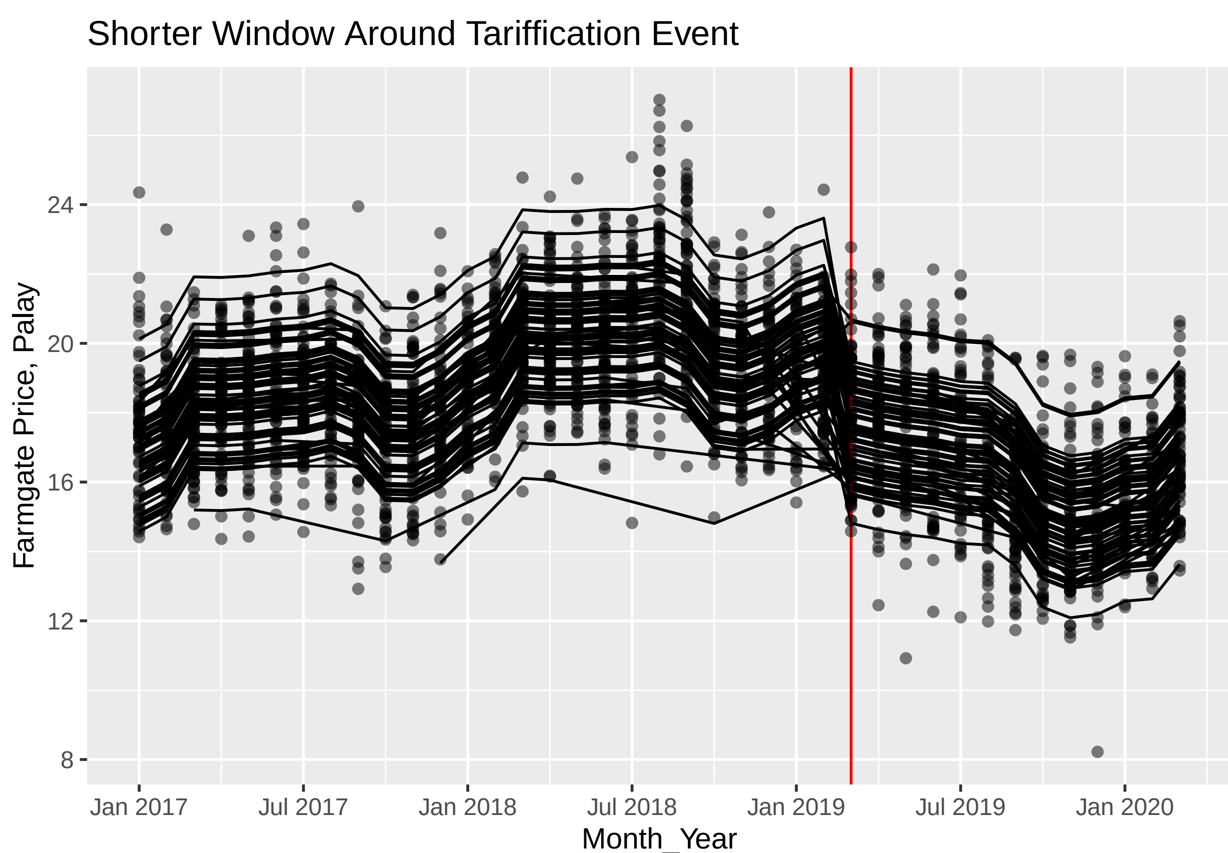
The RD in Time framework's Identification uses the data around the cut-off date of the policy change. Use a global polynomial trend to model farmgate prices, with month dummy variables. We run first a stacked regression, where we estimate dummy variables for each province/city in the dataset, while keeping each province to the same polynomial and month dummies.

```
provp2[,trend:=1:.N,by=.(Geolocation)]
provp2[,treat:=ifelse(Month_Year>as.yearmon("Feb 2019"),1,0)]
provp2[,month:=as.factor(month(provp2$Month_Year))]

provpsr<-provp2[Month_Year>=as.yearmon("Jan 2017") &
```

```
                     Month_Year<as.yearmon("Apr 2020")]

regsr1<-lm(data=provpsr,
   value~treat*Geolocation+I(trend)+I(trend^2)+I(trend^3)+
     factor(month),na.action = "na.exclude")
coef(summary(regsr1))["treat",]
```

```
##      Estimate    Std. Error       t value      Pr(>|t|)
## -3.393894e+00  5.203586e-01 -6.522222e+00  8.527654e-11
```

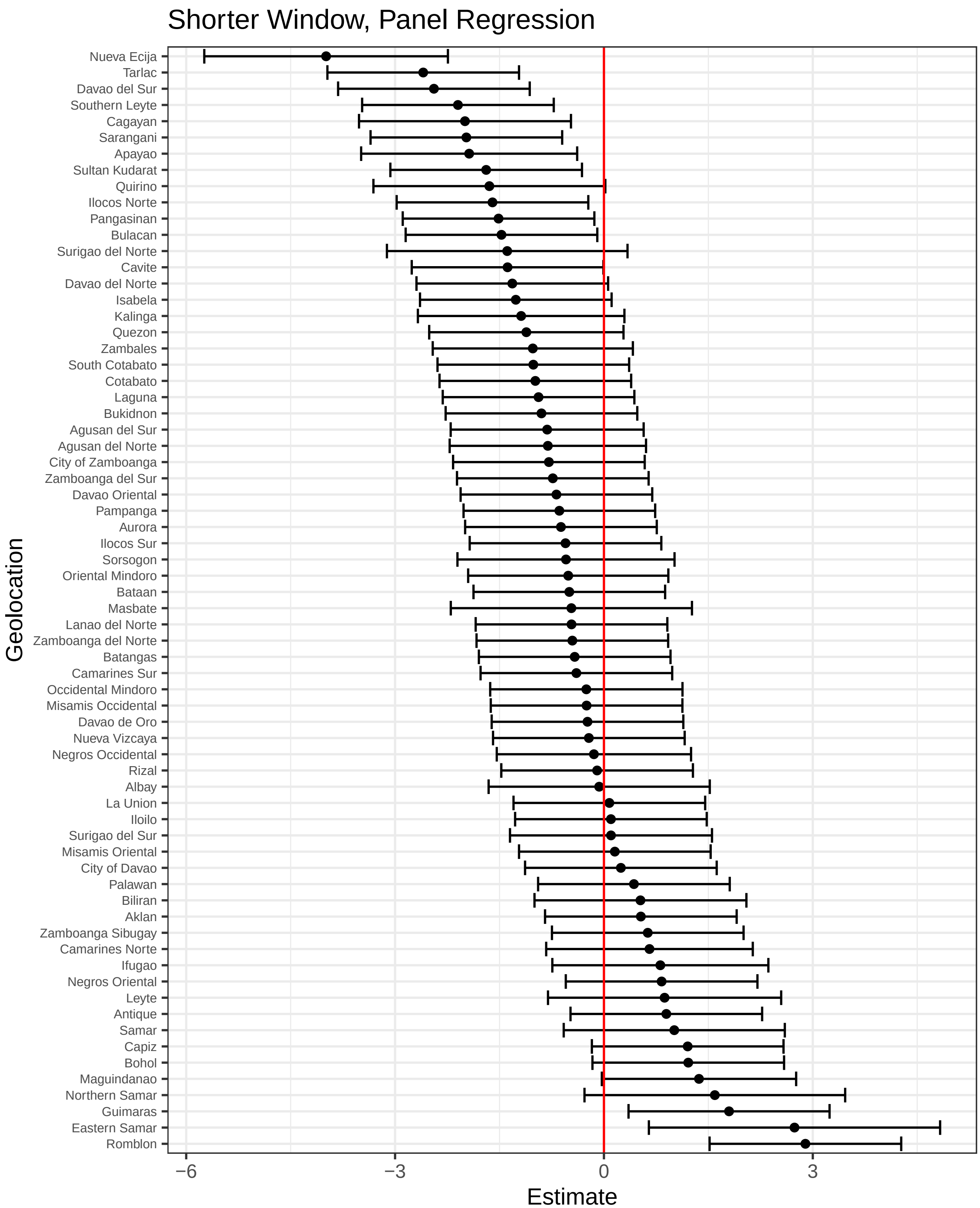
```
provpsr$predregsr1<-predict(regsr1)
ggplot(provpsr[!is.na(value)])+
  geom_line(aes(x=Month_Year,y=predregsr1,group=Geolocation))+
    geom_vline(xintercept = as.yearmon("Mar 2019"),color="red")+
  geom_point(aes(x=Month_Year,y=value,group=Geolocation),alpha=.5)+
  ylab("Farmgate Price, Palay")+
  ggtitle("Shorter Window Around Tariffication Event")
```



2.2 Extracting the Province Tariffication Effect

We extract the Tariffication Effect for the provinces, including the standard errors.

```
regsr1_treat<-data.table(Geolocation=rownames(coef(summary(regsr1))),
                    coef(summary(regsr1))[,1:2])
regsr1_treat<-regsr1_treat[grep("treat:Geolocation",Geolocation),]
regsr1_treat$Geolocation<-gsub(regsr1_treat$Geolocation,
                        pattern="treat:Geolocation",replace="")

regsr1_treat<-regsr1_treat[order(-Estimate)][
  ,Geolocation:=factor(Geolocation,levels = Geolocation)]

ggplot(regsr1_treat)+geom_point(aes(x=Geolocation,y=Estimate))+
  geom_errorbar(aes(x=Geolocation,ymin=Estimate-1.96*`Std. Error`,
                ymax=Estimate+1.96*`Std. Error`))+
  geom_hline(yintercept=0,color="red")+theme_bw()+
  coord_flip()+theme(axis.text.y=element_text(size=6))+
  ggtitle("Shorter Window, Panel Regression")
```



The short-run coefficients are relative to Abra, whose effect is -3.394 and a standard error of 0.52. One issue with the panel fixed effects regression is that the time trend polynomial is fit to be common across provinces, which may not be efficient. It would be better to regress by province.

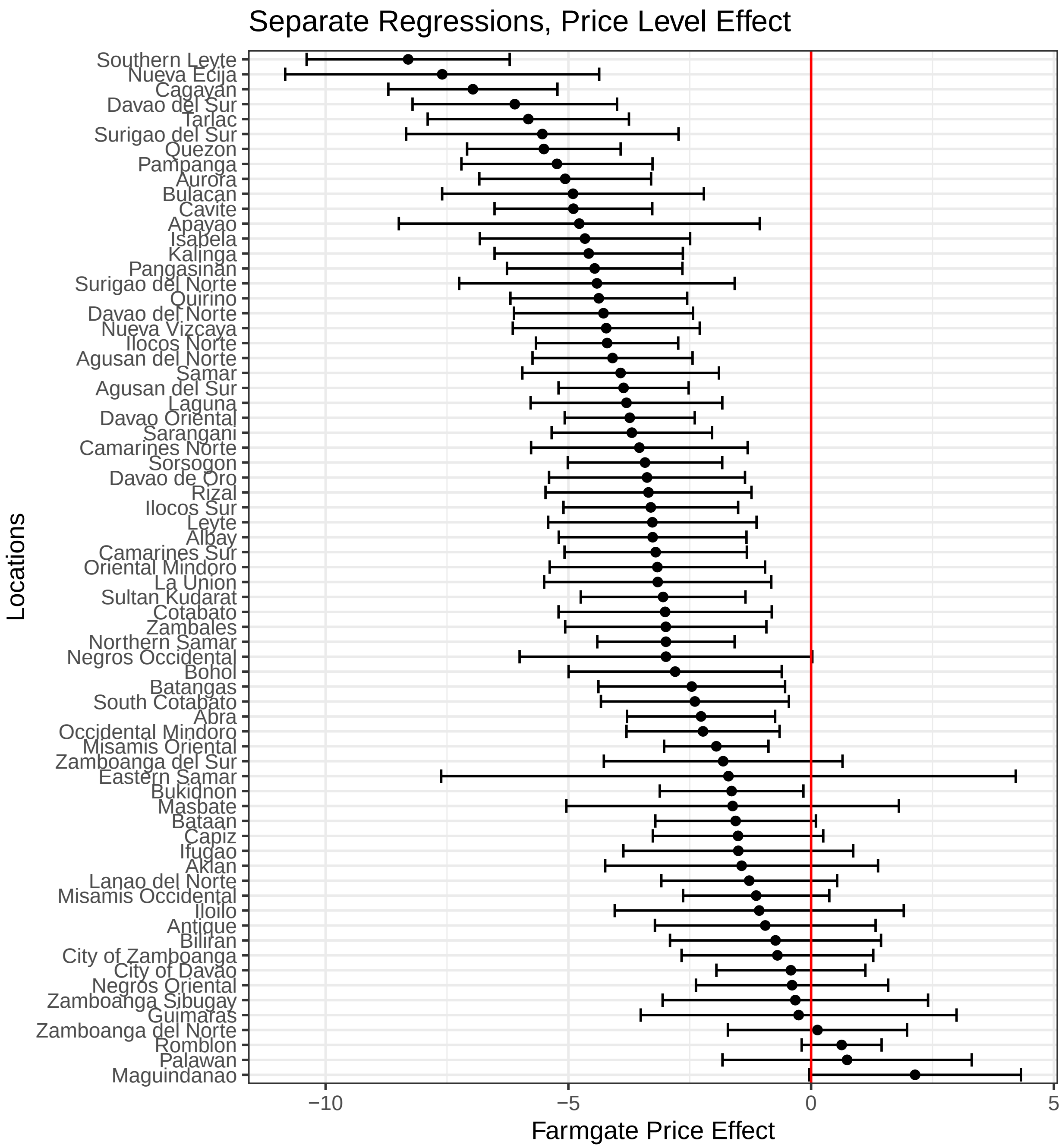
2.3 Regress each Location separately

A better approach would be to regress each constant separately, where the constant would not be constrained by different trend coefficients. We use our preferred short window specification, a linear trend and no month dummies.[2]

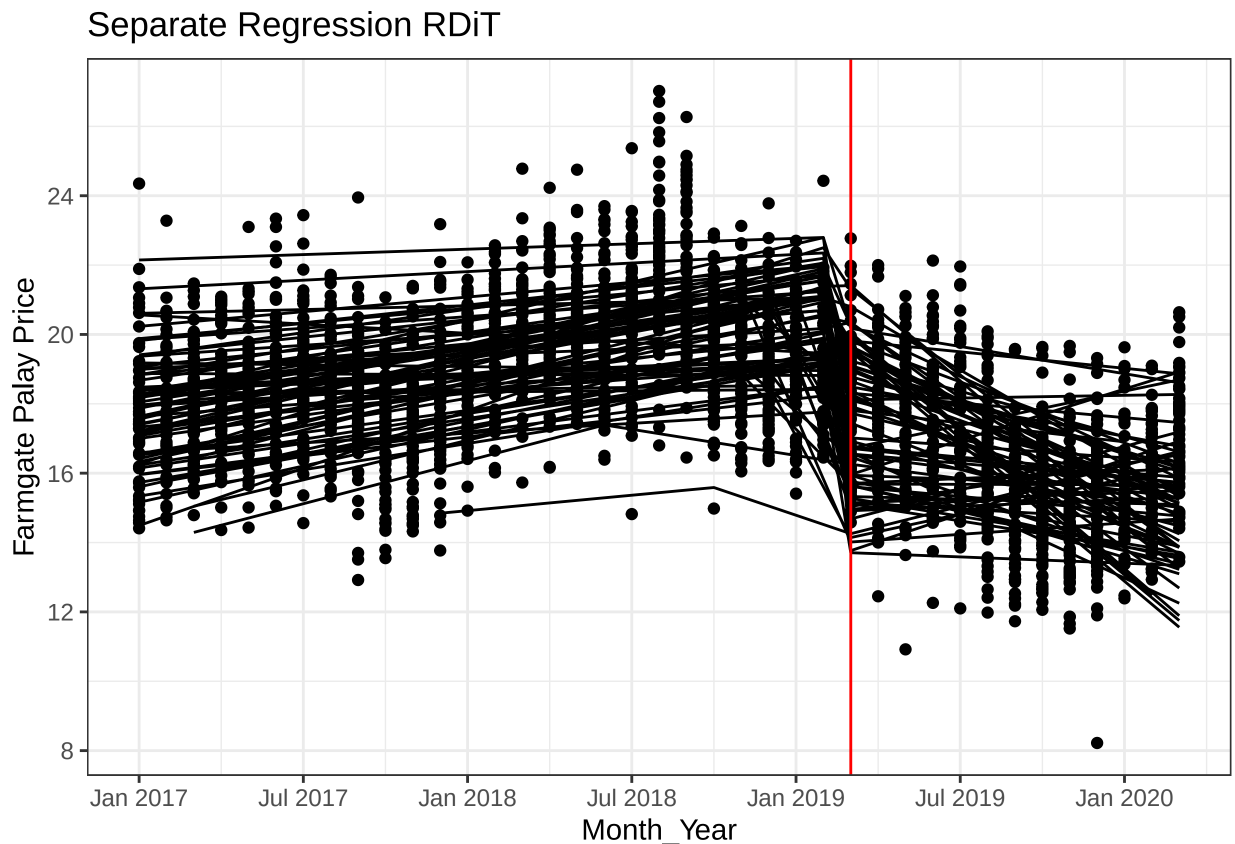
Before we conduct the regression, we set the treatment to be time =0 so that the coefficient on the treatment would simply be: $\gamma_1 + \gamma_1 * (time = 0) = \gamma_1$ when treatment=1. We also interact with the linear trend, which allows for the possibility of change in the trend after tariffication. Hence, we will have two tariffication effects by province. The main effect and the time trend interaction. Looking at the time series plot of palay, we know that the rising price trend was reversed to a falling trend post-tariffication.

```
provpsr[Month_Year>=as.yearmon("Mar 2019"),
        trend:=(1:.N)-1,by=.(Geolocation)]
provpsr[Month_Year<as.yearmon("Mar 2019"),
        trend:=-1*(.N:1),by=.(Geolocation)]

regres<-provpsr[,c(MissObs=sum(is.na(value)),
    reg1=as.list(coef(lm(value ~treat*I(trend)))[c(2,4)]),
    reg1se=as.list(sqrt(diag(vcov(lm(value ~ treat*I(trend))))[c(2,4)]))),
    by=Geolocation]

regres<-regres[order(-reg1.treat)][
  ,Geolocation:=factor(Geolocation,levels = Geolocation)]


ggplot(regres)+
  geom_point(aes(x=Geolocation,y=reg1.treat))+
  geom_errorbar(data=regres,aes(x=Geolocation,
  geom_hline(yintercept=0,color="red")+
  theme(axis.text.x = element_text(size=6))+
  theme_bw()+coord_flip()+ylab("Farmgate Price Effect")+
  xlab("Locations")+
  ggtitle("Separate Regressions, Price Level Effect")
```

[2]Month dummies will not be efficiently estimated as each province will only have 3 unique months.



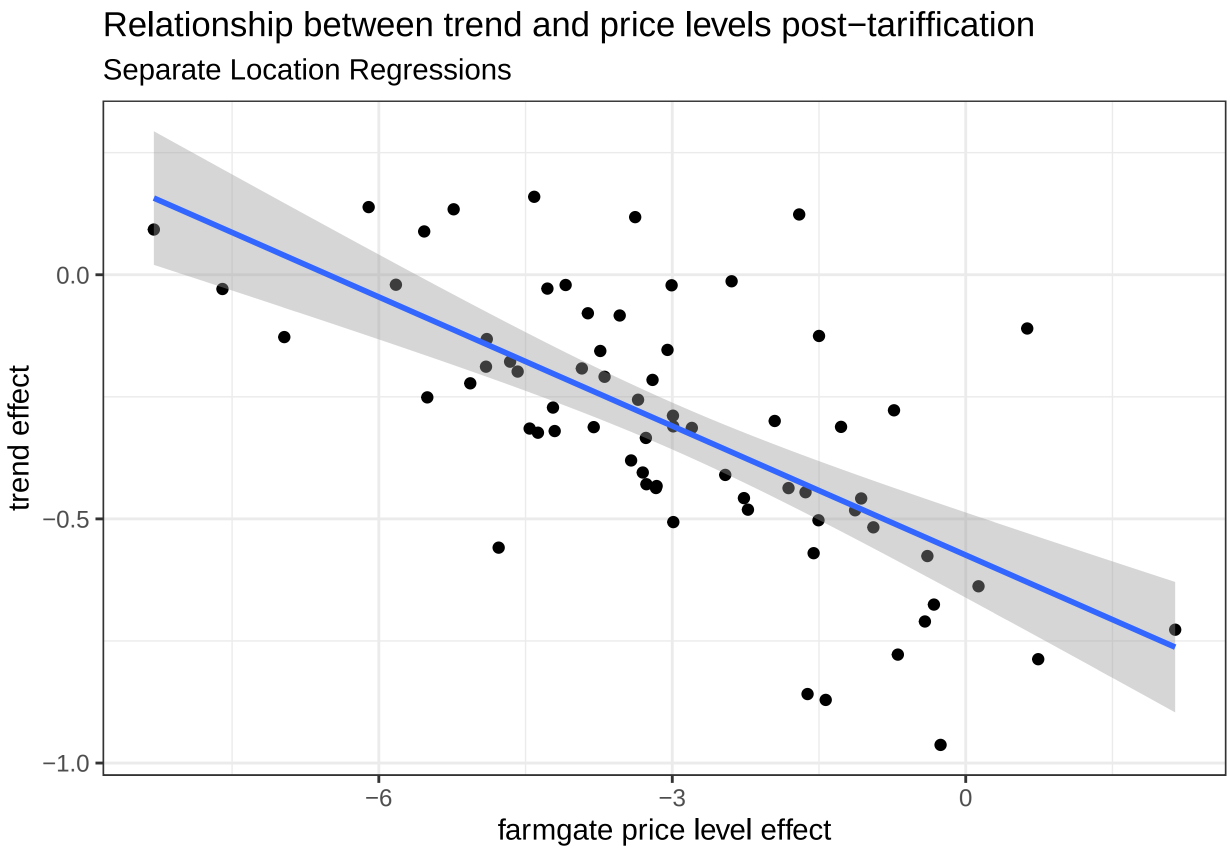

```
provpsr[,reg1.pred:=predict(lm(value ~treat*I(trend),
                na.action="na.exclude")), by=Geolocation]

ggplot(provpsr[ !is.na(value)])+
  geom_line(aes(x=Month_Year,y=reg1.pred,group=Geolocation))+
  geom_point(aes(x=Month_Year,y=value))+
  geom_vline(xintercept = as.yearmon("Mar 2019"),color="red")+
  theme_bw()+ylab("Farmgate Palay Price")+
  ggtitle("Separate Regression RDiT")
```



The relationship between the price level effect and the trend effect of tariffication is interesting. Generally, both happened to each province. However, there is a negative relationship between the two reactions. This implies that while the initial drop in prices may not be immediate, the price trend shifts negatively. For those with the large initial drop, the fall in the trend is less apparent, or practically zero; and vice-versa.

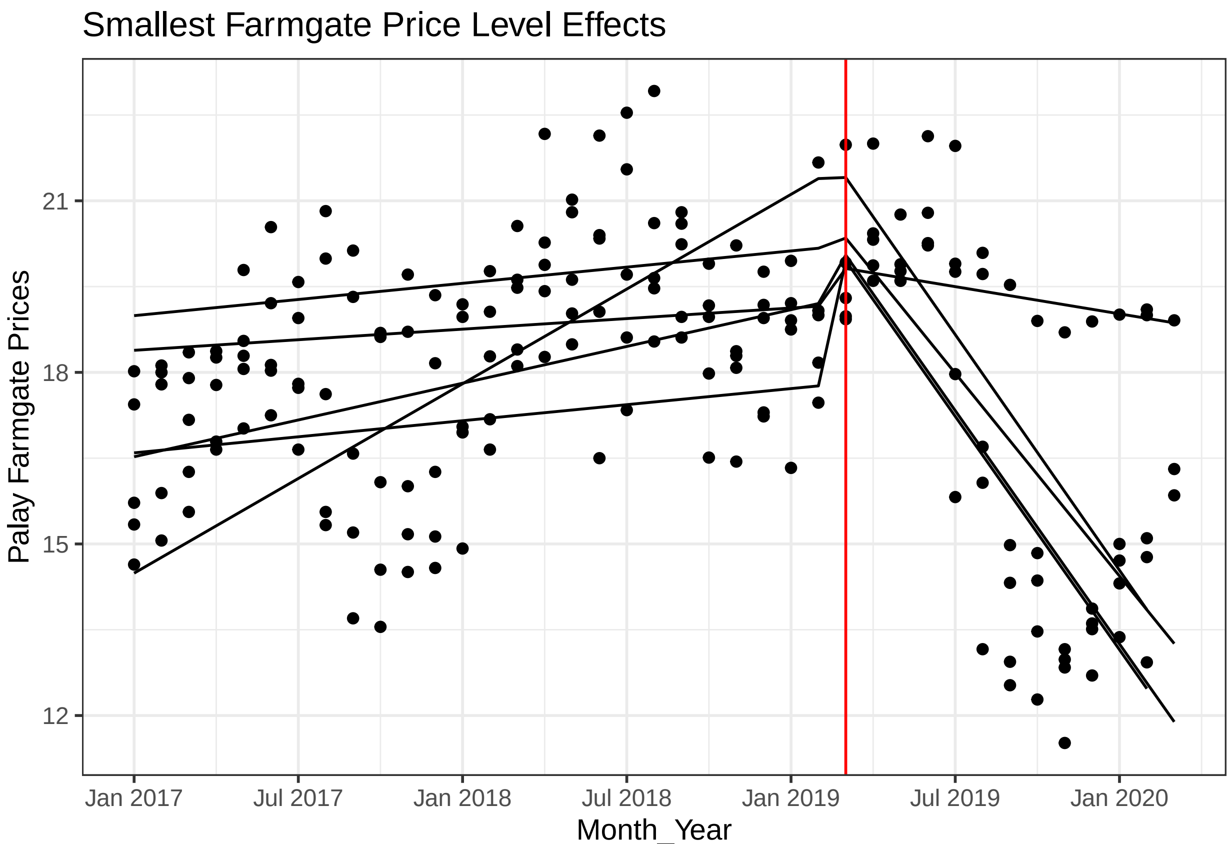
```
ggplot(regres)+geom_point(aes(x=reg1.treat,y=`reg1.treat:I(trend)`))+
  ylab("trend effect")+
  xlab("farmgate price level effect")+
  ggtitle("Relationship between trend and price levels post-tariffication",
          subtitle="Separate Location Regressions")+
  geom_smooth(aes(x=reg1.treat,y=`reg1.treat:I(trend)`),method="lm")+theme_bw()
```

```
## `geom_smooth()` using formula 'y ~ x'
```

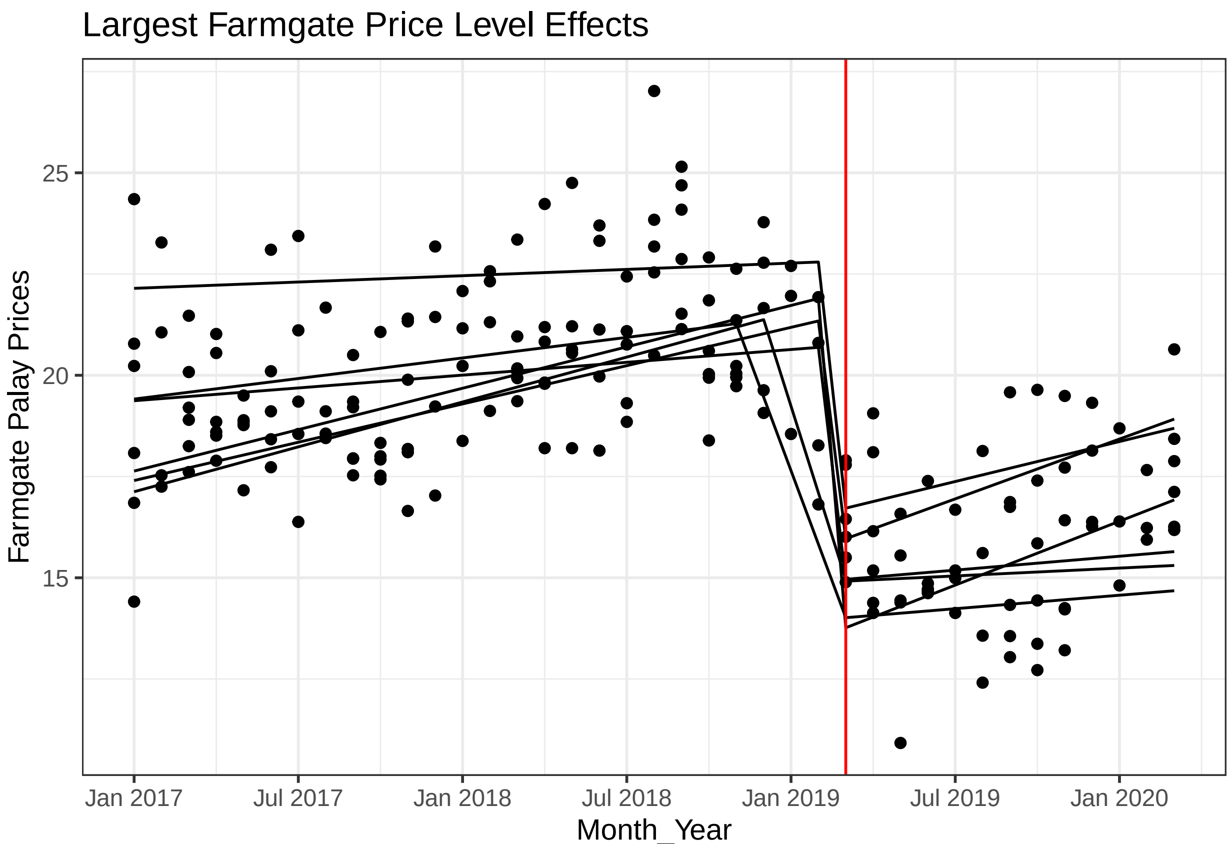


Let us identify the provinces with large treatment effects and/or large se's relative to the treatment effect (the so-called "statistically insignificant estimates"). We can plot those provinces directly.

```
#small level effects
smallevels<-regres$Geolocation[1:5]

ggplot(provpsr[ !is.na(value) & Geolocation %in% smallevels])+
  geom_line(aes(x=Month_Year,y=reg1.pred,group=Geolocation))+
  geom_point(aes(x=Month_Year,y=value))+
  geom_vline(xintercept = as.yearmon("Mar 2019"),color="red")+
  ggtitle("Smallest Farmgate Price Level Effects")+
  ylab("Palay Farmgate Prices")+
  theme_bw()
```




```
largelevel<-tail(regres)$Geolocation

ggplot(provpsr[ !is.na(value) & Geolocation %in% largelevel])+
  geom_line(aes(x=Month_Year,y=reg1.pred,group=Geolocation))+
  geom_point(aes(x=Month_Year,y=value))+    geom_vline(xintercept = as.yearmon("Mar 2019
  ggtitle("Largest Farmgate Price Level Effects")+
  theme_bw()
```



Understanding the heterogenous effects of tariffication over provinces is an important topic. On the one hand, tariffication is a shift down in palay demand, and its effects would be immediate for markets which have easy access to imported rice, or those with large demand for palay. On the other, a delayed effect on palay price by tariffication means that the it takes time for the lower of prices, and prices would tend to fall over the whole year of 2019.

References

Davis, Lucas W. 2008. “The Effect of Driving Restrictions on Air Quality in Mexico City.” *Journal of Political Economy* 116(1): 38–81.

Hausman, Catherine, and David Rapson. 2018. “Regression Discontinuity in Time: Considerations for Empirical Applications.” *Annual Review of Resource Economics* 10: 533–52.