Bài Kiểm Tra KTL\_II

Trương Ngọc Thuỳ Trang

2023-10-06

library(foreign)  
library(carData)  
library(car)  
library(lmtest)

## Loading required package: zoo

##   
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

library(AER)

## Loading required package: sandwich

## Loading required package: survival

library(plm)  
library(stargazer)

##   
## Please cite as:

## Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables.

## R package version 5.2.3. https://CRAN.R-project.org/package=stargazer

library(readxl)  
library(ggplot2)  
library(gplots)

##   
## Attaching package: 'gplots'

## The following object is masked from 'package:stats':  
##   
## lowess

library(caret)

## Loading required package: lattice

##   
## Attaching package: 'caret'

## The following object is masked from 'package:survival':  
##   
## cluster

library(caTools)  
library(fBasics)

##   
## Attaching package: 'fBasics'

## The following object is masked from 'package:car':  
##   
## densityPlot

## Nhập data

PCI2020\_2022 <- read\_excel("E:/KTL/KTL\_02/PCI/PCI2020\_2022.xlsx")  
PCI2020\_2022

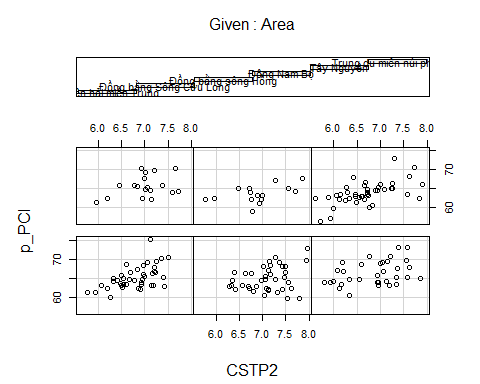
## # A tibble: 189 × 13  
## City p\_PCI CSTP1 CSTP2 CSTP3 CSTP4 CSTP5 CSTP6 CSTP7 CSTP8 CSTP9 CSTP10  
## <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 An Giang 62.4 6.77 7.08 6.19 7.52 7.05 4.84 6.60 5.24 5.05 6.86  
## 2 BRVT 70.3 6.95 7.66 5.93 7.81 7.79 6.99 7.21 6.03 6.31 7.41  
## 3 Bắc Giang 72.8 6.65 7.29 6.30 7.38 8.02 7.31 7.62 6.10 6.80 8.60  
## 4 Bắc Kạn 65.2 6.36 6.28 6.55 7.97 7.24 6.23 7.37 4.59 5.38 8.02  
## 5 Bạc Liêu 60.4 7.34 7.03 4.71 6.46 6.65 4.19 6.27 5.05 4.92 8.02  
## 6 Bắc Ninh 69.1 7.13 7.05 6.59 8.54 6.83 6.93 6.23 6.48 7.57 7.09  
## 7 Bến Tre 68.0 7.26 7.49 5.25 8.03 7.97 6.65 6.88 5.43 5.40 7.83  
## 8 Bình Định 66.7 7.25 7.21 5.97 6.83 6.79 7.36 6.86 5.39 5.46 8.04  
## 9 Bình Dương 65.1 6.32 7.06 6.67 7.35 6.92 6.07 6.32 6.13 5.93 6.84  
## 10 Bình Phước 64.3 7.07 7.12 6.37 7.37 6.39 5.10 7.20 5.82 5.09 7.49  
## # ℹ 179 more rows  
## # ℹ 1 more variable: Area <chr>

## tạo thêm cột thời gian

t22= c(rep(1,63),rep(0,126))  
t21=c(rep(0,63),rep(1,63),rep(0,63))  
time=c(rep(2022,63),rep(2021,63),rep(2020,63))  
#time  
PCI=cbind(PCI2020\_2022,time,t22,t21)  
View(PCI)  
dim(PCI)

## [1] 189 16

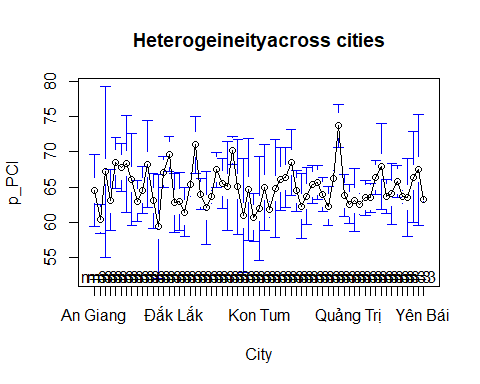
library(graphics)  
coplot(p\_PCI~CSTP2|Area,type="p",data = PCI)



plotmeans(p\_PCI ~ City, main="Heterogeineityacross cities", data=PCI)

## Warning in arrows(x, li, x, pmax(y - gap, li), col = barcol, lwd = lwd, :  
## zero-length arrow is of indeterminate angle and so skipped  
  
## Warning in arrows(x, li, x, pmax(y - gap, li), col = barcol, lwd = lwd, :  
## zero-length arrow is of indeterminate angle and so skipped  
  
## Warning in arrows(x, li, x, pmax(y - gap, li), col = barcol, lwd = lwd, :  
## zero-length arrow is of indeterminate angle and so skipped  
  
## Warning in arrows(x, li, x, pmax(y - gap, li), col = barcol, lwd = lwd, :  
## zero-length arrow is of indeterminate angle and so skipped  
  
## Warning in arrows(x, li, x, pmax(y - gap, li), col = barcol, lwd = lwd, :  
## zero-length arrow is of indeterminate angle and so skipped  
  
## Warning in arrows(x, li, x, pmax(y - gap, li), col = barcol, lwd = lwd, :  
## zero-length arrow is of indeterminate angle and so skipped  
  
## Warning in arrows(x, li, x, pmax(y - gap, li), col = barcol, lwd = lwd, :  
## zero-length arrow is of indeterminate angle and so skipped

## Warning in arrows(x, ui, x, pmin(y + gap, ui), col = barcol, lwd = lwd, :  
## zero-length arrow is of indeterminate angle and so skipped  
  
## Warning in arrows(x, ui, x, pmin(y + gap, ui), col = barcol, lwd = lwd, :  
## zero-length arrow is of indeterminate angle and so skipped  
  
## Warning in arrows(x, ui, x, pmin(y + gap, ui), col = barcol, lwd = lwd, :  
## zero-length arrow is of indeterminate angle and so skipped  
  
## Warning in arrows(x, ui, x, pmin(y + gap, ui), col = barcol, lwd = lwd, :  
## zero-length arrow is of indeterminate angle and so skipped  
  
## Warning in arrows(x, ui, x, pmin(y + gap, ui), col = barcol, lwd = lwd, :  
## zero-length arrow is of indeterminate angle and so skipped  
  
## Warning in arrows(x, ui, x, pmin(y + gap, ui), col = barcol, lwd = lwd, :  
## zero-length arrow is of indeterminate angle and so skipped  
  
## Warning in arrows(x, ui, x, pmin(y + gap, ui), col = barcol, lwd = lwd, :  
## zero-length arrow is of indeterminate angle and so skipped



scatterplot(p\_PCI~time|Area, boxplots=FALSE, smooth=TRUE, reg.line=FALSE, data=PCI)

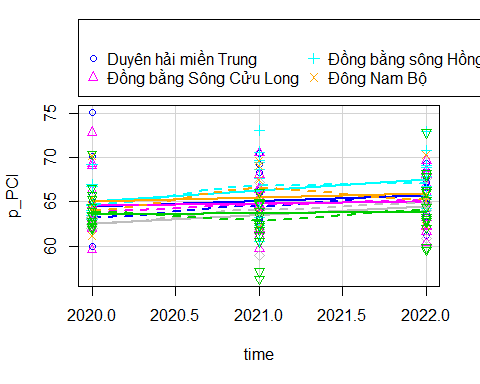
## Warning in plot.window(...): "reg.line" is not a graphical parameter

## Warning in plot.xy(xy, type, ...): "reg.line" is not a graphical parameter

## Warning in axis(side = side, at = at, labels = labels, ...): "reg.line" is not  
## a graphical parameter  
  
## Warning in axis(side = side, at = at, labels = labels, ...): "reg.line" is not  
## a graphical parameter

## Warning in box(...): "reg.line" is not a graphical parameter

## Warning in title(...): "reg.line" is not a graphical parameter



## Panel Model

## 1.Uoc luong gop POLS

pols\_pci=plm(p\_PCI~CSTP1+CSTP2+CSTP3+CSTP4+CSTP5+CSTP6+CSTP7+CSTP8+CSTP9+CSTP10+t22+t21,data=PCI,index = c("City","time"),model = ("pooling"))  
summary(pols\_pci)

## Pooling Model  
##   
## Call:  
## plm(formula = p\_PCI ~ CSTP1 + CSTP2 + CSTP3 + CSTP4 + CSTP5 +   
## CSTP6 + CSTP7 + CSTP8 + CSTP9 + CSTP10 + t22 + t21, data = PCI,   
## model = ("pooling"), index = c("City", "time"))  
##   
## Balanced Panel: n = 63, T = 3, N = 189  
##   
## Residuals:  
## Min. 1st Qu. Median 3rd Qu. Max.   
## -2.225622 -0.277975 0.010008 0.344774 2.215700   
##   
## Coefficients:  
## Estimate Std. Error t-value Pr(>|t|)   
## (Intercept) 2.348817 1.178208 1.9936 0.0477458 \*   
## CSTP1 0.477706 0.087794 5.4412 1.753e-07 \*\*\*  
## CSTP2 0.651396 0.108164 6.0223 9.780e-09 \*\*\*  
## CSTP3 1.560490 0.091569 17.0418 < 2.2e-16 \*\*\*  
## CSTP4 0.235490 0.088149 2.6715 0.0082599 \*\*   
## CSTP5 1.334059 0.104326 12.7874 < 2.2e-16 \*\*\*  
## CSTP6 0.398514 0.065167 6.1152 6.062e-09 \*\*\*  
## CSTP7 0.766604 0.104857 7.3109 8.951e-12 \*\*\*  
## CSTP8 1.960413 0.072105 27.1885 < 2.2e-16 \*\*\*  
## CSTP9 1.692568 0.064341 26.3061 < 2.2e-16 \*\*\*  
## CSTP10 0.566645 0.106008 5.3453 2.772e-07 \*\*\*  
## t22 2.290570 0.184067 12.4442 < 2.2e-16 \*\*\*  
## t21 -0.648020 0.183607 -3.5294 0.0005316 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Total Sum of Squares: 1802.7  
## Residual Sum of Squares: 66.023  
## R-Squared: 0.96338  
## Adj. R-Squared: 0.96088  
## F-statistic: 385.804 on 12 and 176 DF, p-value: < 2.22e-16

### =>có thể thấy các hệ số ước lượng đều có ý nghĩa thống kê

### Kiểm định GT : H0:PSSS mô hình thay đổi

bptest(pols\_pci)

##   
## studentized Breusch-Pagan test  
##   
## data: pols\_pci  
## BP = 46.127, df = 12, p-value = 6.598e-06

### Do p\_value<0.05 nên bác bỏ H0. Vậy với mức ý nghĩa 5%, pols có PSSS thay đổi

### =>các UL là vững những không hiệu quả bởi vì có tương quan chéo giữa ssnn với 1 thực thể đã cho ở thời điểm khác nhau

vif(pols\_pci)# hệ số phóng đại phương sai

## CSTP1 CSTP2 CSTP3 CSTP4 CSTP5 CSTP6 CSTP7 CSTP8   
## 1.697790 1.582941 1.192781 2.087885 2.283452 1.537850 2.010906 1.555472   
## CSTP9 CSTP10 t22 t21   
## 1.463929 2.317541 3.793326 3.774392

### =>có đa cộng tuyến giữa các biến độc lập

## 2.Uoc luong mo hinh tac dong co dinh- FE

fem\_pci=plm(p\_PCI~CSTP1+CSTP2+CSTP3+CSTP4+CSTP5+CSTP6+CSTP7+CSTP8+CSTP9+CSTP10+factor(time),data=PCI,index = c("City","time"),model = "within")  
summary(fem\_pci)

## Oneway (individual) effect Within Model  
##   
## Call:  
## plm(formula = p\_PCI ~ CSTP1 + CSTP2 + CSTP3 + CSTP4 + CSTP5 +   
## CSTP6 + CSTP7 + CSTP8 + CSTP9 + CSTP10 + factor(time), data = PCI,   
## model = "within", index = c("City", "time"))  
##   
## Balanced Panel: n = 63, T = 3, N = 189  
##   
## Residuals:  
## Min. 1st Qu. Median 3rd Qu. Max.   
## -1.653360 -0.310157 -0.010044 0.291652 1.643608   
##   
## Coefficients:  
## Estimate Std. Error t-value Pr(>|t|)   
## CSTP1 0.49596 0.12181 4.0715 8.652e-05 \*\*\*  
## CSTP2 0.71220 0.15126 4.7085 7.082e-06 \*\*\*  
## CSTP3 1.59918 0.15129 10.5701 < 2.2e-16 \*\*\*  
## CSTP4 0.13010 0.13599 0.9567 0.3407519   
## CSTP5 1.40945 0.15945 8.8392 1.384e-14 \*\*\*  
## CSTP6 0.29288 0.08515 3.4396 0.0008148 \*\*\*  
## CSTP7 0.82635 0.15435 5.3536 4.522e-07 \*\*\*  
## CSTP8 2.02635 0.12250 16.5414 < 2.2e-16 \*\*\*  
## CSTP9 1.50251 0.15192 9.8902 < 2.2e-16 \*\*\*  
## CSTP10 0.36610 0.16461 2.2240 0.0281217 \*   
## factor(time)2021 -0.92226 0.23988 -3.8446 0.0001992 \*\*\*  
## factor(time)2022 2.10935 0.26093 8.0840 7.376e-13 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Total Sum of Squares: 431.92  
## Residual Sum of Squares: 48.573  
## R-Squared: 0.88754  
## Adj. R-Squared: 0.81454  
## F-statistic: 74.9759 on 12 and 114 DF, p-value: < 2.22e-16

fem\_pci1=plm(p\_PCI~CSTP1+CSTP2+CSTP3+CSTP4+CSTP5+CSTP6+CSTP7+CSTP8+CSTP9+CSTP10,data=PCI,index = c("City","time"),model = "within")  
stargazer(fem\_pci,fem\_pci1,type = "text")

##   
## ==================================================================  
## Dependent variable:   
## -------------------------------------------------  
## p\_PCI   
## (1) (2)   
## ------------------------------------------------------------------  
## CSTP1 0.496\*\*\* 0.642\*\*\*   
## (0.122) (0.196)   
##   
## CSTP2 0.712\*\*\* 0.647\*\*   
## (0.151) (0.262)   
##   
## CSTP3 1.599\*\*\* 1.657\*\*\*   
## (0.151) (0.269)   
##   
## CSTP4 0.130 0.254   
## (0.136) (0.235)   
##   
## CSTP5 1.409\*\*\* 1.292\*\*\*   
## (0.159) (0.275)   
##   
## CSTP6 0.293\*\*\* 0.045   
## (0.085) (0.142)   
##   
## CSTP7 0.826\*\*\* 0.168   
## (0.154) (0.263)   
##   
## CSTP8 2.026\*\*\* 0.840\*\*\*   
## (0.123) (0.167)   
##   
## CSTP9 1.503\*\*\* 0.831\*\*\*   
## (0.152) (0.222)   
##   
## CSTP10 0.366\*\* 1.570\*\*\*   
## (0.165) (0.255)   
##   
## factor(time)2021 -0.922\*\*\*   
## (0.240)   
##   
## factor(time)2022 2.109\*\*\*   
## (0.261)   
##   
## ------------------------------------------------------------------  
## Observations 189 189   
## R2 0.888 0.638   
## Adjusted R2 0.815 0.413   
## F Statistic 74.976\*\*\* (df = 12; 114) 20.442\*\*\* (df = 10; 116)  
## ==================================================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

### Kiểm tra ảnh hưởng của thời gian tới mô hình

### GT H0:Mô hình fem\_pci1 phù hợp

pFtest(fem\_pci,fem\_pci1)

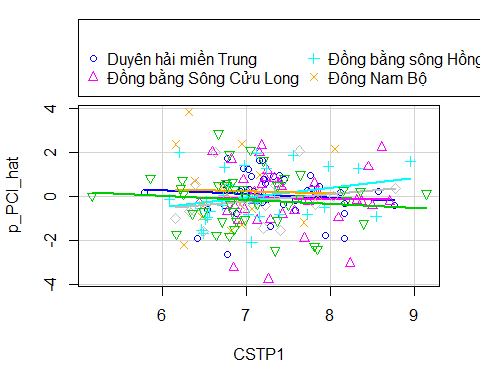
##   
## F test for individual effects  
##   
## data: p\_PCI ~ CSTP1 + CSTP2 + CSTP3 + CSTP4 + CSTP5 + CSTP6 + CSTP7 + ...  
## F = 126.5, df1 = 2, df2 = 114, p-value < 2.2e-16  
## alternative hypothesis: significant effects

### do p\_value < 0.05 nên bb H0. Với mức ý nghĩa 5%, mô hình fem\_pci có sự tác động của thời gian phù hợp

fem\_pci2=plm(p\_PCI~CSTP1+CSTP2+CSTP3+CSTP4+CSTP5+CSTP6+CSTP7+CSTP8+CSTP9+CSTP10+factor(City),data=PCI,index = c("City","time"),model = "within")  
summary(fem\_pci2)

## Oneway (individual) effect Within Model  
##   
## Call:  
## plm(formula = p\_PCI ~ CSTP1 + CSTP2 + CSTP3 + CSTP4 + CSTP5 +   
## CSTP6 + CSTP7 + CSTP8 + CSTP9 + CSTP10 + factor(City), data = PCI,   
## model = "within", index = c("City", "time"))  
##   
## Balanced Panel: n = 63, T = 3, N = 189  
##   
## Residuals:  
## Min. 1st Qu. Median 3rd Qu. Max.   
## -2.330747 -0.605381 -0.038437 0.612174 2.532854   
##   
## Coefficients:  
## Estimate Std. Error t-value Pr(>|t|)   
## CSTP1 0.642255 0.196417 3.2699 0.001417 \*\*   
## CSTP2 0.647138 0.262485 2.4654 0.015149 \*   
## CSTP3 1.657243 0.268663 6.1685 1.038e-08 \*\*\*  
## CSTP4 0.254431 0.234810 1.0836 0.280807   
## CSTP5 1.291526 0.275178 4.6934 7.415e-06 \*\*\*  
## CSTP6 0.045381 0.142332 0.3188 0.750422   
## CSTP7 0.167588 0.263201 0.6367 0.525554   
## CSTP8 0.840202 0.167033 5.0302 1.806e-06 \*\*\*  
## CSTP9 0.831161 0.222207 3.7405 0.000287 \*\*\*  
## CSTP10 1.570143 0.254533 6.1687 1.036e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Total Sum of Squares: 431.92  
## Residual Sum of Squares: 156.37  
## R-Squared: 0.63797  
## Adj. R-Squared: 0.41326  
## F-statistic: 20.4415 on 10 and 116 DF, p-value: < 2.22e-16

#fitted(fem\_pci2) # giá trị ước lượng của biến phụ thuộc p\_PCI  
scatterplot(fitted(fem\_pci2)~PCI$CSTP1|PCI$Area,boxplots=FALSE,xlab = "CSTP1",ylab = "p\_PCI\_hat",smooth=FALSE)  
abline(lm(PCI$p\_PCI~PCI$CSTP1),lwd=3, col="red")



## 3.Uoc luong mo hinh tac dong ngau nhien - RE

rem\_pci=plm(p\_PCI~CSTP1+CSTP2+CSTP3+CSTP4+CSTP5+CSTP6+CSTP7+CSTP8+CSTP9+CSTP10+factor(time),data=PCI,index = c("City","time"),model = "random")  
summary(rem\_pci)

## Oneway (individual) effect Random Effect Model   
## (Swamy-Arora's transformation)  
##   
## Call:  
## plm(formula = p\_PCI ~ CSTP1 + CSTP2 + CSTP3 + CSTP4 + CSTP5 +   
## CSTP6 + CSTP7 + CSTP8 + CSTP9 + CSTP10 + factor(time), data = PCI,   
## model = "random", index = c("City", "time"))  
##   
## Balanced Panel: n = 63, T = 3, N = 189  
##   
## Effects:  
## var std.dev share  
## idiosyncratic 0.4261 0.6527 1  
## individual 0.0000 0.0000 0  
## theta: 0  
##   
## Residuals:  
## Min. 1st Qu. Median 3rd Qu. Max.   
## -2.225622 -0.277975 0.010008 0.344774 2.215700   
##   
## Coefficients:  
## Estimate Std. Error z-value Pr(>|z|)   
## (Intercept) 2.348817 1.178208 1.9936 0.0462011 \*   
## CSTP1 0.477706 0.087794 5.4412 5.291e-08 \*\*\*  
## CSTP2 0.651396 0.108164 6.0223 1.719e-09 \*\*\*  
## CSTP3 1.560490 0.091569 17.0418 < 2.2e-16 \*\*\*  
## CSTP4 0.235490 0.088149 2.6715 0.0075513 \*\*   
## CSTP5 1.334059 0.104326 12.7874 < 2.2e-16 \*\*\*  
## CSTP6 0.398514 0.065167 6.1152 9.642e-10 \*\*\*  
## CSTP7 0.766604 0.104857 7.3109 2.653e-13 \*\*\*  
## CSTP8 1.960413 0.072105 27.1885 < 2.2e-16 \*\*\*  
## CSTP9 1.692568 0.064341 26.3061 < 2.2e-16 \*\*\*  
## CSTP10 0.566645 0.106008 5.3453 9.027e-08 \*\*\*  
## factor(time)2021 -0.648020 0.183607 -3.5294 0.0004165 \*\*\*  
## factor(time)2022 2.290570 0.184067 12.4442 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Total Sum of Squares: 1802.7  
## Residual Sum of Squares: 66.023  
## R-Squared: 0.96338  
## Adj. R-Squared: 0.96088  
## Chisq: 4629.64 on 12 DF, p-value: < 2.22e-16

## 4.Trình bày 3 mô hình trên một bảng

stargazer(pols\_pci,fem\_pci,rem\_pci,type = "text")

##   
## ================================================================================  
## Dependent variable:   
## ---------------------------------------------------------------  
## p\_PCI   
## (1) (2) (3)   
## --------------------------------------------------------------------------------  
## CSTP1 0.478\*\*\* 0.496\*\*\* 0.478\*\*\*   
## (0.088) (0.122) (0.088)   
##   
## CSTP2 0.651\*\*\* 0.712\*\*\* 0.651\*\*\*   
## (0.108) (0.151) (0.108)   
##   
## CSTP3 1.560\*\*\* 1.599\*\*\* 1.560\*\*\*   
## (0.092) (0.151) (0.092)   
##   
## CSTP4 0.235\*\*\* 0.130 0.235\*\*\*   
## (0.088) (0.136) (0.088)   
##   
## CSTP5 1.334\*\*\* 1.409\*\*\* 1.334\*\*\*   
## (0.104) (0.159) (0.104)   
##   
## CSTP6 0.399\*\*\* 0.293\*\*\* 0.399\*\*\*   
## (0.065) (0.085) (0.065)   
##   
## CSTP7 0.767\*\*\* 0.826\*\*\* 0.767\*\*\*   
## (0.105) (0.154) (0.105)   
##   
## CSTP8 1.960\*\*\* 2.026\*\*\* 1.960\*\*\*   
## (0.072) (0.123) (0.072)   
##   
## CSTP9 1.693\*\*\* 1.503\*\*\* 1.693\*\*\*   
## (0.064) (0.152) (0.064)   
##   
## CSTP10 0.567\*\*\* 0.366\*\* 0.567\*\*\*   
## (0.106) (0.165) (0.106)   
##   
## t22 2.291\*\*\*   
## (0.184)   
##   
## t21 -0.648\*\*\*   
## (0.184)   
##   
## factor(time)2021 -0.922\*\*\* -0.648\*\*\*   
## (0.240) (0.184)   
##   
## factor(time)2022 2.109\*\*\* 2.291\*\*\*   
## (0.261) (0.184)   
##   
## Constant 2.349\*\* 2.349\*\*   
## (1.178) (1.178)   
##   
## --------------------------------------------------------------------------------  
## Observations 189 189 189   
## R2 0.963 0.888 0.963   
## Adjusted R2 0.961 0.815 0.961   
## F Statistic 385.804\*\*\* (df = 12; 176) 74.976\*\*\* (df = 12; 114) 4,629.644\*\*\*  
## ================================================================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01