Do US Multinational Firms Prefer Tax Havens?

Nate (Mingrui Xu)

May 15, 2018

Overview

News about US firms stashed money in tax havens country have caught attention of the public. Using tax havens is not new for US multinational corporations, and the strategy can be traced back to 1960s when those small island countries with limited resources wanted to attract foreign investments. But this strategy has not gotten seriously criticized until 2000 when OECD published a list of 35 tax havens and required them to accept the agreement of transparency and effective exchange of information. All of a sudden, uneasiness swept tax havens countries. Moreover, OECD made a further bold move implementing BEPS (Base Erosion and Profit Shifting) action with the intent of eliminating uncooperative tax havens. Although the attack to tax haven countries has been decades, most multinational corporations, especially US multinationals, have taken the strategy for granted. According to a report from the Institute on Taxation and Economic Policy, 366 of the country's 500 largest companies maintain at least 9,755 tax haven subsidiaries where they hold over 2.6 trillion in accumulated profits. Apple is at the very top of the offshore cash pile, booking 246 billion and avoiding 76.7 billion dollars in U.S. taxes in the process (Forbes News, 2017). Not only researchers but also US citizens should concern about this issue, because, if it is true, the US multinational firms' unpaid tax would be a burden of the public.

According to Hines (2010), there are 35 regions are classified as tax havens, of which Switzerland, Ireland, Hongkong and Singapore rank as the Big Four. Compare with 195 countries in the world, the number of tax havens is small. So, the first question–is it real that US multinationals prefer going to tax havens for the benefits in those countries? Empirically speaking, is there really a difference between investing in tax havens and non-havens for US multinationals?

Data and Design

My model is to see that whether being a tax haven would attract more actual FDI which I define as the ratio of capital investment by jobs created. 50% of the data fall into the range between 0.118 and 0.537, after checking the coverage interval.

This paper used the panel data from FDIMarket Database. After cleaning the data, the dataset contains with 22024 observations and 24 variables. The data are project-level and include all industries. The time interval is 2003-2010. The reason of using this time window is that 2003 is the earliest year in the database and I want to include 2008, a year with financial crisis to see if there is more information I would get. Moreover, my research question is about foreign direct investment, this database would provide the appropriate information. The weakness about my data, however, is the lack of firm-level controls such as firm size. The summary statistics of the dataset is given below.

% Table created by stargazer v.5.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu % Date and time: Tue, May 15, 2018 - 19:13:32

Table 1: Summary Statisitcs

Statistic	N	Mean	St. Dev.	Min	Max
Index	22,024	11,012.500	6,357.926	1	22,024
Year	22,024	2,006.724	2.237	2,003	2,010
Capital.Investment	22,024	57.289	449.199	0.000	54,000.000
Capital.Investment.1	22,024	57,289.240	449,199.000	0.000	54,000,000.000
Jobs.Created	22,024	162.824	364.024	1	8,000
Cap.job	22,024	0.546	1.221	0.000	33.333
Haven	22,024	0.098	0.297	0	1

% Table created by stargazer v.5.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu % Date and time: Tue, May 15, 2018 - 19:13:35

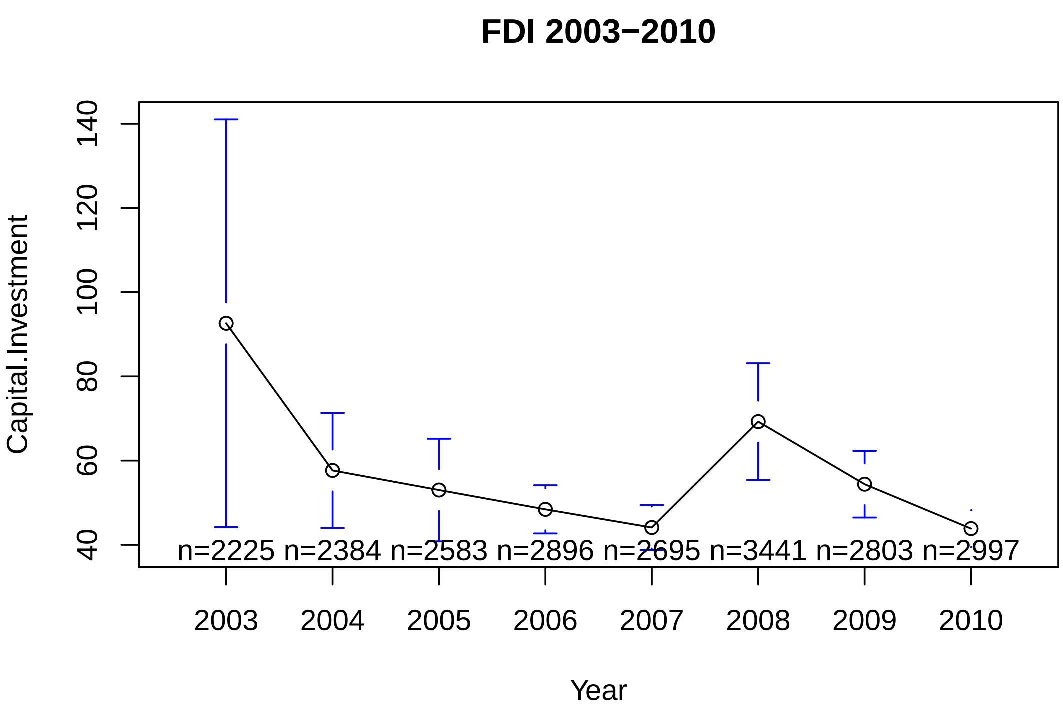
The Index is used in the data to identify each project, and so it has the range from 1 to 22024. Year is an indicator with range from 2003 to 2010. Industry.Sector includes 39 industry categories. Capital.Investment, with the unit of million dollars, is the FDI inflow made by a firm to a host country.

Table 2: Coverage Interval of Cap.job

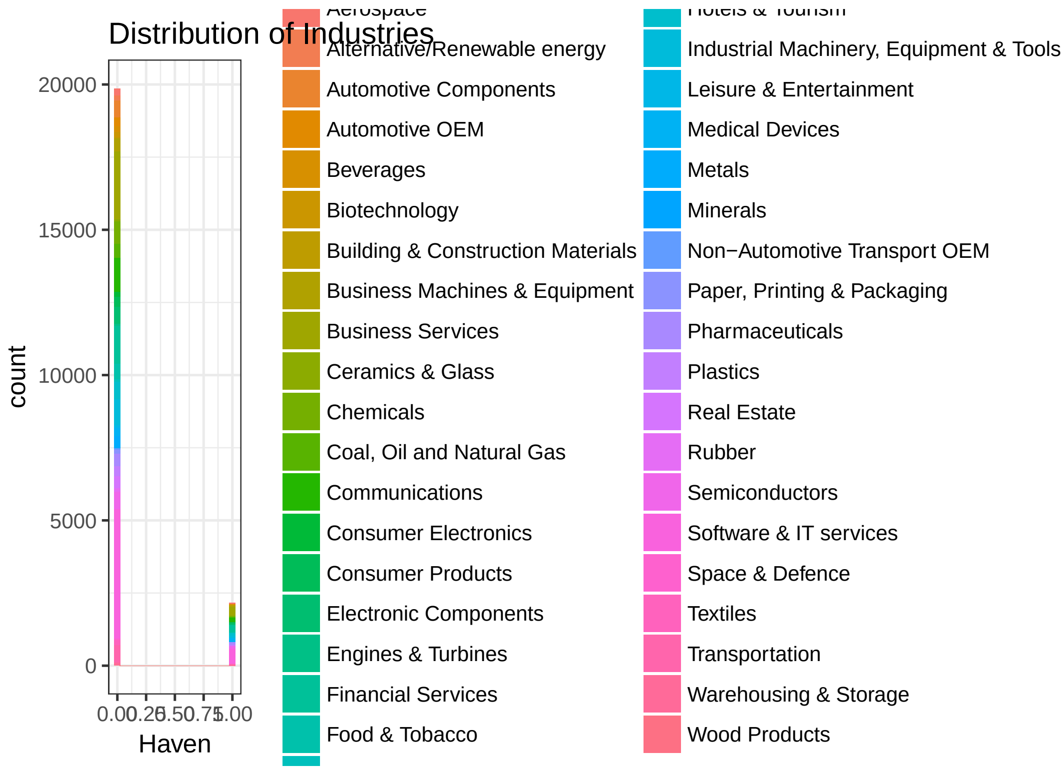
0%	25%	50%	75%	100%
0	0.118	0.253	0.537	33.333

Capital.Investment.1 is created to transfer the unit of million dollars to thousand dollars in case that some FDI data are approximated to be 0. Jobs.Created is the actual and estimated jobs created by each FDI project. Cap.job is my dependent variable, it is the result of Capital. Investment divided by Jobs.Created. To make the ratio meaningful, I eliminated 9 FDI projects with 0 job created. This ratio has the mean of 0.546, meaning the capital investments for each job created by an FDI project, on average, is 0.546 million dollars. Haven is an indicator of the destination country with 1 for a tax haven country and 0 for a non-haven country.

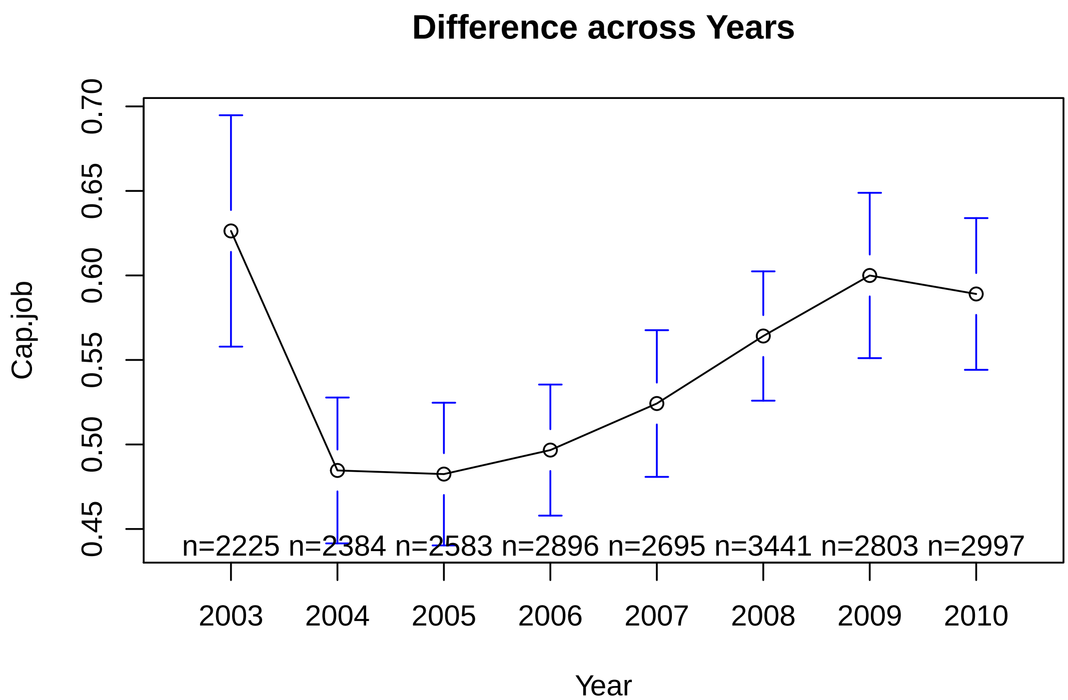
Before going deep into the research design, I would like to explore the data. The graphs below drew a 95% confidence interval around the means. The first graph shows that the general trend of FDI is decreasing from 2003 to 2007, and after a sharp increase between 2007 and 2008, the downward trend shows up again.



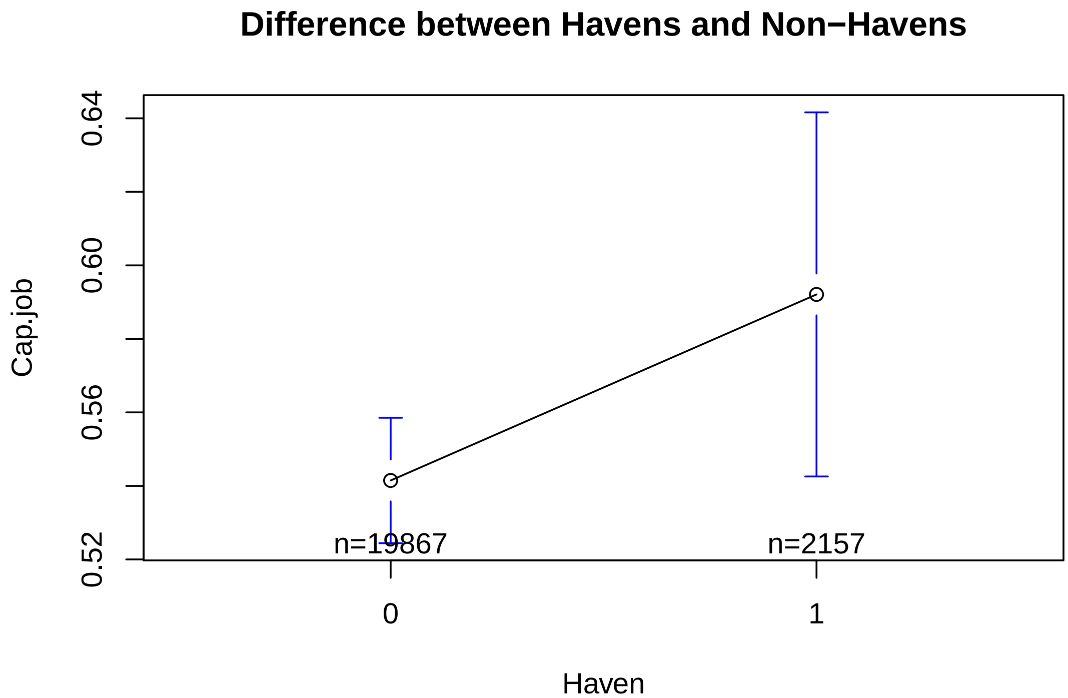
Tax havens are places include the so-called Big Four in which Hong Kong, Singapore, Ireland, and Switzerland, and other small islands such as Bermuda, Cayman Islands and Bahamas, so we may ask that which industry has a preference for tax haven? With the data including all industries, I can take an initial look. It turns out that too many industries make the graph not easy to read, but we can notice that in the tax haven group the reddish part seems bigger indicating industries such as real estate, rubber, semiconductors, etc.



But my dependent variable (labeled as Cap.job) tells another story. The capital investment per job undergoes a big drop from 2003 to 2004 and pick up from 2004. Compared with the previous graph about FDI trend, the ratio between capital investments and jobs created does not conform to a unified trend. The reason to use this ratio as my dependent variable, as mentioned at the beginning, is to capture the actual FDI invested into the tax havens countries. Instead of showing the use of tax havens to hide and transfer company profits or personal wealth, this ratio, I believe, would partly reveal the pull of making investment in tax havens. Additionally, the different trend revealed by the dependent variable leads me to thinking about the comparative change between capital investment and jobs created. Specifically, the ratio could be large due to either the increase of FDI or the decrease of jobs. This consideration could generate another empirical test to explain the reason behind the difference. This paper, however, only focuses on whether being a tax haven offers a country attraction of FDI inflows relative to a non-haven country.



Moreover, the difference in means of capital investments per job between havens and non-havens, as shown in the graph below, is not trivial. The X-axis stands for the tax haven status with 1 for tax haven countries and 0 otherwise. Approximately, the difference in means is \$0.045 million per job.



Model specification. This paper uses fixed-effects regression with industry and year fixed to estimate the effect of tax haven status on real FDI inflows. One of the reasons of using fixed effects is that it is a way to control for unobserved but fixed omitted variables (Angrist & Pischke, 2008). In other words, I believe that there are some unobserved factors in different industries and years affecting the precision of my model.

The regression model is as below

$$Cap.job = Haven + fe(Ind) + fe(Year)$$

The dependent variable is a ratio of capital investments by jobs created, labeled Cap.job in this empirical setting. The explanatory variable is an indicator, tax haven status, with 1 as a tax haven FDI and 0 otherwise.

Furthermore, this paper also tries to explore another approach to do the fixed-effects regression. The traditional two-ways fixed-effects, is to transform data twice, i.e., subtracting the industry-transformed data from the independent variable first and the Year-transformed data later. This paper tested, besides the traditional way, the fixed-effects on the interaction between Industry-transformed data and Year-transformed data. This method has its mathematical support (e.g. Imai & Kim, 2016) and offers an easier interpretation compared with the traditional way.

As a consequence, the other fixed-effects model could be specified as below

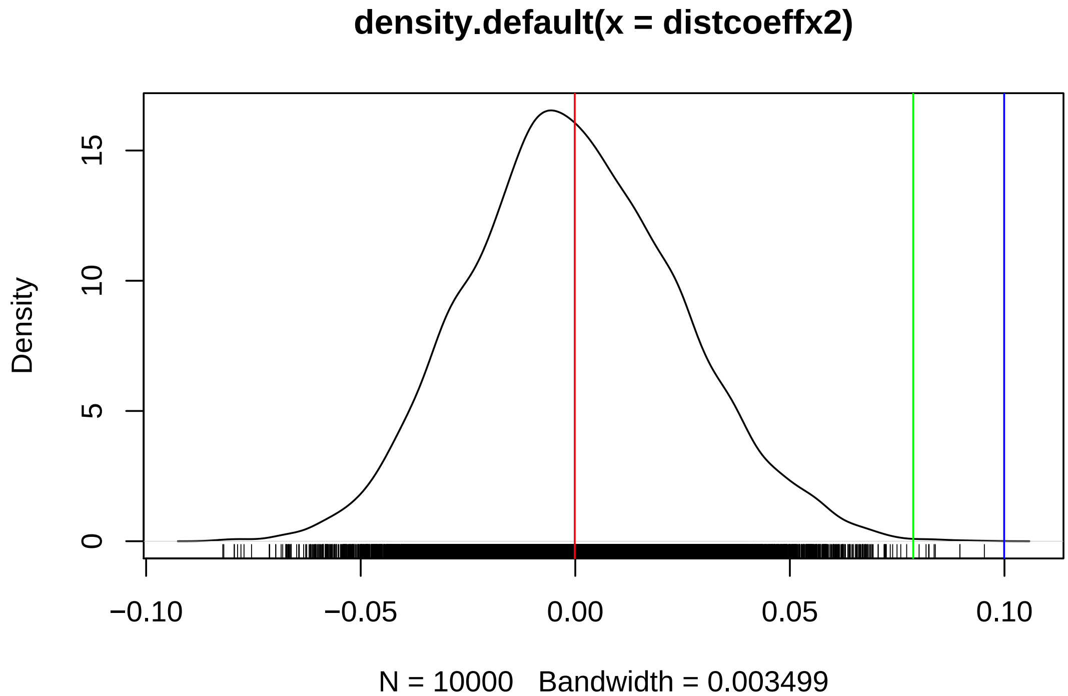
$$Cap.job = Haven + fe(Ind) * fe(Year)$$

Comparisons are necessary in my case when alternative models are available. A potential advantage to compare these two model lies in the generalization to other similar empirical settings. Therefore, I calculate the false positive rate by permutation test. It turns out that the model with fixed-effects on the interaction term performed better, generating smaller false positive rate — 0.0506 for the traditional method, and 0.0495 for the fixed-effects model based on the interaction term. The results of these false positive rates could be interpreted as that according to the fixed-effects model I chose, I would reject the null hypothesis which is actually true 506 (or 495) out of 10000 times.

Statistical Inference

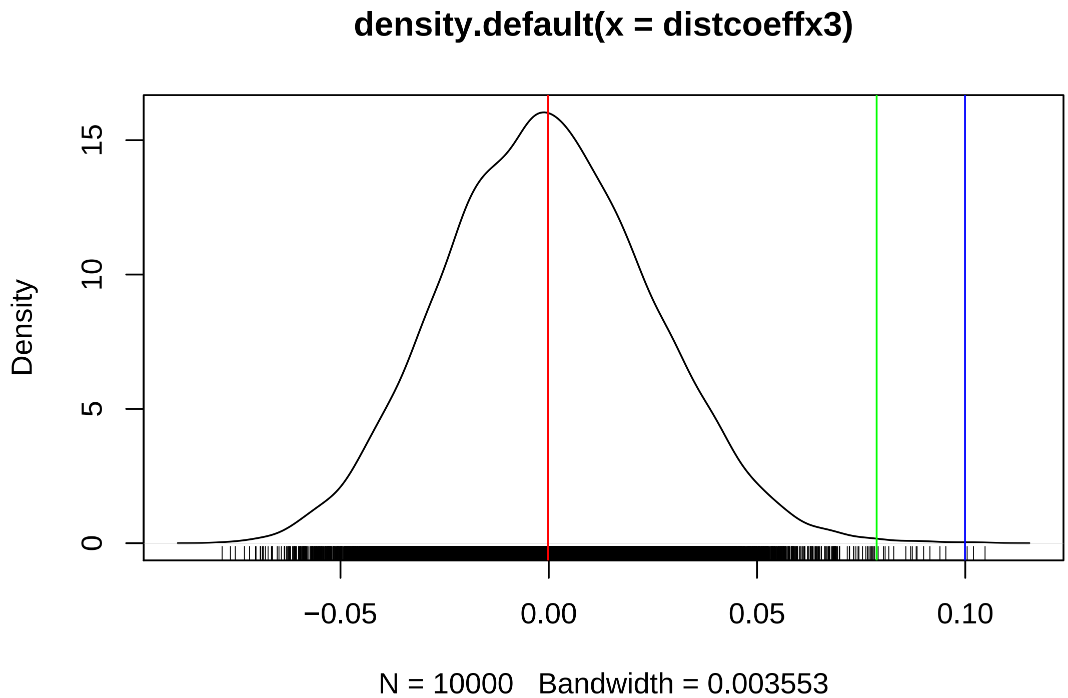
This empirical paper does not rely on the canned OLS model to generate conclusions about statistical inference including p-values. Instead, this paper alternates the question into an experiment design where I use the permutation test to study the relationship between tax haven status and the attraction of actual FDI inflows. The p-values would be calculated after the permutation test where both one-sided and Rosenbaum two-sided p-values are provided. The reasons of doing the as-if experiment method can be summarized in three aspects. First, my research question is about the effect of tax haven status on the attraction of FDI, and thus breaking the relationship by permutation is an appropriate way to answer my question. The p-value, acquired after permutation, serves as a measure of the information against the null hypothesis. Second, the sample size is large enough to take advantage of the merits of design-based method instead of normal OLS regression, and the permutation test would provide me evidence that based on my data whether OLS could offer me a precise result which is as good as using the permutation test. Third, this research design responds to the call of experimental method from JIBS (Zellmer-Bruhn et. al 2016).

Two density plots of the treatment effect after permutations are provided below. The first plot is the fixed effects on Industry first and on Year second. The red line is the mean of the treatment effect acquired from running the permutation 10000 times. The blue line is the coefficient calculated from OLS. Visually, the blue line is far away from 0.



Note: this graph shows the permutation result under the traditional two-ways fixed effects in which the independent variable is transformed by subtracting the industry-transformed and year-transformed data. The X-axis contains the coefficients extracted from the permutation test after 10000 times

The other density plot below is the permutation under the interaction-based fixed effects model. The red line is the mean of the treatment effect extracted from 10000 times permutation. The green line is the coefficient calculated from the fixed effects model based on the interaction between Industry and Year. Visually, the green line is far from 0 but not as far as the blue line.

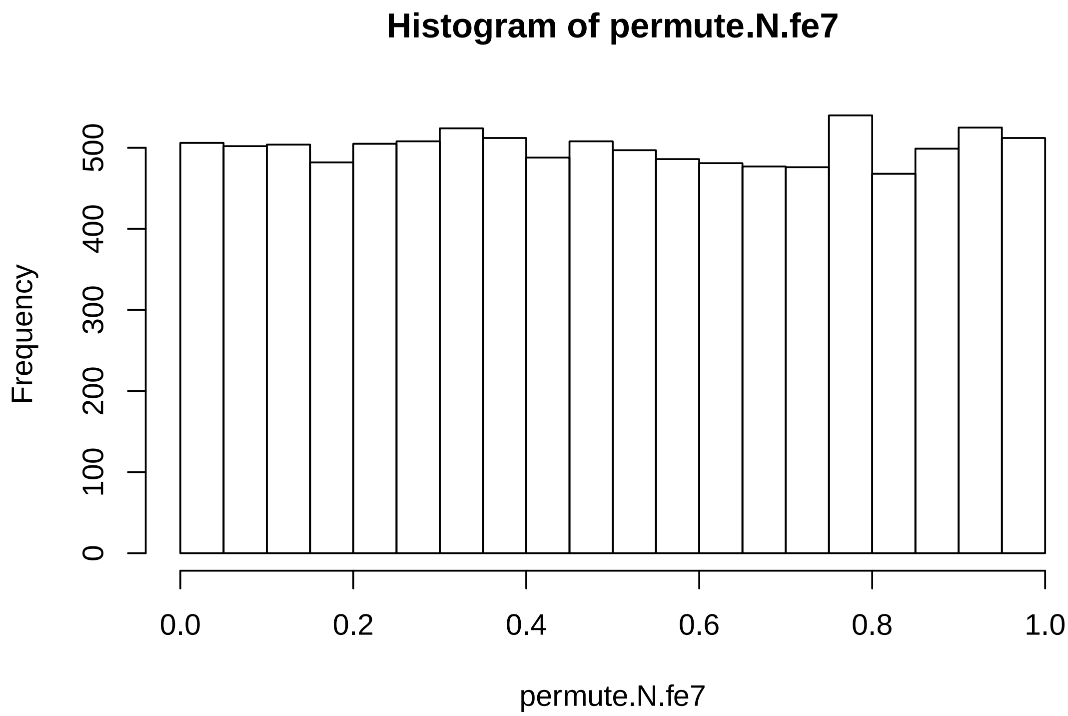


Note: this graph shows the permutation result under the two-ways fixed effects in which the independent variable is transformed by subtracting the interaction between industry-transformed and year-transformed data

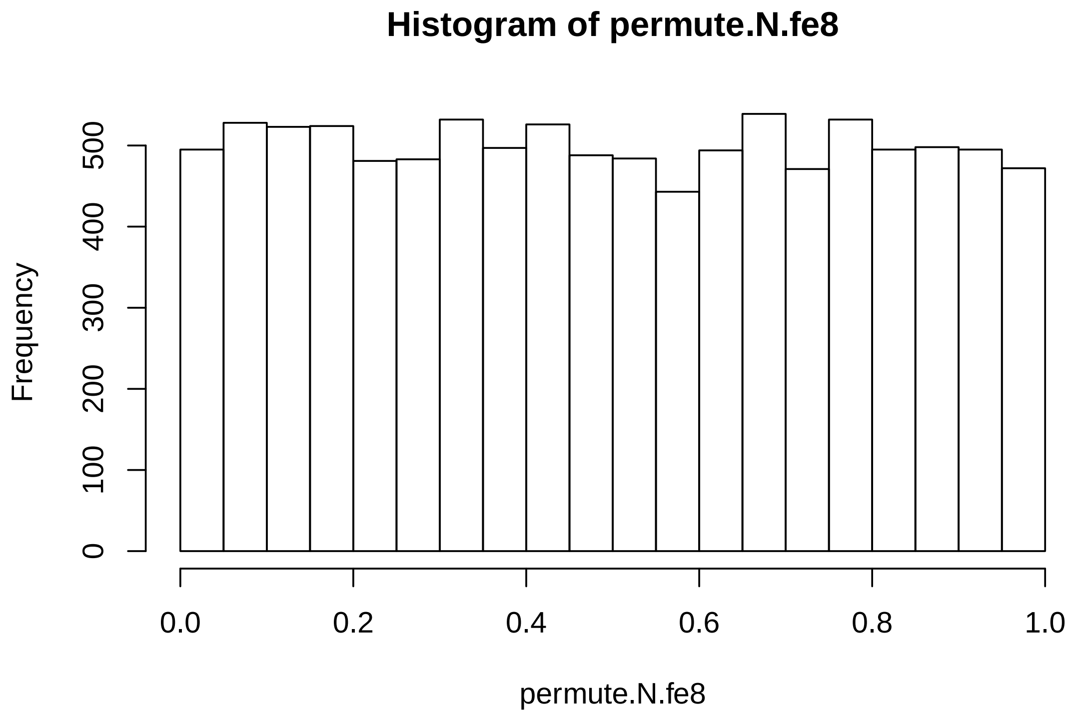
Taking the two plots together, we can see that the fixed-effects model based on the interaction is more conservative in terms of rejecting the null hypothesis. The shape of the first graph is not exactly the same as the second one, but this is not something problematic because my procedure guarantees my results being robust with any shape. After the first permutation, the two-sided p-value with the 5% threshold is 0 under the traditional fixed effects model, meaning than I can reject any null hypothesis and thus there is significant difference in US firms making FDI into tax havens relative to non-havens. The second permutation test generated the p-value equal to 0.0034 under the fixed-effects on the interaction term.

A following question might be “which model performed better?” This is an empirical question instead of a theoretical one and thus false positive rates would be a proper criterion. This question requires another permutation test with the intent to extract p values. Also, a distribution of p-values could offer me information about my null hypothesis because the distribution of p-values should be equally distributed under null hypothesis. As a result, I ran another

10000 times permutation, and I got the two distributions for two models as below.



Note: this distribution of p values is from permutation under the traditional two-ways fixed effects



Note: this distribution of p values is from permutation under the fixed effects based on the interaction between industry-transformed data and year-transformed data

In line with my density plots of treatment effect which are close to normal distributions, the two histograms indicate that the distributions of p-values are nearly equally distributed under the null hypothesis that there is no relationship between tax haven status and the attraction of real FDI. Moreover, this paper uses a large sample size with 22024 observations. As a result, using CLT instead of permutation test can also generate statistical inference as good as doing permutation.

The distribution of p-values under the traditional fixed-effects indicates a higher frequency around 0, compared with the fixed-effects model on the interaction term. This difference suggests that the traditional two-ways fixed-effects method is more likely to reject the null hypothesis.

With the p-values extracted from the two permutations, I calculated the false positive rates under the 0.05 threshold I picked. It turns out that I will reject my null hypothesis which is actually true 506 out of 10000 times under the traditional method, but 495 out of 10000 times under the method based on the interaction term. Predicated on the criterion of false positive rates, the later model outperformed the traditional one.

The results of two fixed-effects models are reported below. In accordance with the results from permutation tests. The canned OLS provided similar results for statistical inference. The two coefficients of Haven with p-value less than 0.05 indicate the significant difference in terms of US firms making FDI between tax havens and non-haven countries. What needs to be concerned is the constants in different fixed-effects models – negative in the traditional two-ways fixed-effects model, 0.24 in the fixed-effects model based on the interaction between Industry and Year. The meaning of the constant is the average capital investments per job made by US firms in the non-haven countries. Therefore, the negative constant although near to 0 means that 1) US firms have not made any further FDI expansion into non-haven countries, holding constant the jobs created by the previous FDIs, and 2) US firms have hired more employees in the non-haven countries, holding constant the FDI in those host countries. Henceforth, I believe that the relative change of jobs and FDIs needs to be further studied to reveal the actual effect of tax havens on other economies.

```
##
## % Table created by stargazer v.5.2 by Marek Hlavac, Harvard University. E-mail:
## % Date and time: Tue, May 15, 2018 - 19:22:33
## % Requires LaTeX packages: dcolumn
## \begin{table}[!htbp] \centering
##   \caption{Two-way Fixed-Effects Regression Results}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lD{.}{.}{-3} D{.}{.}{-3} }
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{2}{c}{\textit{Dependent variable:}} \\
## \cline{2-3}
## \\[-1.8ex] & \multicolumn{1}{c}{FDI/Job_T} & \multicolumn{1}{c}{FDI/Job_I} \\
## \\[-1.8ex] & \multicolumn{1}{c}{(1)} & \multicolumn{1}{c}{(2)}\\
## \hline \\[-1.8ex]
##  Haven(=1) & 0.100^{***} &  \\
##   & (0.024) &  \\
##   & & \\
##  Haven(=1) &  & 0.079^{***} \\
##   &  & (0.025) \\
##   & & \\
##  Constant & -0.000 & 0.240^{***} \\
##   & (0.007) & (0.008) \\
```

```
##   & & \\
## \hline \\[-1.8ex]
## Fixed effects & Yes & Yes \\
## Observations & \multicolumn{1}{c}{22,024} & \multicolumn{1}{c}{22,024} \\
## R$^{2}$ & \multicolumn{1}{c}{0.001} & \multicolumn{1}{c}{0.0005} \\
## Adjusted R$^{2}$ & \multicolumn{1}{c}{0.001} & \multicolumn{1}{c}{0.0004} \\
## Residual Std. Error (df = 22022) & \multicolumn{1}{c}{1.065} & \multicolumn{1}{
## F Statistic (df = 1; 22022) & \multicolumn{1}{c}{16.728$^{***}$} & \multicolumn
## \hline
## \hline \\[-1.8ex]
## \textit{Note:}  & \multicolumn{2}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$
## \end{tabular}
## \end{table}
```

Model adjustment

An important issue of doing fixed-effects lies in the loss of degree of freedom, leading to imprecise standard errors. The standard errors from the canned regression are not the right ones, because the OLS assumes that each observation is independent from others. This assumption does not apply to my dataset. In other words, I cannot assume that an FDI project made by a US firm in Software & IT industry in 2003 is independent from the one in 2004. After calculating HC0-HC4 for each fixed-effects model, I picked HC0 (SE(constant) = 0.007179725, SE(Haven) = 0.02431190) for the traditional model, and HC0 (SE(constant) = 0.007695094, SE(Haven) = 0.02485703). The reason why I chose HC0 over others is that HC0 gave me the smallest standard errors for the independent variable, Haven, although the difference among all the five kinds of standard errors is trivial.

Another adjustment should be made is the permutation test. I did the unrestricted permutation instead of the restricted one. Unrestricted permutation may cause problems of justifying the result of permutation. That being said, I shuffled my data among different industries and years. Although I subtracted from the independent variable the mean of the industry and the mean of year, doing permutation without blocks may still cause the imprecise treatment effects. Therefore, I need to create a block using the interaction between industry and year, and then do the permutation within the block. Unfortunately, I could not make the R code work within the time limit, although I understand the mechanism behind this approach.

Conclusion

Based on the fixed-effects model, permutation test and robust standard error, this paper provided some empirical support that being a tax haven has a significant advantage in terms of attracting US firms' FDI compared with non-haven countries. Two fixed effects are presented and compared by false positive rate. The fixed-effects model based on the interaction between industry and year outperformed the traditional two-ways fixed effects. And this finding may serve as an alternative approach when considering two-ways fixed-effects.

This paper is not without limitations. First, the model needs firm level controls, because I actually assumed that all the US firms making FDI in other countries are comparable. This assumption may not hold in reality. Second, as mentioned in the model adjustment part, an alternative research design should be provided using within-block permutation, because this design would be more convincing by controlling the permutation process.

I think the most important lesson I learned is not taking OLS results for granted. Under some situations, OLS results are misleading for one's research question. And it is not the case that an easy model would generate an easy answer. In this paper, the independent variable is just an indicator, but there is still a lot to discuss about how to justify this indicator does affect the interested independent variable. To illustrate this point, the baseline regression results without any fixed-effects is provided as a comparison. As is shown from the table below, tax haven status is not significant without fixed effects, indicating the effects of unobservable factors which, in my theory, exist in industry and time differences. Furthermore, searching a significant result is never an ultimate goal. Instead, the theory and the proper research design make a research convincing.

References

```
# put all the package I am using here
library(ggplot2)
library(gplots)
library(foreign)
library(plyr)
library(stargazer)
library(lmtest)
```

```
library(hcci)
library(plm)
library(dplyr)
library(permute)
library(perm)
library(vegan)
require(knitr)
library(tidyverse)
library(RefManageR)
library(citr)
download.file(url = "https://raw.githubusercontent.com/NateXu/TaxHaven/master/USta
fdi.df2=read.csv("taxhaven.data.csv")

summary(fdi.df2)
str(fdi.df2)
quantile(fdi.df2$Cap.job)
# summary statistics and coverage rate of the DV in my data

stargazer(fdi.df2,title = "Summary Statisitcs", summary = TRUE )

stargazer(quantile(fdi.df2$Cap.job),title = "Coverage Interval of Cap.job ")

plotmeans(Capital.Investment ~ Year, main="FDI 2003-2010", data =fdi.df2 )

# Look at to which industry foreign FDIs are invested
ggplot(fdi.df2, aes(x = Haven, fill=Industry.Sector))+
    theme_bw() +
    geom_histogram()+
    stat_bin(bins = 20)+
    ggtitle("Distribution of Industries")

plotmeans(Cap.job ~ Year, main="Difference across Years", data =fdi.df2 )


plotmeans(Cap.job ~ Haven, main="Difference between Havens and Non-Havens", data =

#fixed effect
attach(fdi.df2)
fdi.df2$Year<-as.numeric(Year)
```

```
fdi.df2$Industry.Sector<-as.numeric(Industry.Sector)

# I will do two steps to transform my data to do the fixed effects

# First, Industry Transformation.
# Step 1. Split data (Using "ddply") by industry.
# Step 2. Calculating de-mean of X by subtracting from X(Haven) the average of X
# Step 3. Subtracting from Y(Cap.job) the average of Y calculated from the corre

fdi.yeardata4<-ddply(fdi.df2,.(Industry.Sector),
                     transform,
                     gmeanHaven= mean(Haven),
                     gmeanCapjob= mean(Cap.job),
                     dmean_x1= Haven - mean(Haven),
                     dmean_y1= Cap.job-mean(Cap.job))


# Second, Year Transformation. I take the industry-transformed data to do the ye
# Step 1. Split data (Using "ddply") by Year.
# Step 2. Calculating a new de-mean of X by subtracting from dmean_x1 (industry-
# Setp 3. Subtracting from deman_y1 (industry-transformed Y) the average of dmea
fdi.yeardata5<-ddply(fdi.yeardata4,.(Year),
                     transform, y=Cap.job,
                     dmean_x2= dmean_x1 - mean(dmean_x1),
                     dmean_y2= dmean_y1 - mean(dmean_y1))


# fixed effect model
# A problem is that when I transform my data twice, I eat up many degree of free
fe7<-lm(dmean_y2 ~ dmean_x2, data = fdi.yeardata5)
summary(fe7)
#fixed effect on the interaction of IND and Year
attach(fdi.df2)
fdi.df2$Year<-as.numeric(Year)
fdi.df2$Industry.Sector<-as.numeric(Industry.Sector)

# Just as I did in the last chunk using the traditional fixed effects,
# I will take two steps to generate the interaction between the group mean of x
```

```r
# First, using ddply split the data by Industry and generate the group mean of Ha
fdi.yeardata6<-ddply(fdi.df2,.(Industry.Sector),
                     transform,
                     gmeanHaven2= mean(Haven),
                     gmeanCapjob2= mean(Cap.job))

# Second, I take the dataset which I just added two columns to the orginial data
# Using ddply, I generate the group mean of Cap.job and the group mean of Haven
fdi.yeardata7<-ddply(fdi.yeardata6,.(Year),
                     transform,
                     gmeanCapjob3= mean(Cap.job),
                     gmeanHaven3= mean(Haven))

# Last, I run the two-way fixed-effects regression with the interaction between
attach(fdi.yeardata7)
dmean_x4 <- Haven - gmeanHaven2*gmeanHaven3
dmean_y4 <- Cap.job - gmeanCapjob2*gmeanCapjob3


# fixed effect model
# A problem is that when I transform my data twice, I eat up many degree of free
fe8<-lm(dmean_y4 ~ dmean_x4, data = fdi.yeardata7)
summary(fe8) # compared with the traditional fixed effects, using the interactio
summary(fe7) # the traditional fixed effects.

# permutation test under the traditional two-ways fixed effects

myexp2 <- function(y){
  shufx <- sample(fdi.yeardata5$dmean_x2)
  coeffx <- coef(lm(y~shufx))[["shufx"]]
  return(coeffx)
}

set.seed(3333)
distcoeffx2 <- replicate(10000, myexp2(y=fdi.yeardata5$dmean_y2))

plot(density(distcoeffx2))
rug(distcoeffx2,col="black",line=0)
abline(v=mean(distcoeffx2), col="red")
```

```r
abline(v=coef(summary(fe7))[2,1], col="blue")
abline(v=coef(summary(fe8))[2,1], col="green")
# permutation test under the two-ways fixed effects on the interaction

myexp3 <- function(y){
  shufx <- sample(dmean_x4)
  coeffx <- coef(lm(y~shufx))[["shufx"]]
  return(coeffx)
}

set.seed(33334)
distcoeffx3 <- replicate(10000, myexp3(y=dmean_y4))

plot(density(distcoeffx3))
rug(distcoeffx3,col="black",line=0)
abline(v=mean(distcoeffx3), col="red")
abline(v=coef(summary(fe7))[2,1], col="blue")
abline(v=coef(summary(fe8))[2,1], col="green")

# Calculating false positive for fe7, the traditional way of fixed effects
# I use summary(lm()) to acquire p values
attach(fdi.yeardata5)
set.seed(004)
N = 10000
permutefun_fe7<-function(y=dmean_y2, x=dmean_x2){
  model.resample_fe7 = lm(sample(y, replace = F) ~ x)
  summary(model.resample_fe7)$coefficients[2,4] # some online tutorial suggests t
  }
permute.N.fe7<-replicate(N,permutefun_fe7())
hist(permute.N.fe7)

# The false positive rate calculation for fe8(my fixed-effect model fixing on th
# The method is the same as the above chunk
attach(fdi.yeardata7)
set.seed(005)
N = 10000
permutefun_fe8<-function(y=dmean_y4, x=dmean_x4){
  model.resample_fe8 = lm(sample(y, replace = F) ~ x)
  summary(model.resample_fe8)$coefficients[2,4]
```

```
  }
permute.N.fe8<-replicate(N,permutefun_fe8())
hist(permute.N.fe8)

stargazer(fe7, fe8,
          dep.var.labels = c("FDI/Job_T","FDI/Job_I"),
          covariate.labels = c("Haven(=1)","Haven(=1)"),
          title="Two-way Fixed-Effects Regression Results",
          add.lines = list(c("Fixed effects", "Yes", "Yes")),
          align = T )
```