Fundamental Data

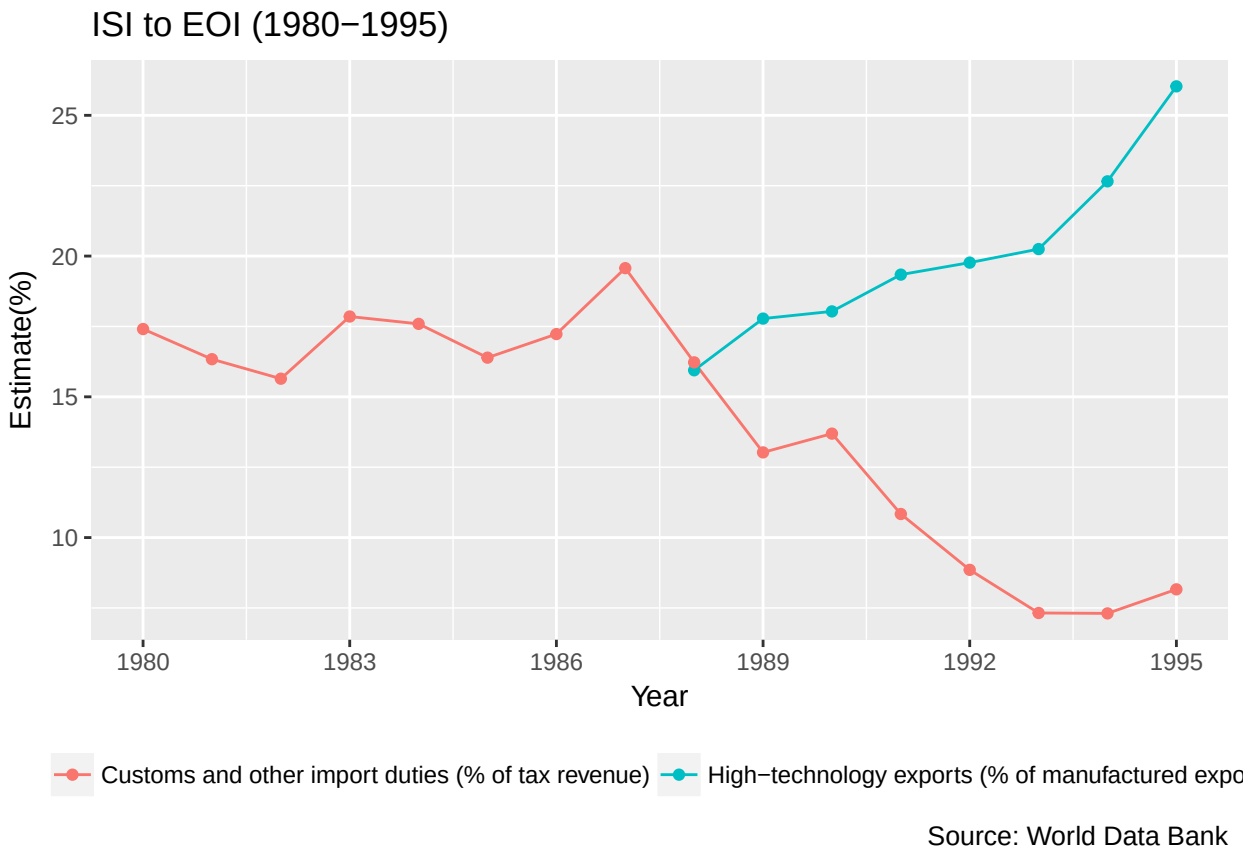
Colin Pi

2018 5 18

Data wrangling and visualization of the World Bank Dataset (which I used for addressing fundamental cause). The World Bank data is a time series data of Manufacturing (% of GDP), Proportion of Textile and Clothing in Manufacturing, Expenditure on Secondary Education (% of government expenditures on education), High-technology export(% of manufactured export), and Customs and other import Duties (% of tax revenue) from 1960 to 2017.

```
library(dplyr)
library(tidyverse)

## Read the World Bank Dataset

final_data <-
  read.csv("~/Documents/2017-18/ECON241/Final_Project/Fundamental/Fundamental.csv")[-c(6:10),]

names(final_data)[5:62] = seq(1960,2017)

## gather the dataset into narrower form

narrow_final_data_func <- function(i){
  outcome <- gather(final_data %>% filter(Series.Name == final_data$Series.Name[i]),
                    key = Year, value = UQ(final_data$Series.Name[i]), 5:62) %>% select(6)
  return(outcome)
}

narrow_final_data <- gather(final_data %>% filter(Series.Name == final_data$Series.Name[1]),
                            key = Year, value = UQ(final_data$Series.Name[1]), 5:62) %>%
  select(-Series.Name,-Series.Code) %>% bind_cols(lapply(2:5, narrow_final_data_func) %>% bind_cols())

narrow_final_data[,3:8] <- sapply(narrow_final_data[,3:8], as.numeric)

## dataset for Figure 1

figure1 <- narrow_final_data %>% filter(Year >= 1980 & Year <= 1995) %>% select(-c(4:6)) %>%
  gather(key = Variables, value = Estimates, 4:5)

## dataset for Figure 2

figure2 <- narrow_final_data %>% filter(Year >= 1970 & Year <= 1990) %>% select(-c(7:8)) %>%
  gather(key = Variables, value = Estimates, 4:6)
```

```
library(ggplot2)
figure1 %>% ggplot(aes(x = Year, y = Estimates, color = Variables)) +
  geom_line() +
  geom_point()+
  scale_x_continuous(breaks = seq(1980, 1995, by=3)) +
  theme(legend.position="bottom", legend.title = element_blank()) +
  labs(title = "ISI to EOI (1980-1995)",y="Estimate(%)",
       caption = "Source: World Data Bank")
```



```r
ggplot(data=figure2[!is.na(figure2$Estimates),],
       aes(x = Year, y = Estimates, color = Variables)) +
  geom_line() +
  geom_point()+
  scale_x_continuous(breaks = seq(1970, 1990, by=2)) +
  theme(legend.position="bottom", legend.title = element_blank()) +
  labs(title = "Manufacturing and Secondary Education (1970-1990)",y="Estimate(%)", caption = "Source: W
```