HW1 - Structural Transformation

Nikhilla B, Ezeriki Emetonjor, Alistair Pyle,

2024-02-16

Question 1

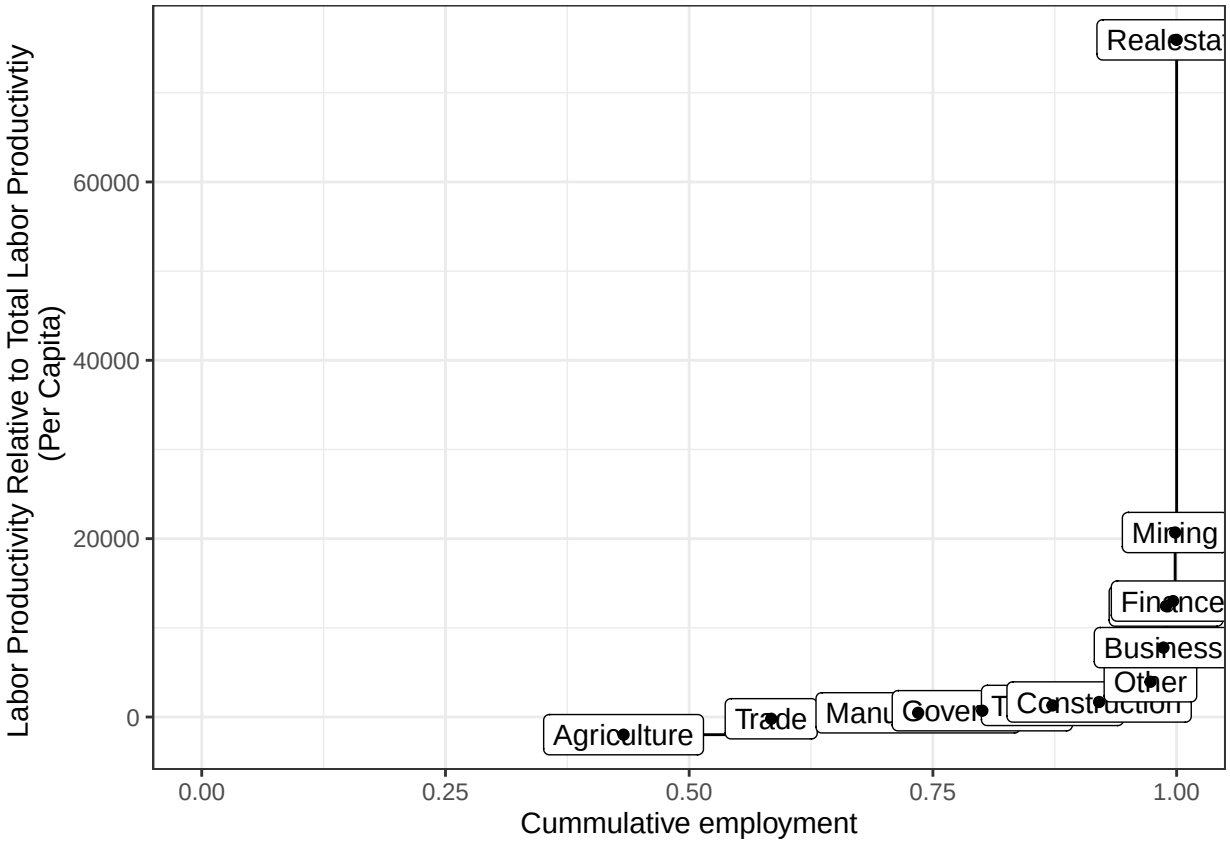
One potential explanation for this difference could stem from the World Bank's focus solely on trade-able goods within markets. As a result, numerous agricultural, locally-produced, or community-owned items may not be included in their calculations, potentially leading to an underestimation of the overall value.

```
## # A tibble: 1 x 3
##   Total_VA_Q15 Total_EMP total_per_worker
##          <dbl>     <dbl>            <dbl>
## 1 183156737500  59696514            3068.
```

Question 2

The Normalized productivity of agriculture is -1983.73, while the employment share of agriculture is 43.2% (as of 2015)

Question 3

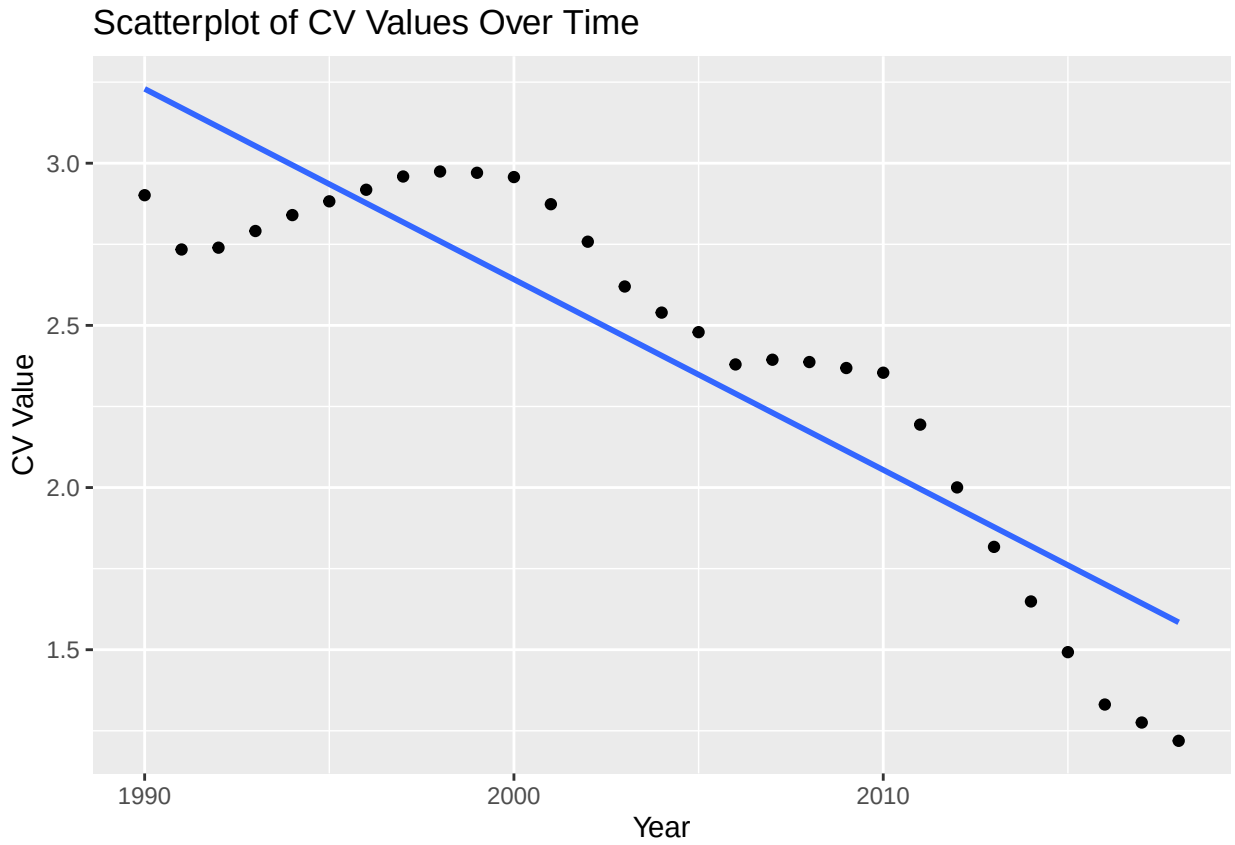


This data illustrates an intriguing trend: despite being the least productive sector, Agriculture boasted the largest share of the labor force in Bangladesh in 2015. Conversely, highly productive sectors like Real Estate, Mining, Finance, and Business commanded only a miniscule fraction of the labor force. This disparity among sectors has the potential to hinder overall labor productivity.

Question 4

The percentage increase after flipping productivity of agriculture and manufacturing is 22.5% (from 3068.131 to 3759.322)

Question 5



As labor productivity across sectors converges towards the mean over time, it suggests that less productive sectors are narrowing the gap with others. This trend results from reduced variation in productivity within each sector, leading to more consistent output. Such stability fosters an environment conducive to progress. Consequently, Bangladesh may resemble advanced countries discussed in previous lectures more closely than other low-income countries, owing to this decreased variation in sector productivity.

Question 6

```
## # A tibble: 3 x 7
##    year Agriculture_VA_Q15 Agriculture_EMP       mfg_va mfg_emp   services_va
##   <dbl>              <dbl>           <dbl>        <dbl>   <dbl>         <dbl>
## 1  1990       12232852827.        26836213  5682913328. 5289198  31048064719.
## 2  2005       18822526115.        22984794 16003341476. 5051255  65661485604.
## 3  2018       30887840910.        25808861 51965884144. 9735080 142898353083.
## # i 1 more variable: services_emp <dbl>
```

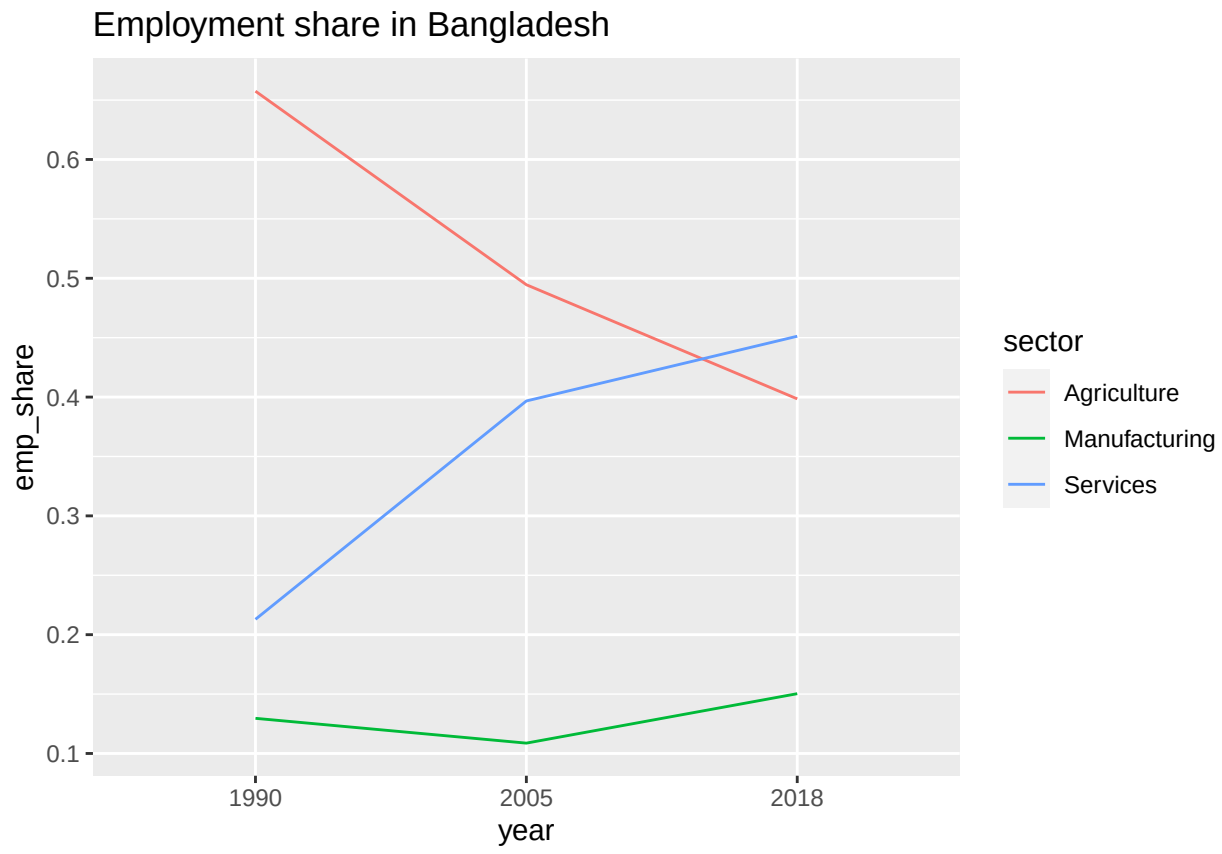
Question 7

Between 1990 and 2005, approximately 47% of productivity decomposition stemmed from inter-sector labor reallocation, illustrating Bangladesh's shift away from agriculture dependency towards manufacturing. Conversely, 52% of the productivity decomposition during this period originated from intra-sector reallocation, likely influenced by advancements in the ready-made garment industry.

In the timeframe from 2005 to 2018, 71% of the productivity decomposition was attributed to within-sector productivity growth, possibly reflecting enhancements in the services sector. This shift, accompanied by a structural change towards the service sector (28%), may encapsulate Bangladesh's narrative of structural transformation.

```
## # A tibble: 6 x 15
##   sector year     emp      va    lp      L emp_share      VA    LP del_lp del_LP
##   <chr>  <chr>  <dbl>   <dbl> <dbl>  <dbl>     <dbl>   <dbl> <dbl>  <dbl>  <dbl>
## 1 Agric~ 2005  2.30e7 1.88e10  819. 4.65e7     0.495 1.00e11 2162.     0      0
## 2 Agric~ 2018  2.58e7 3.09e10 1197. 6.48e7     0.398 2.26e11 3486.   378.  1323.
## 3 Manuf~ 2005  5.05e6 1.60e10 3168. 4.65e7     0.109 1.00e11 2162.     0      0
## 4 Manuf~ 2018  9.74e6 5.20e10 5338. 6.48e7     0.150 2.26e11 3486.  2170.  1323.
## 5 Servi~ 2005  1.84e7 6.57e10 3561. 4.65e7     0.397 1.00e11 2162.     0      0
## 6 Servi~ 2018  2.92e7 1.43e11 4890. 6.48e7     0.451 2.26e11 3486.  1329.  1323.
## # i 4 more variables: del_emp_share <dbl>, product <dbl>, numerator <dbl>,
## #   within <dbl>
```

Question 8



The decomposition within sectors further confirms that the bulk of productivity growth from the 1990s to 2005 occurred within the manufacturing sector, with the service sector contributing the least. In contrast, between 2005 and 2018, the service sector experienced the most significant productivity growth, with agriculture contributing only 14%. Additionally, the graph reveals a sharp decline in the employment share within the agriculture sector from 2005 onwards, accompanied by a proportional increase in the service sector. Notably, while manufacturing does not dominate employment share, it still experiences growth between 2005 and 2018.

Question 9

We withhold final judgement on Bangladesh's developmental prospects based on these findings regarding their structural trajectory. Similar to the approach outlined in Fan et al., our optimism for Bangladesh would be warranted if the expansion of the services sector is driven by productivity growth rather than mere demand. This scenario would suggest the potential for sustained growth. While we lack this specific model applied to Bangladesh, we maintain cautious optimism regarding their structural transition from agriculture to services.

Code

```
# Loading libraries ---------------------------------
library(tidyverse)
library(gt)
library(haven)
library(sf)
library(glue)
library(modelsummary)
library(fixest)
library(fastDummies)
library(ggplot2)



# Loading data --------------------------
structural_change <- readRDS("ETD_230918.RDS")

#structural_change |> colnames()

# functions ----------------------------
perform_division <- function(prefix, your_data) {
  # Column names with "_VA_Q15" suffix
  col_va <- paste0(prefix, "_VA_Q15")
  # Column names with "_EMP" suffix
  col_emp <- paste0(prefix, "_EMP")
  # Perform division
  if (col_va %in% colnames(your_data) & col_emp %in% colnames(your_data)) {
    labor_product <- your_data[[col_va]] / your_data[[col_emp]]
    return(labor_product)
  }
}


# converting variables ----------------------
structural_change_transform <- structural_change |>
  filter((var == "VA_Q15" | var == "EMP") & country == "Bangladesh") |>
  pivot_wider(names_from = var, values_from = c(Agriculture, Mining,
                                                Manufacturing, Utilities, Construction,
                                                Trade, Transport, Business, Finance,
                                                Realestate, Government, Other, Total,
                                                Warflag)) |>
  mutate(across(ends_with("VA_Q15"), ~ (.x * 1000000 /80))) |>
```

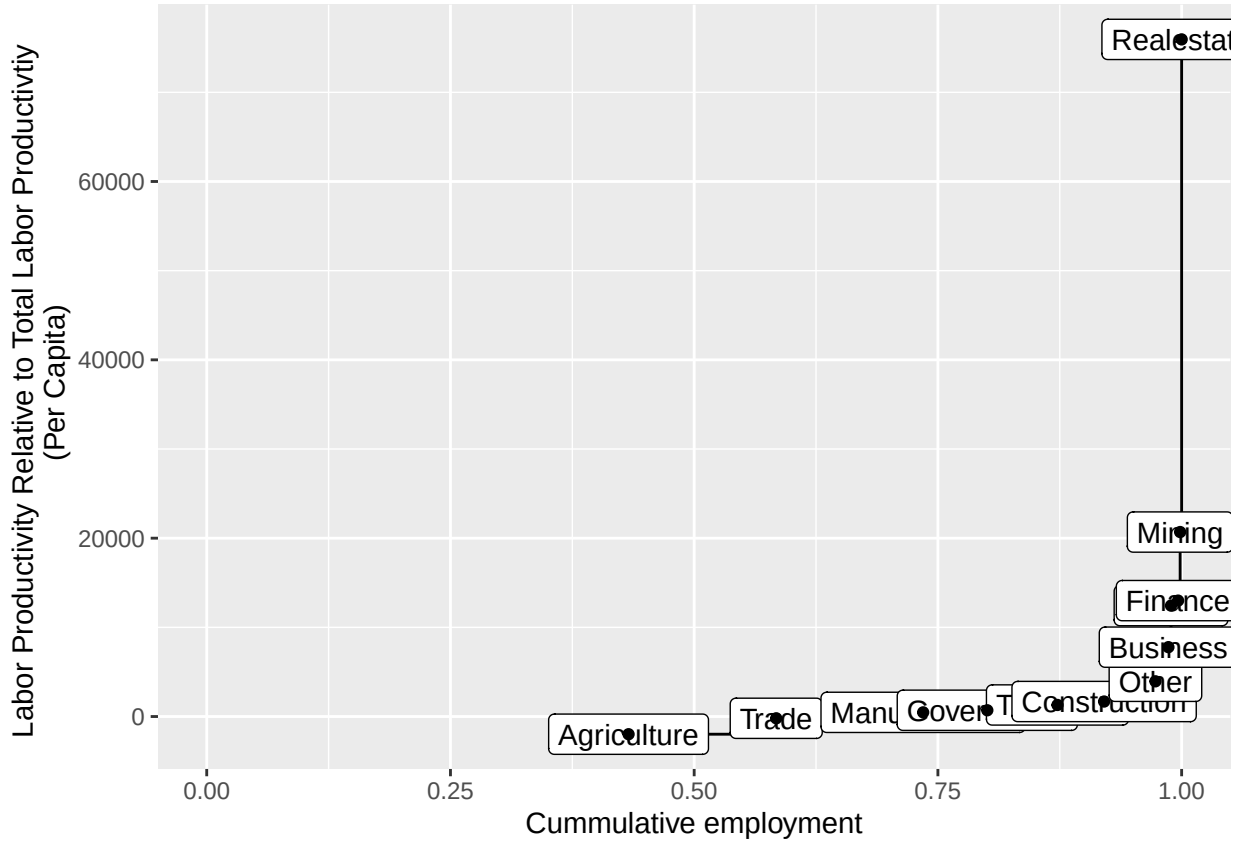
```
  mutate(across(ends_with("EMP"), ~ .x * 1000))


# Question 1 ---------------------------
# total value added per worker in 2015
structural_change_transform |>
  filter(year == 2015) |>
  select(c(Total_VA_Q15, Total_EMP)) |>
  mutate(total_per_worker = Total_VA_Q15/Total_EMP)
```

```
## # A tibble: 1 x 3
##   Total_VA_Q15 Total_EMP total_per_worker
##          <dbl>     <dbl>            <dbl>
## 1 183156737500  59696514            3068.
```

```
# the reason for this discrepancy might be because the World Bank only takes into account
# trade-able goods in markets
# However, a lot of agricultural, home produced or locally owned production might
# not be counted in that, which in turn might underestimate the total value
# added per workers in the country

# Question 2 ----------------------
structure_2015 <- structural_change_transform |>
  filter(year == 2015)

sectors <- list("Agriculture", "Mining", "Manufacturing", "Utilities",
                "Construction", "Trade", "Transport", "Business",
                "Finance", "Realestate", "Government", "Other")


# Loop through sectors and perform division
sector_labor_product = list()
for (sector in sectors) {
  labor_prod = perform_division(sector, structure_2015)
  sector_labor_product = append(sector_labor_product, labor_prod)
}

# Convert sector_labor_product to numeric if it's not already numeric
if (!is.numeric(sector_labor_product)) {
  sector_labor_product <- as.numeric(sector_labor_product)
}

# Perform subtraction
sector_standardized <- sector_labor_product - 3068.131

# naming the standardized values
names(sector_standardized) <- sectors

# Normalized productivity of agriculture is -1983.73
# Employment share of agriculture is 43.2% (code for this in question 4)


# Question 3 ----------------------
```

```r
# Loop through sectors to get employment

#total employment in 2015
total_emp = structure_2015$Total_EMP

# creating employment list
employment = list()

structure_2015_emp <- structure_2015 |>
  select(ends_with("_EMP")) |>
  subset(select = -c(Total_EMP, Warflag_EMP))
for (col in names(structure_2015)) {
  # Append values of each column to the list
  employment[[col]] <- structure_2015_emp[[col]]
}

names(employment) <- sectors

# combining my lists of employment
keys <- unique(c(names(sector_standardized), names(employment)))
combined_prod <- setNames(mapply(c, sector_standardized[keys], employment[keys]), keys)

# changing to dataframe
structure_2015_long <- as.data.frame(combined_prod) |>
  pivot_longer(cols = Agriculture:Other,
               names_to = "Sector",
               values_to = "labor_prod")

# getting employment section
employment <- structure_2015_long[13:24,]

# dropping employment section from combined
structure_2015_long <- structure_2015_long[-c(13:24), ]

# column binding
structure_2015_long <- bind_cols(structure_2015_long, employment) |>
  rename(
    Sector = Sector...1,
    Sector_prod_stand = labor_prod...2,
    Employment = labor_prod...4) |>
  subset(select = -c(Sector...3)) |>
  arrange(Sector_prod_stand) |>
  mutate(cummulative_employment = cumsum(Employment)/sum(Employment)) |>
  arrange(Sector_prod_stand)

# plotting figure
structure_2015_long |>
  ggplot( aes(x = cummulative_employment, y = Sector_prod_stand, label = Sector)) +
  geom_step() +
  geom_label() +
  geom_point() +
  expand_limits(x = 0, y = 0) +
  labs(y = "Labor Productivity Relative to Total Labor Productivtiy \n (Per Capita)",
```

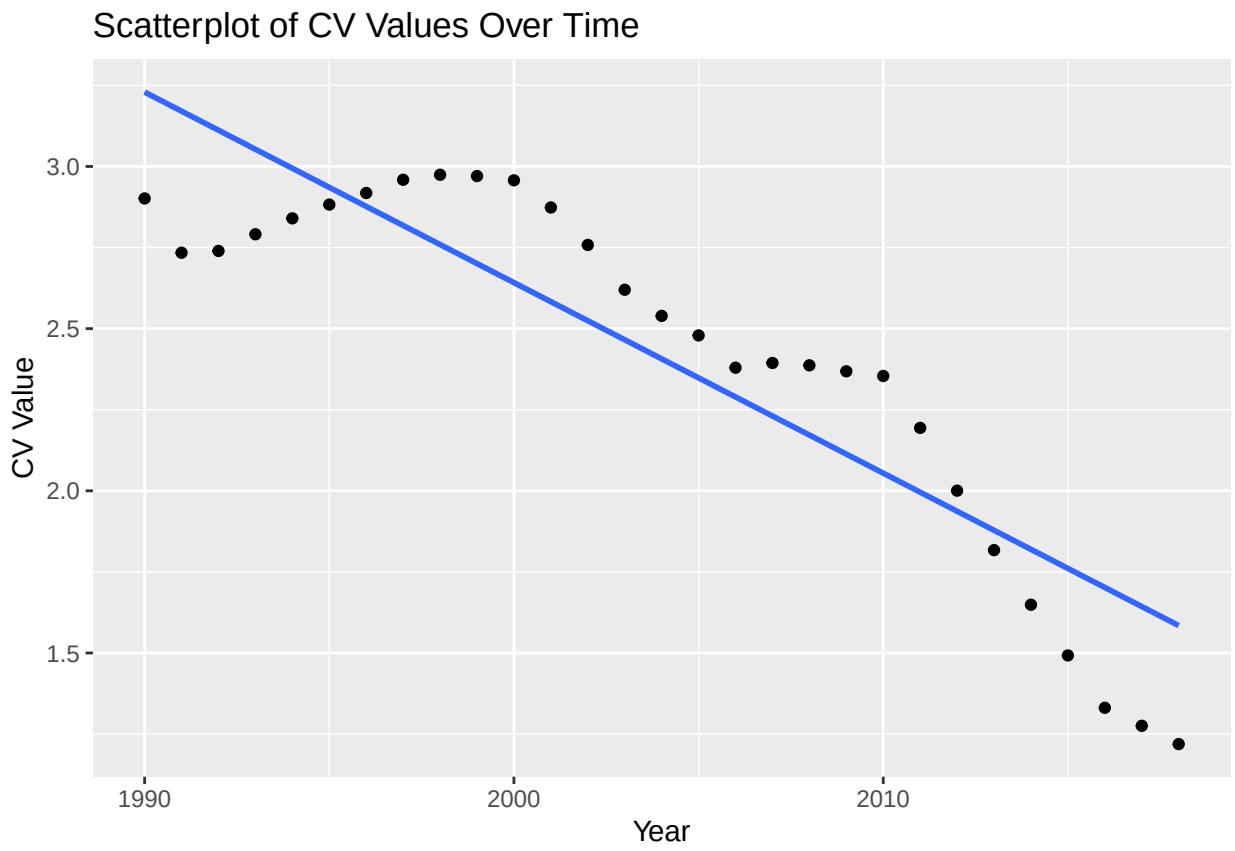
```
       x = "Cummulative employment")
```




```
# This figure shows us that the least productive sector, Agriculture, actually had the largest
# labor force share in Bangladesh in 2015.
# On the other hand, extremely productive sectors, such as Real Estate, Mining,
# Finance, and Business have minuscule shares of the labor force
# The combination of these phenomenons can drag down the overall labor productivity


# Question 4 ----------------------
# naming the non standardized values
# naming the standardized values
names(sector_labor_product) <- sectors

# order the sectors by order of production
sector_labor_product <- sort(sector_labor_product)

# Combining the non standardized values with employment
structure_2015_long <- structure_2015_long |>
  mutate(Sector_prod_non_stand = sector_labor_product,
         Employment_share = Employment/sum(Employment),
         weighted_sector_prod = Sector_prod_non_stand*Employment_share)

# weighted sum of productivity of sectors
weighted_prod <- sum(structure_2015_long$weighted_sector_prod)
```

```r
# flipping employment shares of manufacturing and agriculture
# first getting the values
agriculture_share <- structure_2015_long$Employment_share[1]
manufacturing_share <- structure_2015_long$Employment_share[3]

# now flipping
structure_2015_long$Employment_share[1] <- manufacturing_share
structure_2015_long$Employment_share[3] <- agriculture_share

# calculating the new weighted labor productivity
flipped_prod <- sum(structure_2015_long$Sector_prod_non_stand*structure_2015_long$Employment_share)

# percentage increase
percent_increase <- ((flipped_prod - weighted_prod)/weighted_prod) * 100
# The percentage increase after flipping productivity of agriculture and manufacturing is
# 22.5% (from 3068.131 to 3759.322)



#Question 5 ----------------------------

structural_change_bangladesh <- structural_change |>
  filter((var == "VA_Q15" | var == "EMP") & country == "Bangladesh")

# Find the rows where value added (VA) and employment (EMP) data start and end
va_start <- which(structural_change_bangladesh$var == "VA_Q15")[1]
va_end <- which(structural_change_bangladesh$var == "VA_Q15")[length(which(structural_change_bangladesh$
emp_start <- which(structural_change_bangladesh$var == "EMP")[1]
emp_end <- which(structural_change_bangladesh$var == "EMP")[length(which(structural_change_bangladesh$va

# Create Columns for Each Sector
for (sector in sectors) {
  structural_change_bangladesh[paste(sector, "Labor_Productivity", sep = "_")] <- NA
}

# Getting Labour Productivity for Each Year
for (year_index in va_start:va_end) {
  year <- structural_change_bangladesh$year[year_index]
  for (sector in sectors) {
    # Calculating
    productivity <- structural_change_bangladesh[year_index, sector] /
      structural_change_bangladesh[year_index + (emp_start - va_start), sector]
    # Putting in the column
    structural_change_bangladesh[year_index, paste(sector, "Labor_Productivity", sep = "_")] <- producti
  }
}

#Calculating CV:

#Columns to calculate:
start_col <- 19
end_col <- ncol(structural_change_bangladesh)
```

```r
cv_values <- numeric()

# For Loop To Calculate For Each Sector
for (i in 1:29) {
  # Extract labor productivity values for the current row
  row_values <- unlist(structural_change_bangladesh[i, start_col:end_col])
  # Calculate CV
  cv <- sd(row_values) / mean(row_values)
  cv_values <- c(cv_values, cv)
}

#Creating Dataframe for Graph:
Years <- 1990:2018
graph_data <- data.frame(Years, cv_values)

#Graphing:
ggplot(data = graph_data, aes(x = Years, y = cv_values)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE) +  # Add a linear regression line
  labs(title = "Scatterplot of CV Values Over Time", x = "Year", y = "CV Value")
```

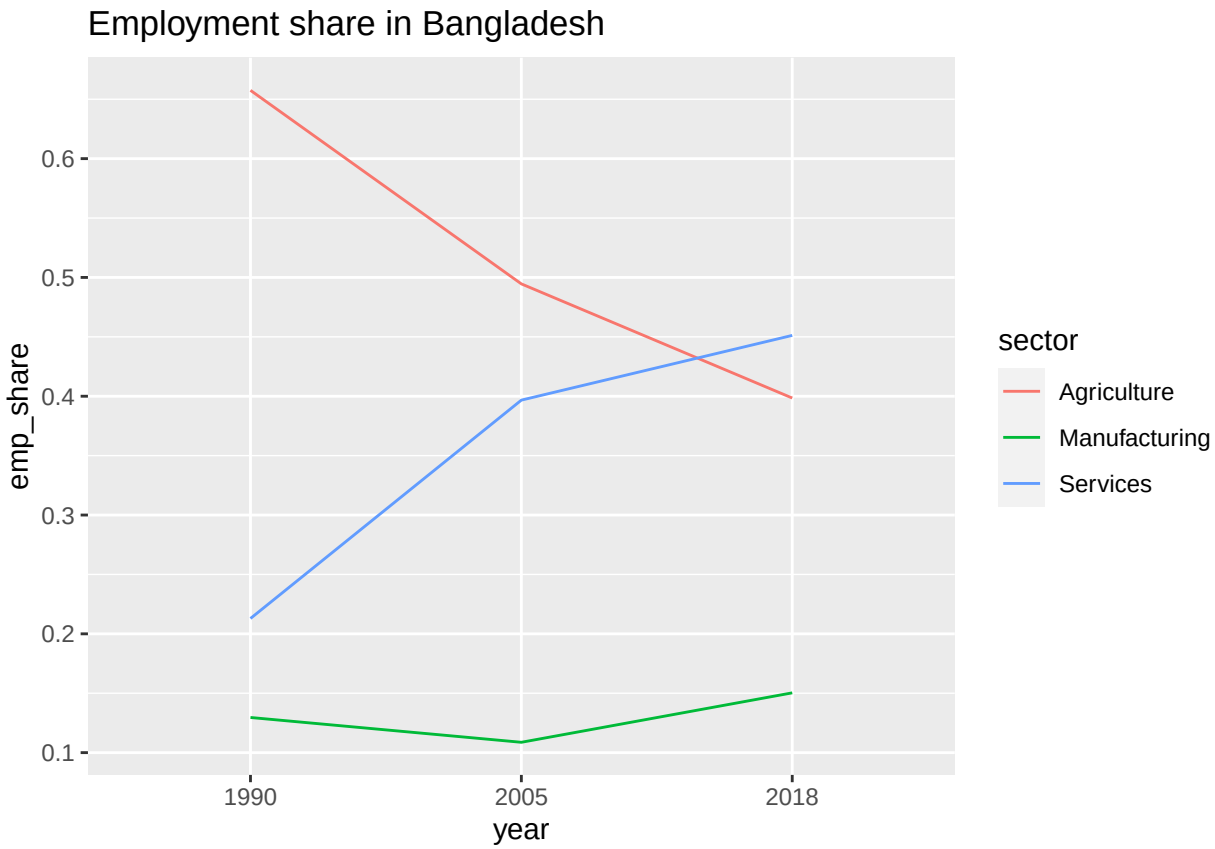



```r
#Interpretation:
# less variation in each sectors productivity, each sector was more
# stable in its output
#There has been reduced variability in the sectors overall, allowing a
```

```r
# stable environment for work to thrive
#Since variation in labour productivity in catching up to mean productivity,
#this could mean that the less productive sectors are catching up to the other
#sectors, creating increased GDP
#This would mean the nation looks more like advanced countries from our previous
#lecture, with less variation in sector productivity, than LIC's
#This could be the result of less productive sectors in 1990s becoming close to high productivity secto
#Since more people are leaving agricultural and other low productivity sectors and entering others, thi
#It could also because of transfer of agricultural technology to LMICs

#Question 6 ----------------------------
#_________________
# using transformed data
structural_change_bangladesh <- structural_change_transform |>
  filter(year == 1990 | year == 2005 | year == 2018)

#Aggregating Sectors:
structural_change_bangladesh <- structural_change_bangladesh |>
  mutate(mfg_va = Mining_VA_Q15 + Manufacturing_VA_Q15 + Utilities_VA_Q15,
         services_va = Construction_VA_Q15 + Trade_VA_Q15 + Transport_VA_Q15 +
           Business_VA_Q15 + Finance_VA_Q15 + Realestate_VA_Q15 + Government_VA_Q15 +
           Other_VA_Q15,
         mfg_emp = Mining_EMP + Manufacturing_EMP + Utilities_EMP,
         services_emp = Construction_EMP + Trade_EMP + Transport_EMP + Business_EMP +
           Finance_EMP + Realestate_EMP + Government_EMP + Other_EMP)


#Creating New Dataframe for further Questions:
productivity_decomp <- structural_change_bangladesh |>
  select(year, Agriculture_VA_Q15, Agriculture_EMP, mfg_va, mfg_emp,
         services_va, services_emp)

productivity_decomp$year <- as.numeric(productivity_decomp$year)
productivity_decomp <- productivity_decomp |>
  filter((year == "1990" | year == "2005" | year == "2018"))

#productivity_decomp$VAR <- c(rep("VA", 3), rep("EMP", 3))


# Q7 ------------------------------------------------------------------------

#Making a data frame in a format convenient to do calculations
sector <- c("Agriculture", "Agriculture", "Agriculture", "Manufacturing","Manufacturing", "Manufacturin
year <- c("1990", "2005", "2018","1990", "2005", "2018","1990", "2005", "2018")
emp <- c("26836213", "22984794", "25808861", "5289198", "5051255", "9735080", "8690267", "18437495", "2
va <- c("12232852827", "18822526115", "30887840910", "5682913328", "16003341476", "51965884144", "3104
df <- data.frame(sector, year, emp, va)

df$year <- as.character(df$year)
df$emp <- as.numeric(df$emp)
df$va <- as.numeric(df$va)

#Mutating necessary variables for calculations
```

```r
df <- df |>
  mutate(lp = va/emp)

df <- df |>
  group_by(year) |>
  mutate(L = sum(emp),
         emp_share = emp/L,
         )

df <- df |>
  group_by(year) |>
  mutate(VA = sum(va),
         LP = VA/L)

#For 1990-2005 across and within productivity decomposition

df_1 <- df |>
  filter(year == "1990" |year == "2005") |>
  group_by(sector) |>
  mutate(del_lp = case_when(year == "2005" ~ lp - lag(lp), TRUE ~ 0),
         del_LP = case_when(year == "2005" ~ LP - lag(LP), TRUE ~ 0),
         del_emp_share = case_when(year == "2005" ~ emp_share - lag(emp_share), TRUE ~ 0))

#within

df_1 <- df_1 |>
  group_by(sector) |>
  mutate(product = case_when(year == "2005" ~ lag(emp_share)*del_lp, TRUE ~ 0)) |>
  ungroup() |>
  mutate(numerator = sum(product),
         within = numerator/del_LP)

#Question 8 - within sector decomposition

df_1 <- df_1 |>
  group_by(sector) |>
  mutate(within_sec_decomp = product/del_LP)

#across

df_1 <- df_1 |>
  group_by(sector, year) |>
  mutate(product_2 = del_emp_share*lp) |>
  ungroup() |>
  mutate(numerator_2 = sum(product_2),
         across = numerator_2/del_LP)


#For 2005-2018 across and within productivity decomposition

df_2 <- df |>
  filter(year == "2005" |year == "2018") |>
  group_by(sector) |>
```

```r
  mutate(del_lp = case_when(year == "2018" ~ lp - lag(lp), TRUE ~ 0),
         del_LP = case_when(year == "2018" ~ LP - lag(LP), TRUE ~ 0),
         del_emp_share = case_when(year == "2018" ~ emp_share - lag(emp_share), TRUE ~ 0))

#within

df_2 <- df_2 |>
  group_by(sector) |>
  mutate(product = case_when(year == "2018" ~ lag(emp_share)*del_lp, TRUE ~ 0)) |>
  ungroup() |>
  mutate(numerator = sum(product),
         within = numerator/del_LP)

#Between 1990 and 2005, about 47% of productivity decomposition comes from across sector labor realloca
#And this illustrates the structural change narrative for Bangladesh going from being agriculture secto
#Whereas 52% of productivity decomposition comes from within sector reallocation accounted for by produ
#the ready made garment industry.

#In 2005 to 2018  timeframe, 71% of the decomposition is accounted for by within sector productivity gr
#This accompanied by structural change towards service sector (28%) could be Bangladesh's structural ch


#Question 8 - within sector decomposition --------------------

df_2 <- df_2 |>
  group_by(sector) |>
  mutate(within_sec_decomp = product/del_LP)

#across

df_2 <- df_2 |>
  group_by(sector, year) |>
  mutate(product_2 = del_emp_share*lp) |>
  ungroup() |>
  mutate(numerator_2 = sum(product_2),
         across = numerator_2/del_LP)

#Time Series Graph for employment share

ggplot(data = df, aes(x = year, y = emp_share, group = sector)) +
  geom_line(aes(color = sector)) +
  labs(title = "Employment share in Bangladesh")
```




```
#Within sector decomposition further helps us confirm that most of the productivity growth between 1990
#On the other hand, between 2005 and 2018, service sector accounted for most of the productivity growth
#It can also be noted from the graph that employment share in agriculture sector falls sharply 2005 onw
#But this is accompanied by a balanced increase in employment share in service sector.
#It can also be noted that manufacturing while not accounting for a major share in employment, still do


#Question 9 -------------------------------

#We reserve judgement for Bangladesh's prospects from these findings
#of their path of structural development. Like the methodology in
#Fan et al, we would be optimistic for Bangladesh if the growth
#in the services sector stemmed from productivity growth and not demand.
#This would mean that sustained growth for Bangladesh is a possibility. Since
#we do not have this particular model applied to Bangladesh, we reserve our judgement, but are
#cautiously optimistic of their path of structural development going from agriculture
#to services.
```