Assignment 4b- Zamima Islam Sabaa

STEP 1: Loading my packages

```
library(haven)
library(dplyr)

library(micEconAids)

library(haven)

#Loading all the consumer spending datasets and the CPI dataset

set2006<- read_sav("/Users/zamimaislam/Documents/SFU courses/Spring
2022/Econ836 - A. Econometrics/Assignm/Assignment 4b- Zamima Islam
Sabaa/shs2006.sav") %>% mutate(year=2006)
set2005<- read_sav("/Users/zamimaislam/Documents/SFU courses/Spring
2022/Econ836 - A. Econometrics/Assignm/Assignment 4b- Zamima Islam
Sabaa/shs2005.sav") %>% mutate(year=2005)
set2007<- read_sav("/Users/zamimaislam/Documents/SFU courses/Spring
2022/Econ836 - A. Econometrics/Assignm/Assignment 4b- Zamima Islam
Sabaa/SHS2007.sav") %>% mutate(year=2007)
set2008<- read_sav("/Users/zamimaislam/Documents/SFU courses/Spring
2022/Econ836 - A. Econometrics/Assignm/Assignment 4b- Zamima Islam
Sabaa/SHS2008.sav")  %>% mutate(year=2008)
set2009<- read_sav("/Users/zamimaislam/Documents/SFU courses/Spring
2022/Econ836 - A. Econometrics/Assignm/Assignment 4b- Zamima Islam
Sabaa/shs2009.sav") %>% mutate(year=2009)

prices <- read.csv("/Users/zamimaislam/Documents/SFU courses/Spring
2022/Econ836 - A. Econometrics/Assignm/Assignment 4b- Zamima Islam
Sabaa/prices table.csv")

#Binding all the consumer spending datasets into 1

mainset <- bind_rows(set2005,set2006,set2007,set2008,set2009)
```

Step 2: Cleaning the consumer spending dataset

```
mainset <-
  mainset %>%
  rename(province = 'PROVINCP',
         household_size = 'HHSZTOTP',
         household = 'CASEID',
         rooms = 'NUMRMP',
         type = 'TYPDWELP',
         tenure = 'TENURYRP',
         age_reference = 'RPAGEGRP',
```

```
         sex_reference = 'RPSEX',
         Food = 'F001',              #Total expenditure for food.
         Rented_shelter = 'G004',    #Expenditure on rented shelter.
         Clothing = 'J001',           #Total expenditure for clothing.
         Transportation = 'K001')     #Total expenditure for transportation.

# Grouping each year each province and each household size (filter for only
full year)

mainset <- mainset %>% filter(province !=0, #filters out the masked provinces
                              FYPYFLAG == 1) #filters households that have at
least one full-year member


# Transform consumer spending dataset from wide to long for later matching
with cpi ######################

mainset_long <-
  mainset %>% select(household,
                     year,
                     household_size,
                     province,
                     rooms,
                     type,
                     tenure,
                     age_reference,
                     sex_reference,
                     Food,
                     Rented_shelter,
                     Transportation,
                     Clothing) %>% gather(key='good_type',   #new column name
with the goods
                                          value='expenditure', #new column
that gives the spending from the goods
                                          c(Food, Rented_shelter,
Transportation, Clothing))
```

Step 3: Cleaning the prices got from Statcan data

```
#renaming provinces for matching

prices_long <- prices %>%
  gather('year',
         value='price',
         X2005,X2006,X2007,X2008,X2009) %>%
  mutate(year =
        case_when(
            year== 'X2005' ~ '2005',
            year== 'X2006' ~ '2006',
            year== 'X2007' ~ '2007',
```

```
            year== 'X2008' ~ '2008',
            year== 'X2009' ~ '2009'))

prices_long <- prices_long %>%
  rename(
         good_type= 'Products.and.product.groups')

prices_final <- prices_long %>%
  mutate(province = case_when(
    province== 'Newfoundland and Labrador' ~ 10,
    province== 'Prince Edward Island' ~ 11,
    province== 'Nova Scotia' ~ 12,
    province== 'New Brunswick' ~ 13,
    province== 'Quebec' ~ 24,
    province== 'Ontario' ~ 35,
    province== 'Manitoba' ~ 46,
    province== 'Saskatchewan' ~ 47,
    province== 'Alberta' ~ 48,
    province== 'British Columbia' ~ 59
  ))


prices_final$province <- as.numeric(prices_final$province)
prices_final$year <- as.numeric(prices_final$year)

mainset_long$province <- as.numeric(mainset_long$province)

matched_final <- prices_final %>% left_join(mainset_long,
by=c('year','province', 'good_type'))

matched_final <- filter(matched_final, good_type %in% c('Food', 'Shelter',
'Clothing', 'Transportation'))
matched_final1 <- matched_final %>% select(year,
                                           province,
                                           good_type,
                                           price,
                                           household,
                                           household_size,
                                           age_reference,
                                           sex_reference,
                                           expenditure)



matched_final1 <- matched_final1 %>%
  mutate(province = case_when(
    province== 10 ~ 'Newfoundland and Labrador',
    province== 11 ~ 'Prince Edward Island',
    province== 12 ~ 'Nova Scotia',
```

```
    province== 13 ~ 'New Brunswick',
    province== 24 ~ 'Quebec',
    province== 35 ~ 'Ontario',
    province== 46 ~ 'Manitoba',
    province== 47 ~ 'Saskatchewan',
    province== 48 ~ 'Alberta',
    province== 59 ~ 'British Columbia'
  ))

#Notes: All the datasets contain annual expenditure figures
```

Step 4: Calculating the impact of GST increase from 5 to 10 perc

```
#GST only effects clothing and transportation

#TYPE 1#############################################################################
#Crawford and Smith approach, pick a single region-year of data (that is
households facing just one price vector) and evaluating the cost-of-living
impact

Deaton_dataset <- matched_final1

matched_wide <- matched_final1 %>%
pivot_wider(names_from=good_type,values_from = c(price,expenditure))

matched_wide <- matched_wide %>%
  mutate(new_price_Transport=price_Transportation*1.10,
         new_price_Clothing=price_Clothing*1.10,
         price_Clothing=price_Clothing*1.05,
         price_Transportation=price_Transportation*1.05,
         quantity_Clothing_old=expenditure_Clothing/price_Clothing,
         quantity_Clothing_new=expenditure_Clothing/new_price_Clothing,

quantity_Transportation_old=expenditure_Transportation/price_Transportation,

quantity_Transport_new=expenditure_Transportation/new_price_Transport,
         quantity_Food=expenditure_Food/price_Food,
         Crawford_cost_of_Living_Index = ((price_Food*quantity_Food +
new_price_Transport*quantity_Transportation_old +
new_price_Clothing*quantity_Clothing_old)/(price_Food*quantity_Food +
price_Transportation*quantity_Transportation_old +
price_Clothing*quantity_Clothing_old)))

Crawford_dataset <- matched_wide %>% filter(province=='British Columbia',
                                            year %in% c('2009'))

write.csv(Crawford_dataset,file= 'Crawford Method Cost of Living Index.csv')
Deaton_dataset <- matched_wide

write.csv(Deaton_dataset,file='Deaton Dataset.csv')
```

```
Average_BC_Inflation_2009 <-
mean(Crawford_dataset$Crawford_cost_of_Living_Index)
Standard_Deviation <- sd(Crawford_dataset$Crawford_cost_of_Living_Index)
```

Step 5: I do the Deaton and Muelbauer Cost of Living Index using Stata

You will find my analysis in a separate STATA Do and Log file that I submitted along with these

Step 6: Compile the R and Stata files to create one data file to get the merged cost of living index from Crawford (I) and Deaton and Muelbauer (II)

I submittied the final data file as well, please focus on the columns named:

```
deaton_new <- read_dta("/Users/zamimaislam/Documents/SFU courses/Spring
2022/Econ836 - A. Econometrics/Assignm/Assignment 4b- Zamima Islam
Sabaa/Stata Deaton Analysis Data.dta")


write.csv(deaton_new, file = 'Crawford and Deaton and Muelbauer Dataset.csv')
```

Comments

In this assignment, I have looked at the effect of an increase in GST from 5 to 10 perc using SHS data from 2005–2009 using two methods, one followed by Crawford and Smith (I) and the other following Deaton and Muelbauer(II).

For the Crawford and Smith (I), I have estimated the effect on all years and all provinces however, I will focus to commentary on the effect only for the province of British Columbia in 2009 (because the assignment was to do this method for only single region-year)

There is some variation in the effect of the increase in GST among different households with different income distributions in British Columbia at the same point in 2009. The standard deviation in inflation or the cost of living increase is around 0.96 per cent compared with an annual average rate of inflation of about 2.52 per cent. This tells us that the average rate of inflation is not necessarily a good guide to the actual rates of inflation faced by individual households with different income distributions.

I have conducted the Deaton and Muelbauer (II) using STATA, but we can see the column for the Cost of Living Ratio in the Excel File and it effect of GST increase from 5 to 10 perc is done for all regions in all years.

What we see is that there is an effect on the total cost of living because of the GST increase in different households with different income distributions in there. This estimation is Deaton and Muelbauer demand equations, the AIDS, in which the budget shares of the various commodities are linearly related to the logarithm of real total expenditure and the logarithms of relative prices. The model is shown to possess most of the properties usually thought desirable in conventional demand analysis, and to do so in a way not matched by any single competing system. Since we control here for one demographic characteristic

(household size), what the values reflect is that households with different sizes also have an effect on how GST effects the Cost of Living since the values vary depending on the size of the Households too. If other demographic characteristics were controlled for, it would also be reflected in the data variations, suggesting that not just the current prices and current expenditures are the determinants of price changes.