Marginal propensity to consume

Alessandro Bramucci

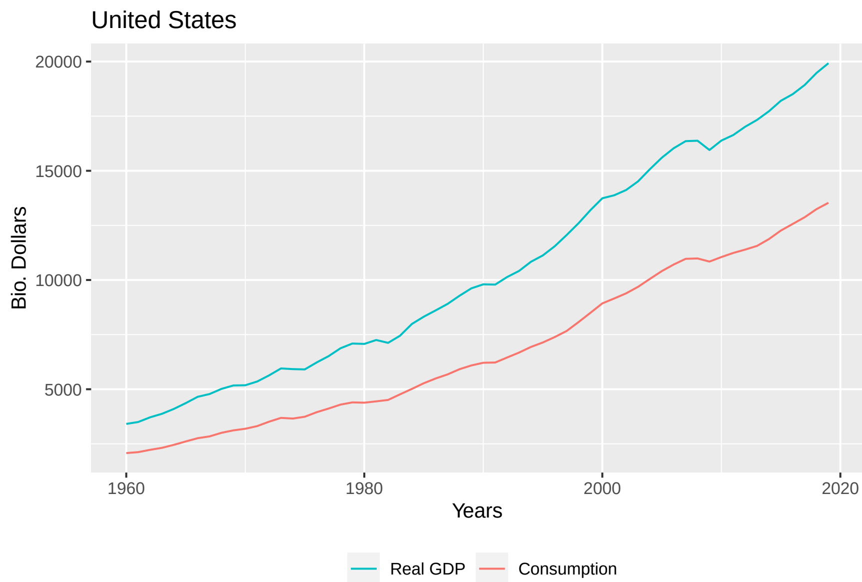
In this brief exercise we are going to estimate a Keynesian aggregate consumption function for the United States from 1960 to 2019. We will regress private consumption (at constant prices) on real GDP. The slope of the regression will be what in the economic literature is called the marginal propensity to consume (MPC for short). What is MPC? MPC is that value that tells us how much consumption will increase as income increases. Theory tells us that the MPC is a value between 0 and 1. When income increases by 1$, consumption will increase by a value less than the initial increase of 1$ but greater than 0.

From our Macroeconomics lessons we remember that the Keynesian consumption function is represented by the following equation, where C indicates the aggregate household consumption, Y the national income (both expressed in billions of constant dollars), c_Y is the MPC and c_{aut}, autonomous consumption, i.e. the part of consumption that does not depend on income.

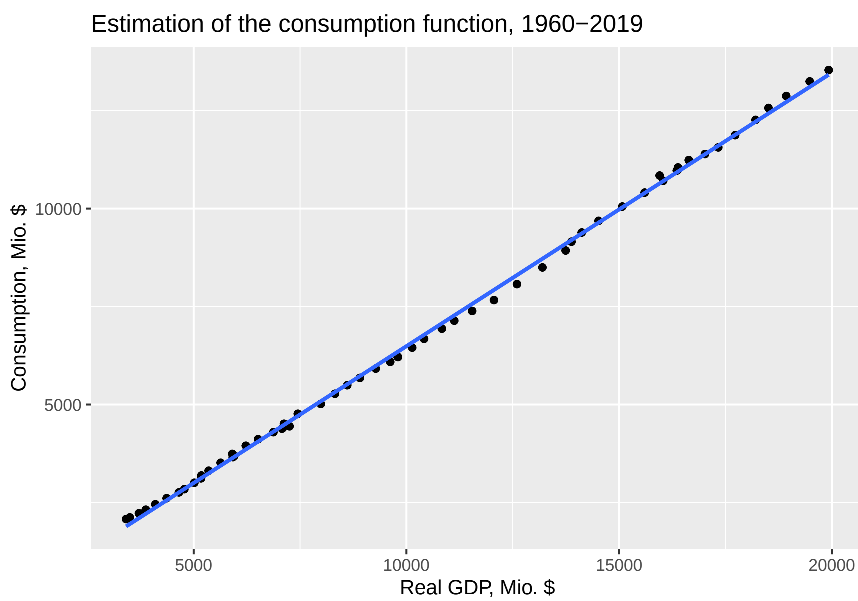
$$C = c_{aut} + c_Y Y$$

We first get the data using the *rdbnomics* package and then we will produce a graph of the time series from 1960 to 2019 with the *ggplot2* package. Both time series are expressed in billions of 2015 dollars.

```r
# Private final consumption expenditure at 2015 prices
df_cons <- rdb(ids = "AMECO/OCPH/USA.1.1.0.0.OCPH") %>%
  select(original_period, value) %>%
  rename(Year = original_period,
         cons = value) %>%
  filter(Year >= 1960 & Year <= 2019)

# Gross domestic product at 2015 reference levels [OVGD]
df_rGDP <- rdb(ids = "AMECO/OVGD/USA.1.1.0.0.OVGD") %>%
  select(original_period, value) %>%
  rename(Year = original_period,
         rGDP = value) %>%
  filter(Year >= 1960 & Year <= 2019)

# Joining data
df_final <- df_cons %>%
  left_join(df_rGDP, by = c("Year")) %>%
  mutate(ldcons = c(NA, diff(log(cons))),
         ldrGDP = c(NA, diff(log(rGDP))))

# Preparing the data for the graphs
df_graph_level <- df_final %>%
  pivot_longer(!Year, names_to = "Variable", values_to = "Value") %>%
  filter(Variable %in% c("rGDP", "cons")) %>%
  arrange(Variable)
```



We are now ready to perform the linear regression. We will also produce the scatter plot of the data to get the graphical intuition of the regression.



From the linear regression, we obtained the MPC. If income increases by 1$, consumption will increase by approximately 70 cents on average. Once the MPC is obtained, we can calculate the value of the Keynesian multiplier (m) as learned in Macroeconomics classes.

$$m = \frac{1}{1 - c_Y} = \frac{1}{1 - 0.7} = 3.33$$

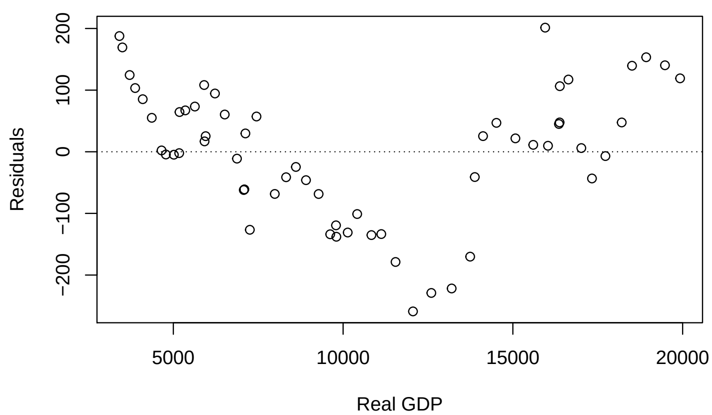
However, this regression has a number of problems from an econometric perspective.[1] The two series have a tendency to grow over time and the variables might seem highly correlated to us because they both have the same tendency to increase with time. This could lead to the conclusion that there is a correlation when in fact there is not (not the case here). This problem is known as **spurious regression**. We need to see if

[1]The value of R^2 almost equal to one must immediately raise doubts.

Table 1: Estimation of the consumption function in levels, 1960-2019.

	Dependent variable:
	cons
rGDP	0.698^{***}
	(0.003)
Constant	-492.110^{***}
	(32.683)
Observations	60
R^2	0.999
Note:	*p<0.1; **p<0.05; ***p<0.01

the independent variable is correlated with the error term. This is certainly an undesirable behavior. The residuals must be completely random and contain no predictive power.



We can observe a clear pattern between residuals and income. Our coefficients are biased. If we were in a cross-sectional context, we would have to figure out which variable is missing from the model and where possible include it. Or it could be that the model is misspecified. To get to the point, econometricians would say that our two series follow a **unit root process**. The series are highly persistent over time and contain a, in our case positive, time trend.[2] Wanting to simplify a lot we can say that our time series are not stationary and must be transformed before being used in a regression. A quick fix that works is to use logarithmic differences.[3]

[2]The two concepts, trending behaviour and persistent behaviour, should not be confused. Please refer to the Wooldridge textbook (Chapter 11).

[3]Differencing will removes the linear trend.

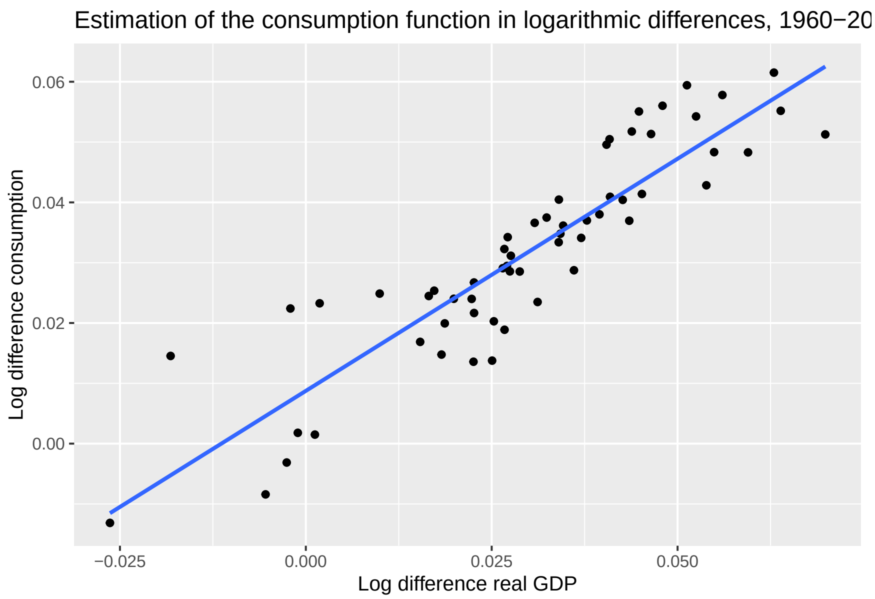


Table 2: Estimation of the consumption function in logarithmic differences, 1960-2019.

	Dependent variable:
	ldcons
ldrGDP	0.770***
	(0.050)
Constant	0.009***
	(0.002)
Observations	59
R^2	0.808
Note:	*p<0.1; **p<0.05; ***p<0.01

We can see that the estimate in the slope has changed slightly. This time it is somewhat higher at 0.77. However, the interpretation of the model has changed. This time we estimated an elasticity. How can we calculate the MPC having estimated the elasticity of consumption with respect to income? From courses in Microeconomics we remind that elasticity is nothing more than the ratio of two percentage rates and that on the demand curve elasticity varies depending on where you measure it.[4] In our case, our elasticity parameter (just called ϵ_Y for simplicity) is given by:

$$\epsilon_Y = \frac{\frac{\partial C}{C}}{\frac{\partial Y}{Y}} = \frac{\partial C}{\partial Y}\frac{Y}{C}$$

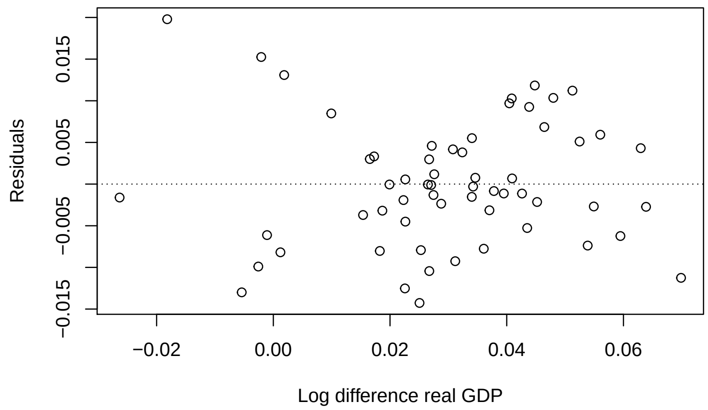
Rearranging the terms, we get:

$$\frac{\partial C}{\partial Y} = \epsilon_Y \frac{C}{Y}$$

[4]For a review of the elasticity concept in economics, see here.

The elasticity parameter is equal to 0.77 while the C/Y term that we will calculate as the average over the entire period from 1960 to 2019 is equal to 0.64. The marginal effect, the MPC, calculated on the avare over the entire period is therefore 0.4928.

How do the residuals perform this time?



There are actually a whole host of econometric issues that we have left out but that need to be properly addressed when estimating a model. Is there serial correlation in the residuals and if so what is the consequence? Is there heteroschedasticity in the residuals and if so what problems could it cause? There are tests appropriately developed by statisticians to identify these problems and solutions to fix them. All of these things will be the subject of the course in the following semester.