

Savings Rate Analysis

By undecided

msuga, yliang6, nsathan1, zliang4

Hypothesis

Savings rate is the ratio of savings to nominal GDP of the country. Too low savings rate can be a sign of trouble, for example, national savings may be insufficient to sustain a high level of long-run economic growth. Too high savings rate could mean low economic growth due to the lack of money circulating. We believe not only the two extremes but also any value of savings rate is an interesting component to investigate when observing the country's economy. Thus, we were going to see if there is a significant correlation between a general macroeconomic phenomena, e.g. unemployment rate, and the savings rate.

Data

We collected and joined data sets from existing .csv files from a few websites, the majority being from Worldbank.org, on each macroeconomic feature of each country in the world from 1960 to 2018. The data sets are on savings rate, as well as macroeconomic features of our interests, including unemployment rate, male-to-population ratio, female-to-population ratio, total population, population density, nominal GDP, nominal GDP per capita, GDP growth, age dependency ratio of young working age, literacy rate, and education. The data was very dirty, with many missing values from the early 1960s to 1990s and some rows not even being countries, but rather regions or groups of countries. We cleaned our data and decided to only use data from 2010 to 2018.

Findings

Claim #1: Multicollinearity exists among macroeconomic factors of our interests.

Support for Claim #1: At first, we ran multiple regression with all eleven macroeconomic features as independent variables and savings rate as a dependent variable and concluded the correlations based on p-value. However, the results were counter-intuitive. We suspected that some of the independent variables might be highly correlated, so we performed a correlation matrix on all macroeconomic features to find out. We considered any pair with absolute value of correlation coefficient greater than 0.3 to be correlated. The result, as shown in Figure 1, did indicate that there is multicollinearity in our macroeconomic factors.

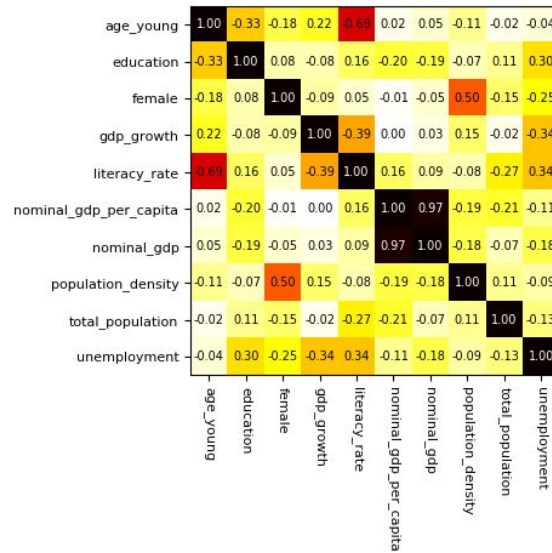


Figure 1. The correlation matrix

Claim #2: Age dependency ratio of young working age, GDP growth, total population, and population density are significantly correlated with savings rate, while nominal GDP per capita is not.

Support for Claim #2: Due to the finding in Claim #1, we removed redundant macroeconomic features that possibly added further noise to the model, based on the correlation matrix in Figure 1. Then, we ran multiple linear regression on the remaining features. We found that, for age dependency ratio, GDP growth, total population, and population density, there is a significant correlation with savings rate because their p-values are less than 0.05. The correlation between nominal GDP per capita and savings rate is, however, not significant, with p-value = 0.672. Figure 2 depicts our regression result.

OLS Regression Results						
=====						
Dep. Variable:	y	R-squared (uncentered):		0.642		
Model:	OLS	Adj. R-squared (uncentered):		0.640		
Method:	Least Squares	F-statistic:		369.8		
Date:	Sat, 02 May 2020	Prob (F-statistic):		3.35e-227		
Time:	18:55:25	Log-Likelihood:		-4231.8		
No. Observations:	1038	AIC:		8474.		
Df Residuals:	1033	BIC:		8498.		
Df Model:	5					
Covariance Type:	nonrobust					
=====						
		coef	std err	t	P> t	[0.025 0.975]

age dependency ratio, young (%)		0.2070	0.014	14.557	0.000	0.179 0.235
GDP growth (%)		1.3591	0.151	9.017	0.000	1.063 1.655
nominal GDP per capita		6.426e-08	1.52e-07	0.423	0.672	-2.33e-07 3.62e-07
population density (people per square km)		0.0046	0.001	6.702	0.000	0.003 0.006
total population		6.854e-08	9.98e-09	6.867	0.000	4.9e-08 8.81e-08
=====						
Omnibus:	32.414	Durbin-Watson:		1.828		
Prob(Omnibus):	0.000	Jarque-Bera (JB):		77.382		
Skew:	-0.062	Prob(JB):		1.57e-17		
Kurtosis:	4.332	Cond. No.		1.72e+07		
=====						

Figure 2. The multiple linear regression result