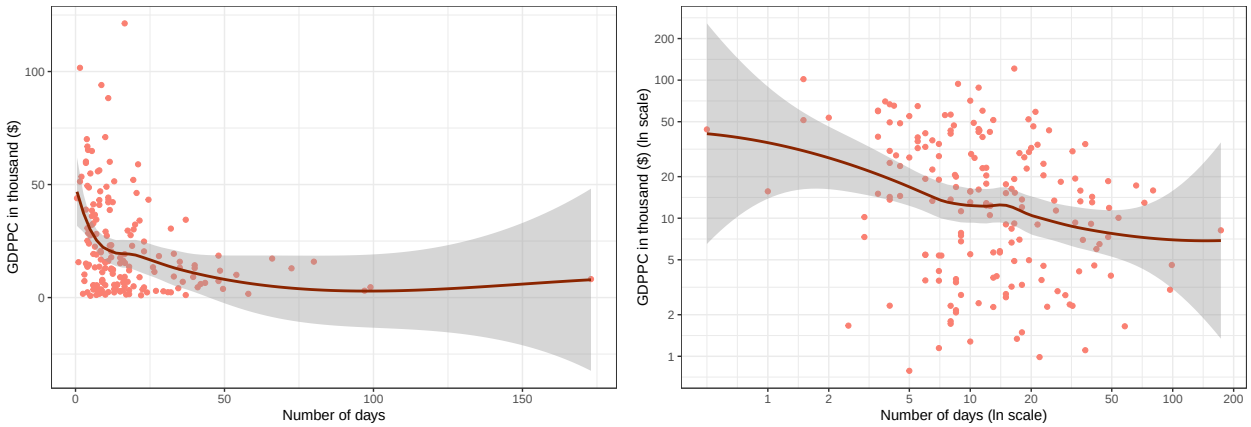
Analysis between the GDPPC (in PPP) and the days required to start a new business

Maeva Braeckevelt

Introduction

This analysis aimed at analyzing the pattern of association between the GDPPC (in PPP) and the days required to start a new business. The data used was gathered on the World Bank'sites. The main variables that I used were: the GDDPPC of countries in thousand of dollars US (y) and the number of days required to start a new business (x). I decided to also compare level-level to log-log transformation. I will do a non-parametric regression and then a linear regression

Non-parametric regression : level-level and log-log



I observe that the slope is negatif. But it is not easy to see and not very precise. I can already say that there few chances that there is a higher association between GDPPC and the days required to start a business. The trend of the regression is a bit easier to see in the log-log models. I will do a regression to have some quantitative analysis.

Linear regression level-level and log-log

level-level

Formula : GDPPC= 26,32 -0,26 * days

Alpha : 26,32 is the average of GDppc (in thousand $) of a country when the days to start a business is equal to 0.

Beta : One additional day is associated with a 0,26 (thousand $) smaller GDPPC

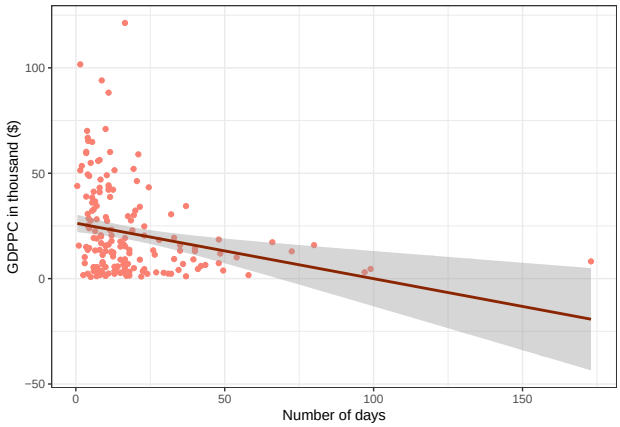
log-log

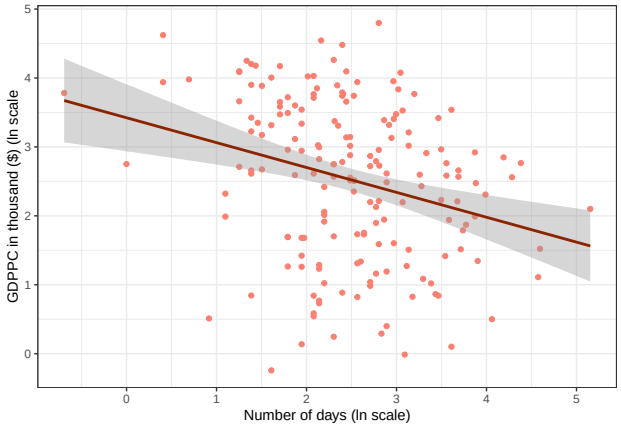
Formula : GDPPC= 26,32 -0,26 * days

Alpha : 3,42 is the average of the ln_GDppc (in thousand $) of a country when the days to start a business is equal to 1 (ln(0)=1.

Beta : 1 % higher change in days, is associated with a 0,36% smaller GDPPC, on average.

The R squared adjusted are both very small. 5% pour level-level and 7% for log-log. So the log-log model fit a bit better but the regression still doesn't captures very well the regression. To conlude, the business-friendliness of a country affect the income it generateS but there is probably some other factors that have a higher association with the GDPPC of a country.





	Linear	Linear log-log
(Intercept)	26.32 ***	3.42 ***
	(2.41)	(0.23)
days	-0.26 **	
	(0.08)	
ln_days		-0.36 ***
		(0.08)
nobs	178	178
r.squared	0.06	0.08
adj.r.squared	0.05	0.07
statistic	9.71	19.78
p.value	0.00	0.00
df.residual	176.00	176.00
nobs.1	178.00	178.00
se_type	HC2.00	HC2.00

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.