The effect of corruption on economic growth in developing countries

Qian Zhang | timqzhang@uchicago.edu | Github: timqzhang
Master in Computational Social Science, University of Chicago
June 3, 2020

MASTERS IN COMPUTATIONAL SOCIAL SCIENCE
THE UNIVERSITY OF CHICAGO

Research Question

How does corruption affect economic growth in developing countries?

- How to obtain a consistent and continuous measurement of corruption?
- How to estimate the impact of corruption on economic growth in developing countries?
- What is the best statistical learning model for prediction on economic growth?

Limitation of current corruption index

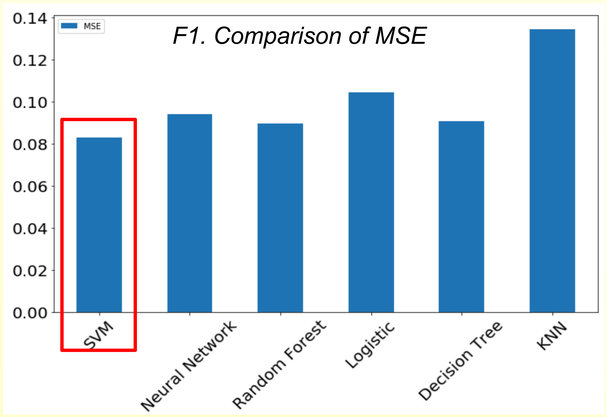
1. Different methodologies in each existing index.
2. Failing to include some important factors.
3. Current indices are not continuous.

Reconstruction of corruption: SVM

- **Theoretical framework: SVM [1]**
  - Corruption $c_{i,t}$ for certain country i at period t:
  - $c_{i,t} = F\left(x_{i,t}^1, \cdots, x_{i,t}^m\right) \quad \forall(i,t)$
- **Algorithms**

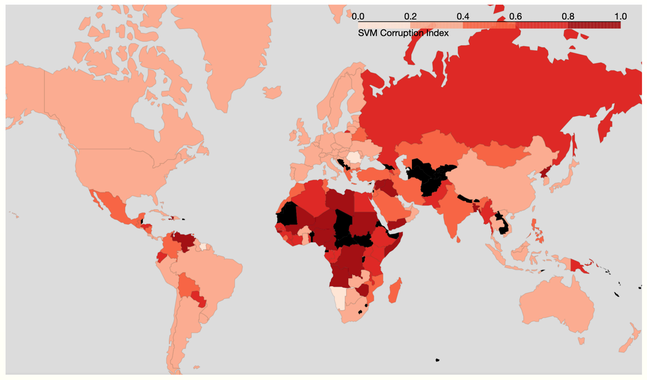
1. Select ***m*** conditions for function
2. Get a sample set by coding country-year pairs as corruptive (1) or not corruptive (0).
3. Fit the SVM model and apply the approximated function to all country-year pairs.

- **Statistical justification:**

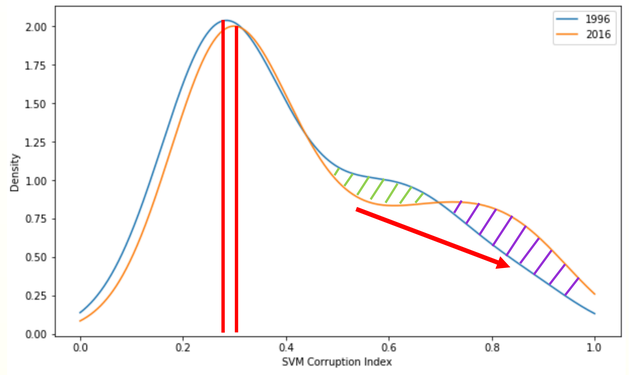


F1. Comparison of MSE

Estimated SVM Corruption index: Visualization



F2. SVM Corruption index around world, 2016



F3. Kernel density estimate of SVM corruption, 1996 and 2016

Impact of corruption on growth: Model specification

- **A Dynamic Panel Data (DPD) model** is implemented. Baseline model is written as

$$y_{i,t} = \beta y_{i,t-1} + \gamma c_{i,t} + \theta \boldsymbol{X}_{i,t} + \mu_i + \xi_t + \epsilon_{i,t},$$

- $y_{i,t}$ denotes the log form of GDP per capita, $c_{i,t}$ denotes the corruption indicator, $\boldsymbol{X}$ denotes all other control variables, μ_i denotes country's fixed effect, and ξ_t denotes time's fixed effect .
- **Two-step "Difference" GMM** is applied for estimation. [2]

Data and Variables: Statistic summary

Variables	Obs	Mean	Std.Dev.
GDP Per Capita	2630	5448.30	8370.52
Corruption Index	1974	0.50	0.22
Regulatory Quality	1948	2.21	0.86
Inflation%	2623	13.44	111.67
Gov Consumption%	2464	14.66	7.21
Trade Openness%	2553	82.55	50.86
Capital Investment%	2464	22.51	8.23
FDI inflow%	2605	4.38	7.46
Natural Resource%	2630	10.19	12.85

T1. Summary Statistics

Empirical results

Dependent: lgdp_pcap	(1)	(2)	(3)	(4)
Lagged lgdp_pcap	0.96*** (0.0017)	0.96*** (0.0018)	0.97*** (0.0026)	0.97*** (0.0026)
Corruption	-0.085*** (0.0062)	0.37*** (0.030)	0.46*** (0.026)	0.50*** (0.026)
Corr^2		-0.44*** (0.023)	-0.50*** (0.022)	-0.41*** (0.044)
Reg_Qual			-0.015*** (0.0031)	0.014** (0.0067)
Corr ×Reg_Qual				-0.053*** (0.012)

** p<0.1, ** p<0.05, *** p<0.01*

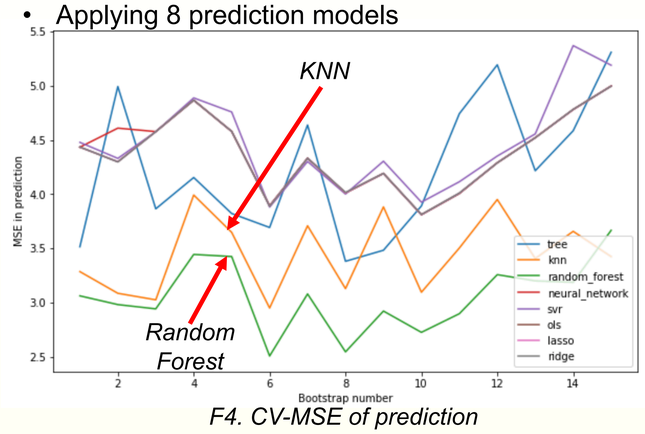
T2. GMM estimation results

Corruption's overall effect on column 4 :
0.50-0.82×Corr-0.053×Reg_Qual

Given high level corruption (0.50-0.82×Corr < 0),
Poorer Reg_Qual enlarges corruption's negative effect.

More results: Prediction on growth

- Randomly splitting dataset to get 15 bootstraps
- Applying 8 prediction models



F4. CV-MSE of prediction

- Best prediction models: **Random Forest & KNN**

Conclusions

1. There is a downward quadratic relation between corruption and growth.
2. Poorer regulatory quality could enlarge the negative influence of corruption on growth.
3. Best prediction models on economic growth are Random Forest and KNN.

Limitations & Future work

1. The diff-GMM, though much better than sys-GMM, still has IV proliferation which weakens the power of IV related test (Sargan test).
2. More approaches and algorithms could be explored to measure the corruption more accurately other than SVM in this work.

References

[1] Gründler, K., & Krieger, T. (2016). Democracy and growth: Evidence from a machine learning indicator. *European Journal of Political Economy*, *45*, 85–107.
[2] Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, *58*(2), 277–297.

Acknowledgements

I would like to express my sincere gratitude to Dr. Richard Evans for his continuous support and helpful comments on my work. Also, I want to thank all classmates in MACS 30250 this Spring.