How does corruption affect economic growth in the developing countries? A machine learning approach

MACS 30250 Project proposal

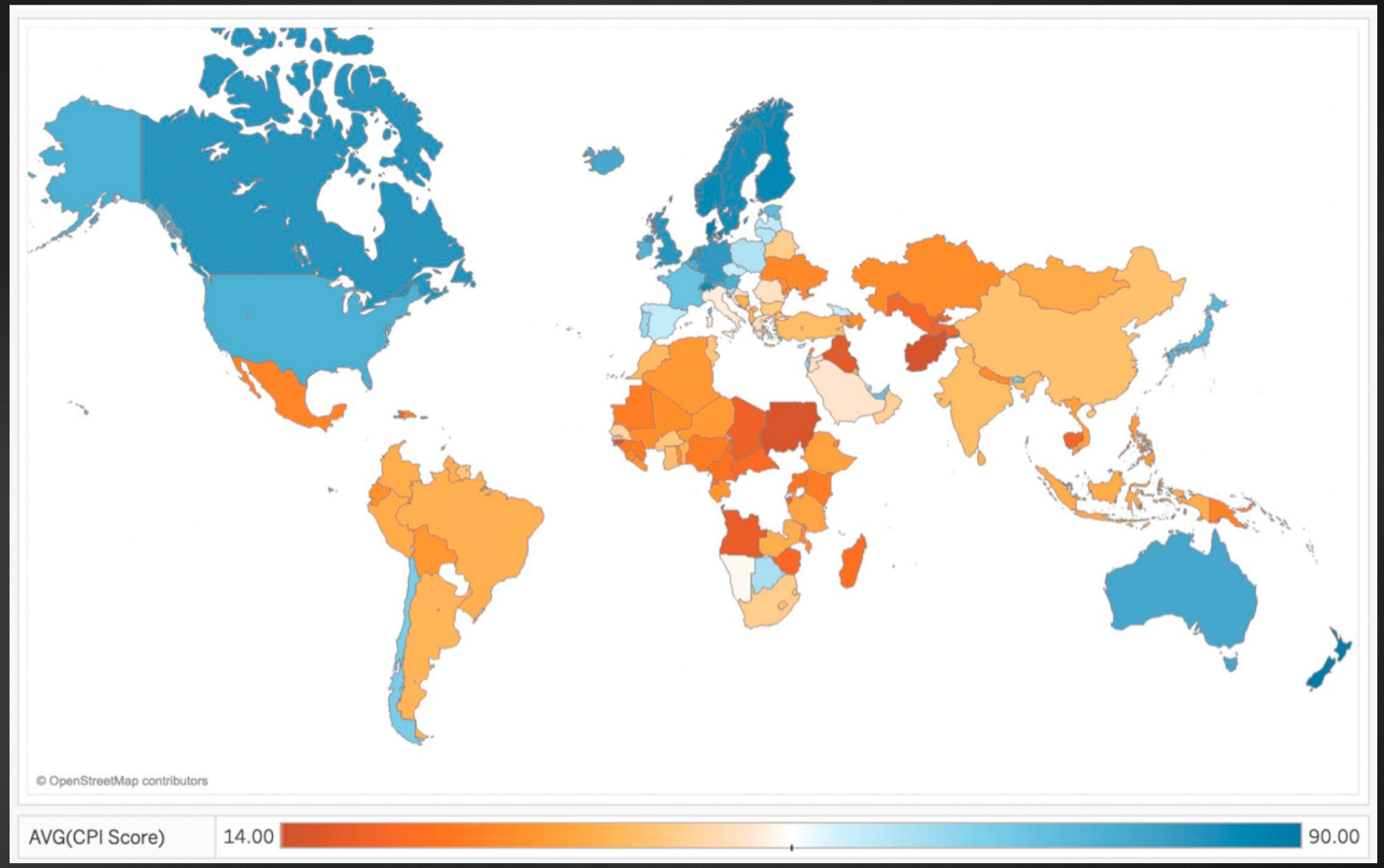
Tim, Qian ZHANG

Research Question

- This project mainly researches on how the economic growth would be influenced if the government is corruptive in developing countries.

- Main tasks:
  - **Interpretation:** Figure out the impact of corruption on economic growth in developing countries via GMM estimations.
  - **Prediction:** Introduce machine learning models and compare prediction capabilities on economic growth using variables included in project.

Motivation



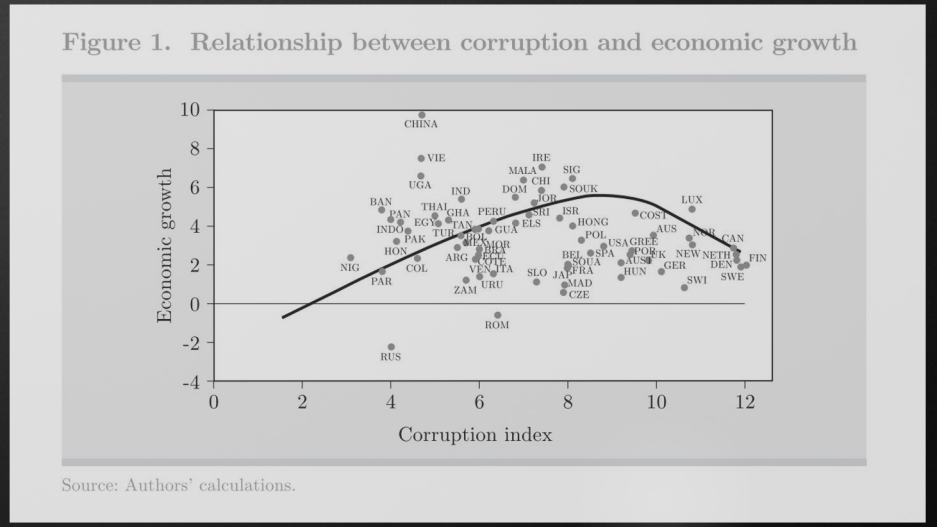
Corruption around the world: Retrieved from Lima & Delen (2020)

Theory

- Contradicting views first appeared in 20th century:
  - **Leff (1964): "Grease the wheel" hypothesis**
    - Corruption may be beneficial in **a second best world** by alleviating the distortions caused by **ill-functioning institutions**.
  - **Mauro (1995): "Sand the wheel" hypothesis**
    - Corruption will only **reduce and does not** improve economic growth.
- It continued in 21st century, but two views may integrate:
  - **Méon & Weill (2010):** Corruption is beneficial in countries where their governance are defective, but remains detrimental elsewhere.

Empirical results

- **Positive role:** Grandgirard, et al. (2003), Huang (2016), Biru (2010)
- **Negative role:** Gründler & Potrafke (2019) , Frimpong, et al. (2019), Sharma & Mitra (2019), Swaleheen (2011)
- **No significant effect:** Glaeser & Saks (2006), Treisman (2007)
- **An optimal level of corruption (inversed U shape):** Acemoglu & Verdier (2001), Ahmad, et al. (2012)



Source: Authors' calculations.

Figure retrieved from Ahmad, et al. (2012)

Divergence in empirics

- **Why such a long-lasting debate?**
  - Usage on proxies of corruption from different sources
    - Corruption Perceptions Index
    - World Bank World Governance Indicator
    - The International Country Risk Guide
    - ……
  - Results comes from specific techniques of estimations, or focus on too specific regions.

Possible contributions

- Explore a uniform corruption indicator by machine learning construction.
  - Gründler & Krieger (2016)
  - Smola & Schölkopf (2004)
  - Lima & Delen (2020)

- Consider the methodologies from different papers to form base model.
- Bring new data to focus on the whole group of developing countries, instead of just one/two specific ones.
- Introduce statistical learning models to make growth prediction.

Data and Variables

Indicators	Definition	Representation
GDP	GDP Per Capita	Economic Growth
Corruption	Corruption Index	Corruption level
Trade	(Imports + Exports)/GDP	Trade openness
Government consumption	General government final consumption Expenditure/GDP	Government consumption level
Inflation	Inflation rate	Inflation level
Foreign Direct Investment	Net flow of FDI/GDP	Foreign Direct Investments
Capital Investment	Gross fixed capital formation/GDP	Infrastructure
Natural resource	Total natural resources rent/GDP	Natural resource level
Regulator quality	Index of regulator quality	Quality of governance

Data Source:
World Development Indicators (WDI)
Transparency International
World Bank World Governance Indicator
Political Risk Service Group

* The selection of developing countries is based on the classification in *World Economic Situation and Prospects 2019* by United Nations.

Methods/Models

- Baseline model - inspired by Frimpong et al. (2019), Sharma & Mitra (2019)

- $ln\,GDP_{it} = \beta lnGDP_{it-1} + \gamma_1 CI_{it} + \gamma_2 (CI^2)_{it} + \delta CI_{it} \times RQ_{it} + \theta \boldsymbol{X_{it}} + \mu_i + \epsilon_{it}$

* GDP_{it} is in the log form, CI_{it} represents corruption index, RQ_{it} represents index of regulator quality, μ_i is country's fixed effect, and $\boldsymbol{X_{it}}$ are all other control variables.

- Estimations
  - difference-GMM (Arellano & Bond, 1991)
  - system-GMM (Blundell & Bond, 1998)

Expected result

- Probably verify “sand to wheel” hypothesis: Corruption has a **negative** effect on economic growth.
- “Grease to wheel” and “optimal corruption” are still possible.

Beyond interpretation: prediction

- Baseline model
- Machine learning algorithms
  - Neural Network (Lima & Delen (2020)) , Decision Tree, Random Forest.