Progress Report for MPhil Thesis

Tilman Graff – University of Oxford

November 23, 2017

Research Question

Which factors influence the global distribution of trade network optimality?

In this thesis, I aim to create a (potentially) global dataset of trade network efficiency. Taking the spatial distribution of current economic activity and population as given, I use a model from a recent working paper to determine the optimal trade network for each country. I then compare each country's current road network to its optimal one and derive a measure of how far a country is currently away from its ideal self.

In a second step, I will then investigate the origins of this global distribution. Which factors led to the heterogeneities among countries today? Specifically, I will look at:

- Do networks with large colonial infrastructure investments do better or worse today?

- Does tribal favoritism explain why some areas are lacking lucrative investment?

Process

In order to conduct this research, I will need to follow a series of steps and transfer data between multiple programming softwares. Here is a step-by-step guide:

- ✓ Find global raster data on population, night-lights, ruggedness, and colonial infrastructure investments.

- Grid the world on 50x50km squares and aggregate finer-resolution data into those grids.

- Locate the maximum population point within each grid and call this the (population) centroid of the grid.

- ✓ Use OpenStreetMap to find distance and average speed between neighboring gridcells.

- Use distance and ruggedness to calculate Infrastructure Building Cost Matrix $\delta^I_{i,k}$ for every country.

- Use average speed to calculate current Infrastructure Matrix $I_{i,k}$ for every country.

- Use distance to calculate (iceberg) Trade Cost Matrix $\delta^\tau_{i,k}$ for every country.

- Use $\delta_{i,k}^{I}$, $\delta_{i,k}^{\tau}$, and $I_{i,k}$ to find the optimal trade network $I_{i,k}^{*}$ and optimal tradeflows $Q_{i,k}^{*}$ for every country. This directly follows the Schaal Working Paper. I know how to do this, I am however afraid that this might take too long.

- Compare $I_{i,k}$ and $I_{i,k}^{*}$ for every country to obtain a measure of network optimality ζ_c for every country c.

- Investigate heterogeneity in ζ_c

This flowchart visualises the process. It shows the path input data (red circles) take through various programming languages (white rectangles) and intermediate datasets (blue circles) into eventual findings (green circle).