

Simulating How Remote Learning and Shared Education Could Reduce Regional Education Inequality

Yuming Liu

MACSS Project Proposal
April 14, 2020

Motivation

- Regional Poverty has been a major issue in China for decades since the reform and opening-up.
- Spatial Inequality is in areas like income, education, and health care.

Introduction

- Focus on the economical perspective on education. The data would be closely related to economy.
- Introduce a new method to measure education inequality inspired by the Education Gini model from Thomas, Wang, and Fan(2000).

Regions of Focus

Beijing, Shanghai, Hebei, and Henan.

Theory

- The Direct Method for Calculating Gini Coefficient (Deaton 1997).

$$GINI = \frac{1}{\mu N(N-1)} \sum_{i>j} \sum_j |y_i - y_j|$$

- Where $GINI$ is the Gini index;
 μ is the mean of variable;
 N is the total number of observations;
for education Gini, y_i and y_j are years of schooling of individuals.

Theory

- The formula for Calculating Education Gini (Thomas, Wang, Fan 2000).

$$E_L = \frac{1}{\mu} \sum_{i=2}^n \sum_{j=1}^{i-1} p_i |y_i - y_j| p_j$$

- Where E_L is the education Gini based on education attainment distribution;
 μ is average years of schooling for the sample population;
 p_i and p_j stand for the proportions of population with certain levels of schooling;
 y_i and y_j are years of schooling of specific levels;
 n is the number of levels.

Data

- The data we use includes:
 - Education Census data (2019)
 - Income data (2019)
 - Educational Consuming data (2018)
 - Number of Active Educational Institutions (2020)
 - Number of New Remote Learning Resources during COVID-19 pandemic (2020).

Data Table

Education Levels	Beijing	Shanghai	Henan
Uneducated	0.021	0.028	0.053
Primary School	0.086	0.124	0.241
Middle School	0.207	0.315	0.441
High School	0.200	0.215	0.172
College and Above	0.487	0.317	0.094
Consuming Data per person	Beijing	Shanghai	Henan
Annual Total	42926	43351	15169
Education and Entertainment	4209	4263	1685

Note: Education Data (%) from 2019, Consuming Data (RMB) from 2018

Regression Model

- Base linear regression model

$$E_{Li} = \beta_0 + \beta_1 \text{Income}_i + \beta_2 \text{EduCon}_i + \beta_3 \text{km}^2 \text{Edu}_i + \epsilon$$

- Where the dependent variable E_{Li} is the education Gini based on education attainment distribution in area i ;
 Income_i is the average income per household in area i ;
 EduCon_i is the average educational consuming per household in area i ;
 $\text{km}^2 \text{Edu}_i$ is the number of active educational institutions per km^2 in area i .

Results

- Apply the data to the regression model and check what the coverage rate is.
- Analysis how remote learning could have changed the Education Gini index for regions with lower Education Gini.
- Check how education inequality could be balanced by remote learning and education sharing.

Summary

- The results of the model suggest a way to deal with education inequality and help the government to formulate policies.
- Weakness of the model
 - The data about remote learning is limited at this moment.
- Future discussion about the research
 - More data could be added to enrich the regression model.
 - Other types of computational models could also be tested for the data.