ERSA Presentation

Re-estimate economic convergence in China

Using satellite nightlight data

Presenter: YANG Wenxuan (Tim)

PhD Student
Graduate School of International Development
Nagoya University , Japan

yang.wenxuan@h0.nagoya-u.jp

Outline

Objective

To fill in missing GDP data at city level in China, and explore how spatial relation impact economic convergence from 2012 to 2019

Methods

Absolute beta convergence, spatial dependence model, spatial heterogeneity model

Data

GDP and nightlight data from 2012-2019, 344 units at city level

Outline

Finding

1. Nightlighting data can be a proxy for **GDP per capita** in China. It will solve missing value issue at city level.
2. There is **absolute beta convergence** in China from 2012 to 2019. Spatial effect **accelerates convergence** during this period.
3. Local convergence speed is heterogenous, with higher convergence speed in **west and central China**, lower convergence speed in **north and southeast.**

Outline

Research Steps

1. Construct a new and completed dataset by using nightlight data(NTL data)
2. Compare and see how spatial effects influence economic convergence
3. Explore spatial heterogenality in economic convergence process

Step 1. Why we need a new dataset?

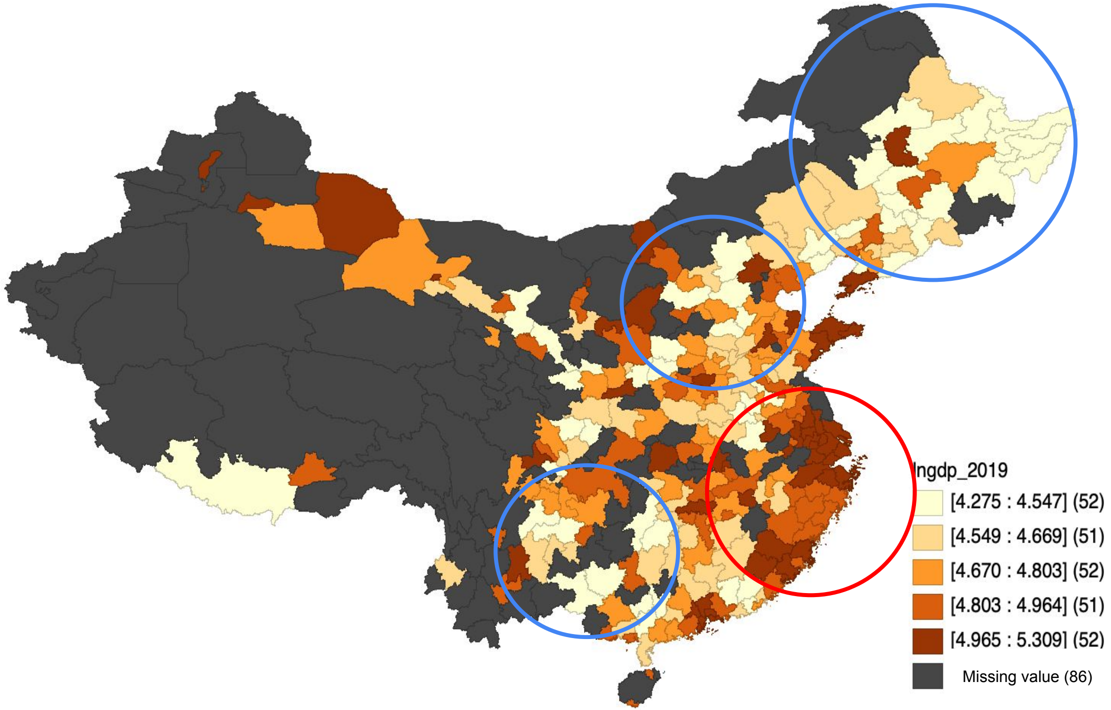
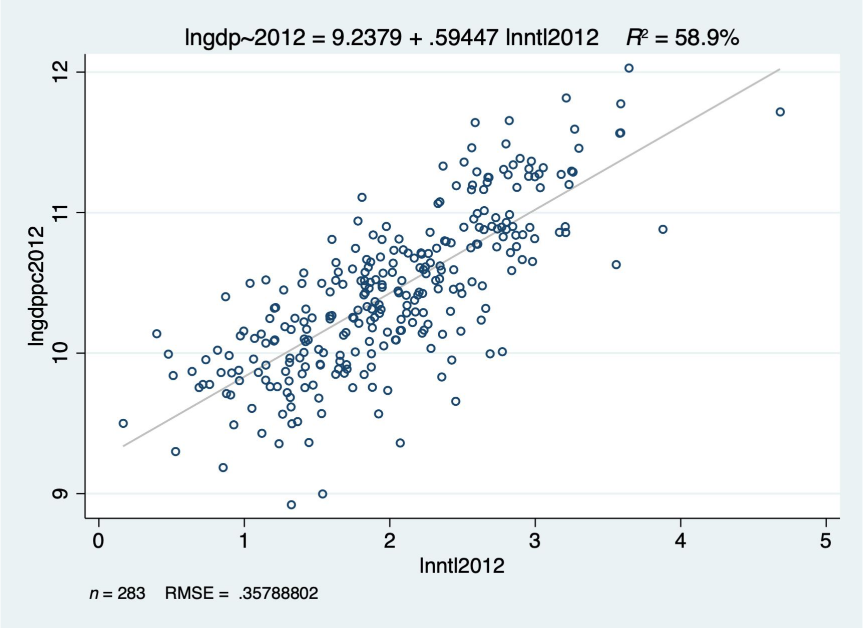


Fig.1 Regional GDP per capita in China in 2019(from statistical yearbook)

Fewer research is using <u>city level data</u>, this is because there are too many missing values at city level(over 25%).

Step 1. Why I use nightlight(NTL) data?



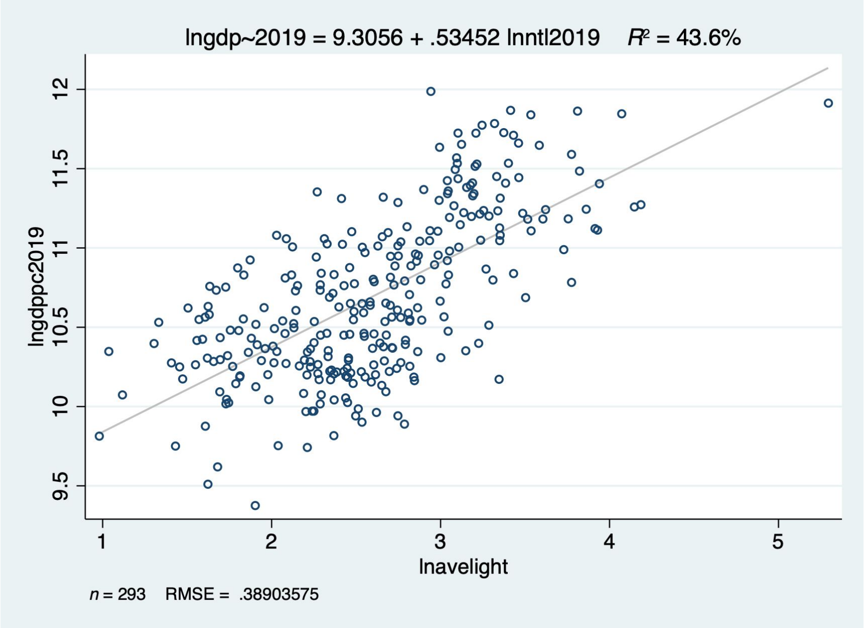


Fig.2 Scatter plot of light p.c. and GDP p.c. in 2012 and 2019

Step 1. Include **provincial data** for a better estimation

$$\log(GDPpc_i) = \beta_0 + \beta_1 \log(light_i) + \beta_2 \log(GDPpc_j) + \varepsilon$$ **i:city, j: province**

*This equation refers to Lessmann and Seidel(2017)

Table 2. Estimation of city GDP p.c. in 2019

Dependent variable: $\log(GDPpc_i)$	2012		2019		Difference (2012-2019)	
	OLS	Controlled provincial GDP	OLS	Controlled provincial GDP	OLS	Controlled provincial GDP
$\log(light_i)$	0.594***	0.529***	0.535***	0.448***	0.193***	0.169***
$\log(GDPpc_j)$		0.432***		0.650***		1.682***
Constant	9.238***	4.869***	9.306***	2.514***	0.154***	−0.488***
Obs.	283	283	293	293	282	282
R-Squared	$\boxed{0.590}$	$\boxed{0.645}$	$\boxed{0.436}$	$\boxed{0.607}$	0.027	0.537
AIC	$\boxed{223.527}$	$\boxed{184.066}$	$\boxed{280.258}$	$\boxed{176.193}$	116.120	-91.261

Step 1. Build a completed dataset

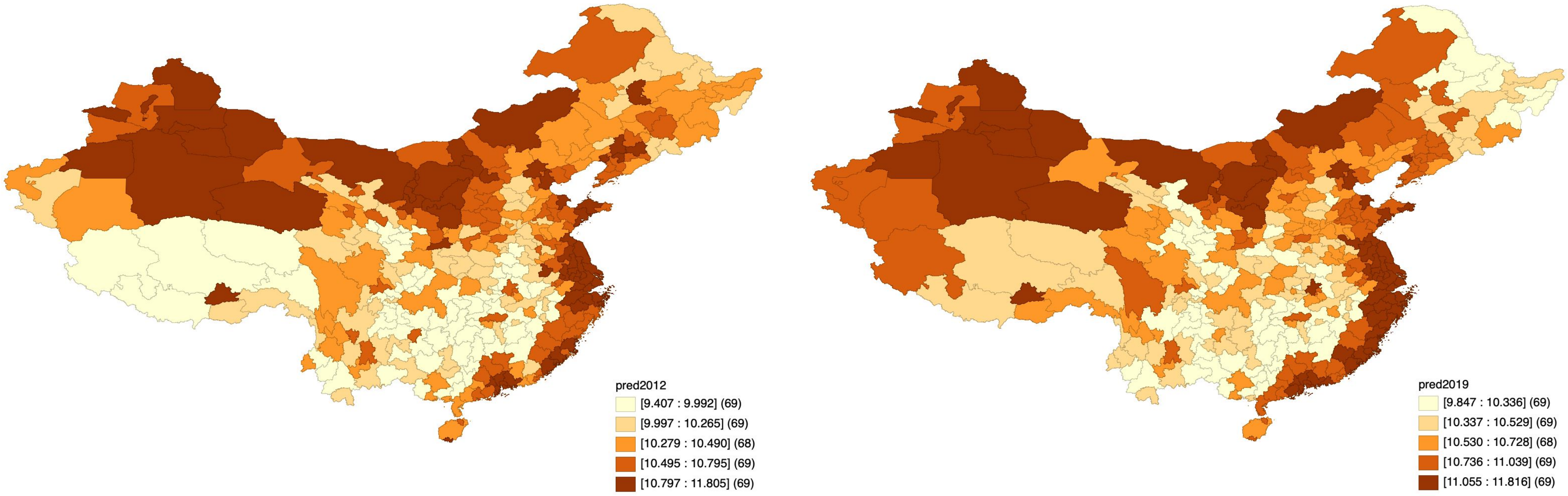


Fig.3 Compare of estimated GDP p.c. in 2012 and 2019

Step 2. Convergence analysis

Beta convergence: poorer economies' per capita incomes will grow faster than rich ones.

Simplified beta convergence model

$$g(y)_i = \alpha + \beta \cdot \log(y_{i0}) + \epsilon_i$$

Y: Average annual growth rate
X: GDP p.c. of initial year(in log form)

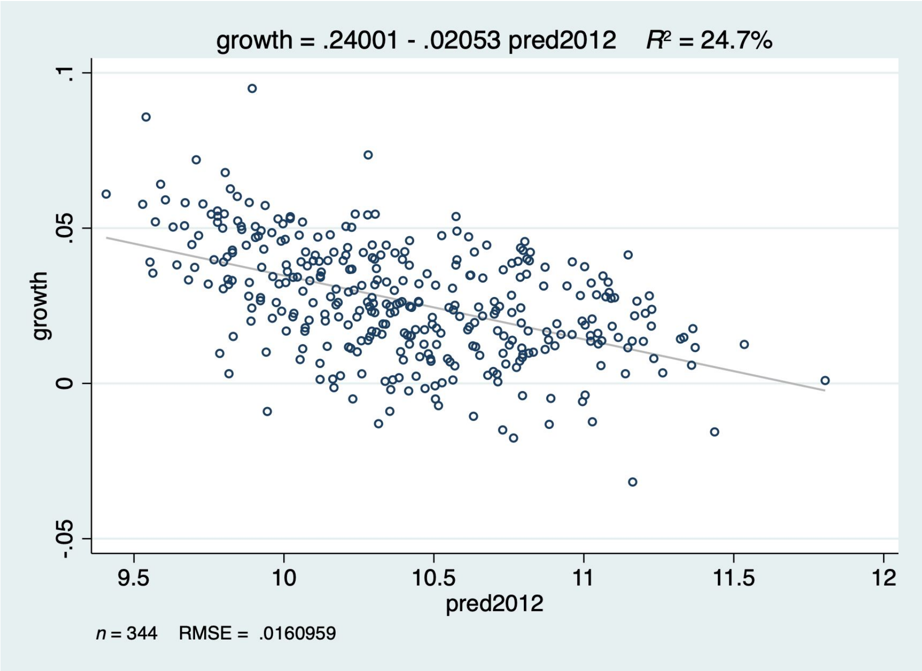


Fig.5 Scatter plot of growth rate and GDP p.c. of 2012(initial year)

Step 2. Convergence analysis with control variables

Conditional beta convergence

$$g(y)_i = \alpha + \beta \log(y_{i0}) + \underbrace{\sum_{j=1}^{k} \beta_j (x_{i0})_j}_{\text{control variables}} + \epsilon_i$$

Y: Average annual growth rate
X: GDP p.c. of initial year(in log form)

i: city, j: control variables

Control variables:
SK: physicial capital
SH: human capital
GOV: government expenditure
FDI: foreign inverstment
INDUSTRY: percentage of 2nd industry

Table 2. Beta convergence results with control variables

	Dependent variable: Growth rate (2012-2019)						
	Absolute conv.	SK ↑	LnSH ↓	GOV ↑	FDI ↑	INDUSTRY ↓	ALL ↑
β	-0.197^{***}	-0.225^{***}	-0.185^{***}	-0.205^{***}	-0.204^{***}	-0.188^{***}	-0.198^{***}
β *of SK*		-0.225^{***}					-0.223^{***}
β *of LnSH*			-0.130				-0.342^{**}
β *of GOV*				-0.135^{**}			-0.129^{*}
β *of FDI*					0.449		1.21^{**}
β *of INDUSTRY*						-0.019	-0.037^{*}
Speed of conv.	3.13%	3.64%	2.92%	3.27%	3.26%	2.98%	3.15%
Halflife time	22.11	19.04	23.72	21.15	21.27	23.30	21.99
R-Squared	0.255	0.330	0.258	0.268	0.259	0.257	0.357
Obs.	255	255	255	255	255	255	255

Step 2. Spatial convergence analysis

Non-spatial convergence model

$$g(y)_i = \alpha + \beta \cdot \log(y_{i0}) + \epsilon_i$$

Spatial lag model(SAR)

$$g(y)_i = \alpha + \beta \cdot \log(y_{i0}) + \rho W \cdot g(y)_i + \epsilon_i$$

Spatial error model(SEM)

$$g(y)_i = \alpha + \beta \cdot \log(y_{i0}) + \lambda W \epsilon_i + u_i$$

Spatial effects **accelerate** convergence process！

Table 3. Spatial dependence model results

	Dependent variable: Growth rate (2012-2019)		
	Non-spatial	Spatial lag	Spatial error
β	-0.191^{***}	-0.137^{***}	-0.266^{***}
α	2.267^{***}	1.517^{***}	3.040^{***}
Speed of convergence	3.03%	2.10%	4.42%
Half – life time (year)	22.89	32.93	15.69
ρ		0.649^{***}	
λ			0.769^{***}
R-Squared	0.202	0.521	0.617
AIC	-238.994	-377.282	-440.084
P-value of LM test			
LM test SAR	0.00000		
LM test SEM	0.00000		
Robust LM test SAR	0.00193		
Robust LM test SEM	0.00000		
Obs.	344	344	344

Step 3. Measuring spatial heterogeneity in convergence

There are two important features about China's regional economic convergence.

First is a time feature that China has experienced "Narrow-Expand-Narrow" period from the 1990s (Pan, 2010).

Second is the regional feature that <u>there is homogeneity within regions and heterogeneity between regions</u> (Zhu and Yu, 2014). It means there is convergence in either east, middle or west part of China, but the gap between different parts of China is enlarging.

Even though I have found absolute beta convergence, there may still be some difference across regions, since they have different initial status and economic environment. This part is to <u>explore the regional heterogeneity in beta convergence.</u>

Step 3. Measuring spatial heterogeneity in convergence

Geographically weighted regression(GWR model)

$$y = \beta_0(u_i, v_i) + \sum \beta_k(u_i, v_i)x_{ik} + \varepsilon_i$$

Where (u_i, v_i) the co-ordinate location of i, k is the number of explainary variables, ε_i is the error term.

<u>In another word, GWR model is doing 344 local regression for each city in China.</u>

In Fig.6, **b** is bandwidth, it decides the scale of doing local regression(the number of points including in local regression)

In my model, bandwidth is 46, <u>it implies high heterogeneity.</u>

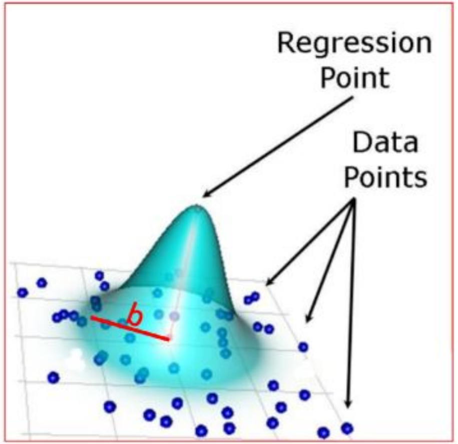


Fig.6 Illustration of local regression

Step 3. Mapping local convergence coefficient

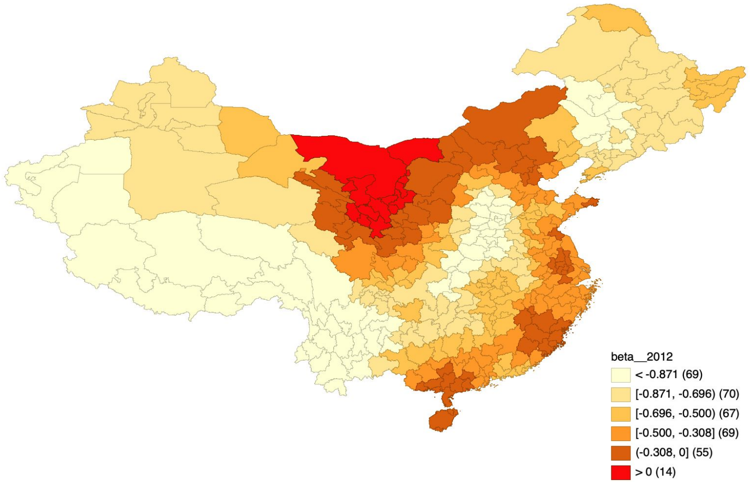


Fig.7 Local convergence **coefficient**

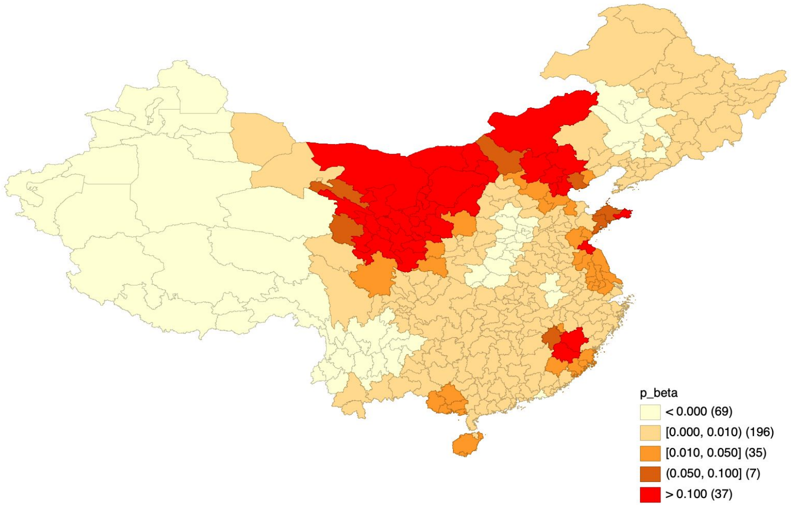


Fig.8 Local convergence **significance level**

Step 3. Local convergence half-life time

Half-life time represents the time that it would take for an economy to reduce by 50% the income gap between its initial state and its longrun equilibrium.

A higher convergence speed means this economy requires a shorter half-life time.

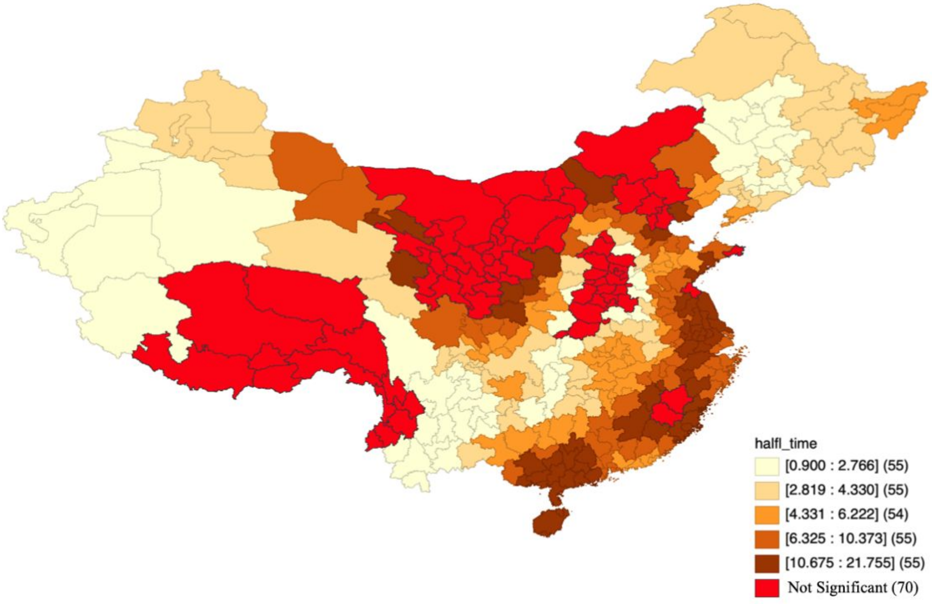


Fig.9 Local convergence half-life time

Concluding remarks

1. After filling the missing value, we have a clear overview of China's regional development. In China, Northwest, Southeast and Inner Mongolia Province are more developed areas.
2. China has absolute beta convergence, **physicial capital, government expenditure and FDI** promote convergence process. Spatial error model is more approriate in spatial beta convergence. Spatial effect accelarates convergence speed.
3. There is heterogenity in China's regional convergence process. Convergence speed is faster in west, southwest and northeast provicnes, lower in coastal provinces(southeast).

Thank you!

Acknowledgement

This work was financially supported by JST SPRING, Grant Number JPMJSP2125. The author (Initial) would like to take this opportunity to thank the “Interdisciplinary Frontier Next-Generation Researcher Program of the Tokai Higher Education and Research System.”