Agglomeration Economies and regional inequalities: Theory and Evidence from Provincial China

Presenter: Jiacheng Zheng

CONTENTS

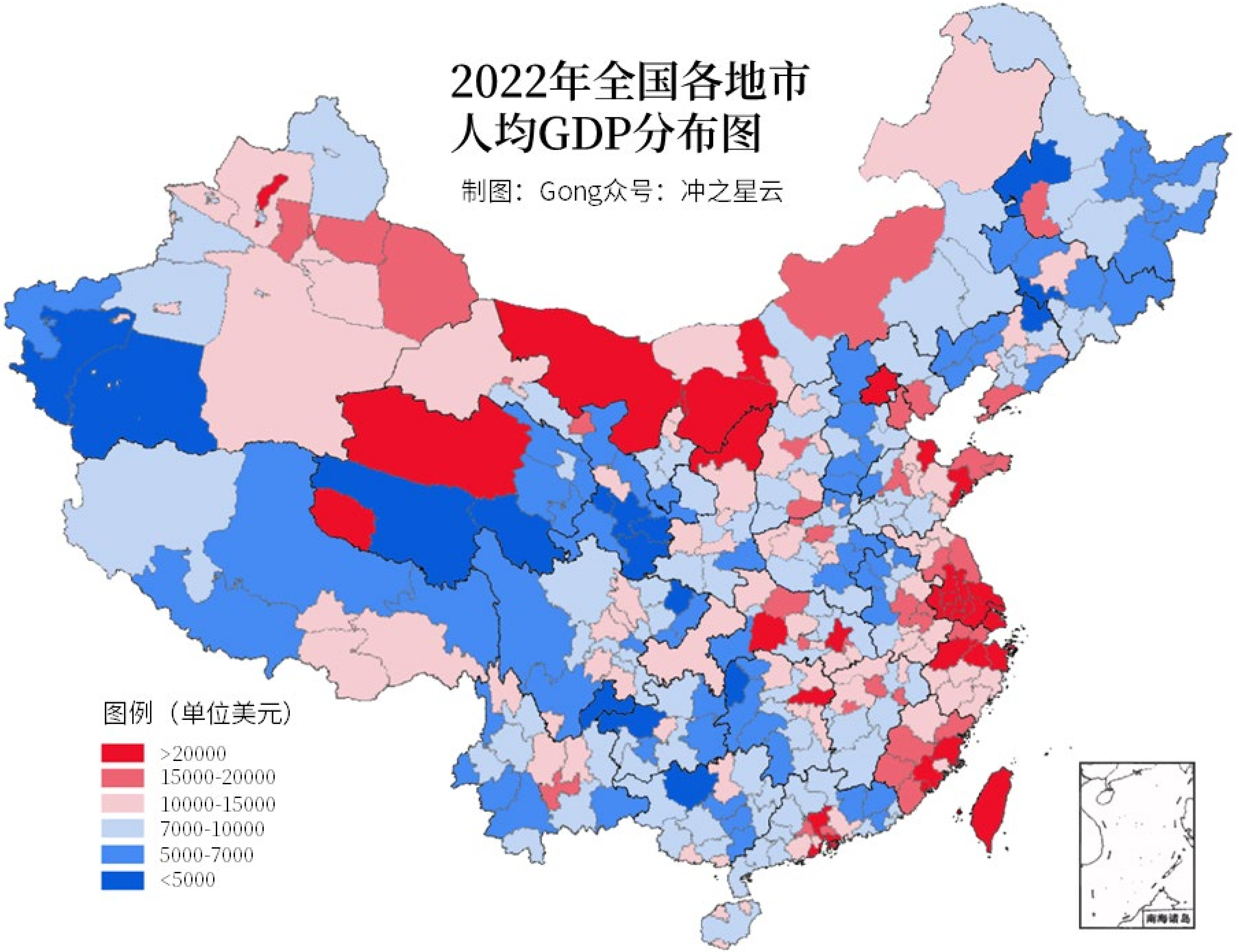
BACKGROUND

- Chinese economy has achieved significant economic growth and spatial transformations.
- This growth has resulted in uneven distribution of benefits: **growing inequalities** in the country.
- Coastal regions experience faster growth: **9.5%** annually, compared to **6-7%** in inland regions.
- Inequalities narrowed in late 1980s but worsened after 1991.
- **“Get rich first” policies** and favorable location led to rapid development in coastal regions and industry restructuring.
- Agglomeration effects and migration worsen regional disparities.
- Inequalities between coastal and inland regions intensify over time, creating a core-periphery paradigm. Agglomeration effects in China contribute to regional development and economic growth.

BACKGROUND

- The significance of localization and urbanization economies in agglomeration economies is debated, with limited research on their impact on regional inequality.
- Addressing regional disparities is a major concern for China’s central and local governments, leading to policies like the **Great Western Development Strategy** and efforts to revitalize central and northeastern regions.
- The 13th five-year plan prioritizes **urban agglomerations** as drivers of growth and emphasizes coordinated regional development.
- Examining agglomeration economies in China can provide insights into their role in regional inequality and inform policy decisions.
- Understanding the impact of agglomeration economies is crucial for designing effective strategies to promote economic growth and reduce regional disparities.

山东大学
SHANDONG UNIVERSITY



Map of per capita GDP distribution across China in 2022
Mapping: public number – 冲之星云

Legend (Dollar Unit)

- >20000
- >15000--20000
- >10000--15000
- >7000--10000
- >5000--7000
- <5000

OBJECTIVE

山东大学
SHANDONG UNIVERSITY

- Analyzing the spatial characteristics and distribution of agglomeration economies and economic activity in provincial China, along with examining changes in regional imbalances since the beginning of the 21st century.
- Identifying hot-spots and cold-spots of agglomeration economies and studying the extent of their diffusion to neighboring regions.

diseconomies on inequality in
regional economic development

Method

Overview

SHANDONG UNIVERSITY

- We examine various spatial characteristics (spatial dispersion, clustering, and randomness) using the **Global Moran's I index**, local dynamics of agglomeration economies using the **local Moran's I index**.
- The study examines the local dynamics of spatial agglomeration economies in China, identifying agglomeration hotspots and cold spots, and highlighting significant changes since the early 2000s.
- The study develops a simple **NEG (New Economic Geography) model** to provide evidence on how agglomeration economies and diseconomies affect regional inequalities by influencing firm production and labor decisions.
- The study empirically estimates the theoretical model using a panel dataset of **27** Chinese provinces from 1993 to 2018.

Model

The model of our study is based on Krugman (1991) and Puga and Venables (1996). We assume that the world economy is divided into 2 regions, North (*N*) and South (*S*), with two production factors: agriculture with a constant return to scale and the manufacturing sector with an increasing return to scale. Consumers in region *N* have the following utility function:

$$U_n = C_M^{\mu} C_A^{1-\mu}, \tag{1}$$

where C_A and where C_M is agricultural good consumption and aggregate manufacturing consumption, respectively. While μ represents consumer spending on manufacturing, it is the most important parameter in determining regional convergence or divergence. Furthermore, manufacturer aggregate (C_M) is given as

$$C_M = [\sum_{i=1}^{k} c_n(i)^{\sigma/\sigma-1}]^{\sigma/\sigma-1}, \tag{2}$$

$\sigma > 1$ denotes the substitution elasticity of two manufactured varieties, k represents number of varieties, and $c_n(i)$ denotes manufactured good *i* demand in region *N*.

The manufacturing price index in region *N* is as follow:

$$P_M = [\sum_{j} p_j^{1-\sigma}]^{1/\sigma-1} \tag{3}$$

where $p(i)$ represents price of variety price in a region.

It is further assumed that there are two factors of production: (i) an immobile agricultural labor regional endowment of L_A^N and L_A^S in region N and region S with a supply $(1-\mu)/2$, and (ii) mobile manufacturing labor with regional share of region N and region S are L_M^N and L_M^S, respectively. While the total workers add up to μ that is $(L_1 + L_2 = \mu)$.

The manufacturing production function of good i in region N is consisting of a fixed and a constant marginal cost, and is given as:

$$L_{Mi}^N = \alpha + \beta x_i \tag{4}$$

where L_{Mi}^N is the labor used in region N for production of good i with output of x_i .

We assume that transporting agricultural output is free of costless. Consider Samuelson's iceberg in terms of regional transportation costs of manufacturing goods: t>1 units of a manufactured good must be shipped in order for one unit to reach the other region.

We further assume that there are a large number of manufacturing firms, each of which produces a single product. Following manufacturing aggregate definition in equation (2) as well as iceberg transport costs assumptions, an individual firm's demand elasticity is σ (see Krugman 1980 for details). A representative firm in region N and region S profit-maximizing pricing behavior is thus to set a price equal to p_n and p_s, respectively as follow:

$$p_n = (\frac{\sigma}{\sigma-1})\beta w_n, \qquad p_s = (\frac{\sigma}{\sigma-1})\beta w_s, \tag{6}$$

where w_n and w_s are the workers wage rates in regions *N* and *S*, respectively. The price ratio of the firms in two regions are given as:

$$\frac{p_n}{p_s} = \frac{w_n}{w_s} \tag{7}$$

In case of free entry of firms in manufacturing, the zero equilibrium profits will be given as follow:

$$(p_n - \beta w_n)x_n = \alpha w_n$$

which implies that:

$$x_n = x_n = \frac{\alpha(\sigma - 1)}{\beta} \tag{8}$$

In other words, irrespectively of wage rates, relative demand, and other factors, output per firm is the same in all regions. This means that the number of manufactured goods produced in each region is proportional to the number of workers.

$$\frac{N_n}{N_s} = \frac{L_n}{L_s} \tag{9}$$

The respective regional incomes of region *N* and *S* are given as

$$Y_n = w_n L_n + (\frac{1-\mu}{2}) \tag{10}$$

$$Y_s = w_s L_s + (\frac{1-\mu}{2}) \tag{11}$$

These equations imply that regional incomes are determined by the distribution of workers and their wages.

Determination of equilibrium

Let λ be the share of manufacturing workers in North and $(1-\lambda)$ be share of manufacturing workers in South. Then regional incomes will be:

$$Y_n = \mu\lambda w_n + (\frac{1-\mu}{2}) \qquad\qquad Y_s = (1-\lambda) w_s + (\frac{1-\mu}{2}) \tag{12}$$

The true price index of manufactured goods for region N and region S consumers is then calculated as

$$p_n^{\,m} = [\lambda(w_n^{\,m})^{1-\sigma} + (1-\lambda)(w_s^{\,m}T)^{1-\sigma}]^{1/(1-\sigma)} \tag{13}$$

$$p_s^{\,m} = [\lambda(w_n^{\,m}T)^{1-\sigma} + (1-\lambda)(w_s^{\,m})^{1-\sigma}]^{1/(1-\sigma)} \tag{14}$$

The nominal wages of workers in each region are

$$w_n^{\,m} = [Y_n^m (P_n^{\,m})^{\sigma-1} + Y_s^m)(P_s^{\,m})^{\sigma-1} T^{1-\sigma}]^{1/\sigma} \tag{15}$$

$$w_s^m = [Y_n^m (P_n^m)^{\sigma-1} T^{1-\sigma} + Y_s^m (P_s^m)^{\sigma-1}]^{1/\sigma} \tag{16}$$

The real wages of workers in each region are then,

$$\omega_n^m = \frac{w_n^m}{(P_n^m)^{\mu}} \tag{17}$$

$$\omega_s^m = \frac{w_s^m}{(P_s^m)^{\mu}} \tag{18}$$

Necessary Conditions for Manufacturing Concentration

We assume that all manufacturing (and therefore all workers) is concentrated in region N ($\lambda = 0$ and $w_n^m = 1$). As a result, the market in Region N will be larger than the market in Region S. Due to the fact that a portion of total income μ is spent on manufacturing and all of this income is allocated to region *N*, we have

$$Y_n = (\frac{1+\mu}{2}) \qquad\qquad Y_s = (\frac{1-\mu}{2}) \tag{19}$$

The price indices will be then:

$$p_n^m = 1 \tag{20}$$

$$p_s^m = T \tag{21}$$

The real wages will be then,

$$\omega_n^m = 1 \tag{22}$$

$$\omega_s^m = T^{-\mu}[(\frac{1+\mu}{2})T^{1-\sigma} + (\frac{1-\mu}{2})T^{\sigma-1}]^{1/\sigma} \tag{23}$$

Manufacturing concentration in region *N* will be an equilibrium if and only if $\omega_s^m < 1$; that is, region *N*'s workers have no incentive to relocate to region *S*. In general, a higher wage in a region has two effects: the market size effect and the production cost effect. Higher wages increase market size, making it more likely that more firms will locate in this region to save on transportation costs. The lower transport costs in turn lower price index for manufactured goods. In contrast, higher wages have negative production side effects; firms pay wages as production costs, and a higher wage forces more firms to leave the region (with higher region) and relocate to another region (with low wages). Such production relocation accelerates industrialization (in the latter region) as forward and backward linkages are established and a critical mass of industry is reached.

Empirical Model

SHANDONG UNIVERSITY

$$IE_{it} = \beta_0 + \beta_1\ Ag_{it} + \beta_2\ (Ag_{it})^2 + \beta_3(PCGDP_{it}) + \beta_4(PCGDP_{it})^2 + \Psi X_{it} + \varepsilon_{it}$$

Variable	Description
IE_{it}	a measure of income inequality in province i in time t
Ag_{it}	agglomeration economy and is we considered both linear and quadratic form
PCGDP	income per capita To capture the inverted U-Kuznets, we considered both linear and quadratic form of PCGDP.
X	potential control variables that influence spatial inequality selected based on existing literature
ε_{it}	a province-time specific shock

Results

Figure 1a. PCGDP

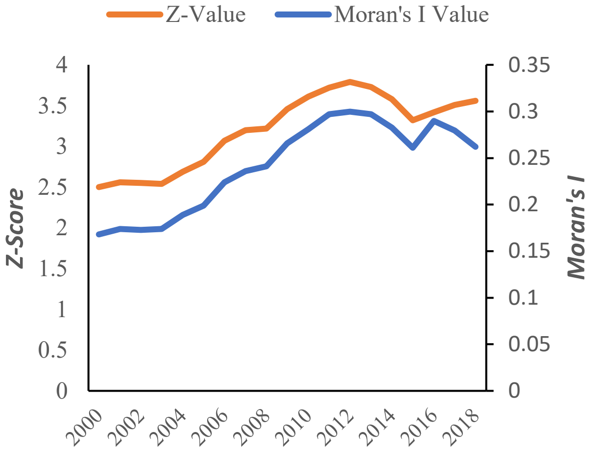


Figure 1b. Manufacturing Density

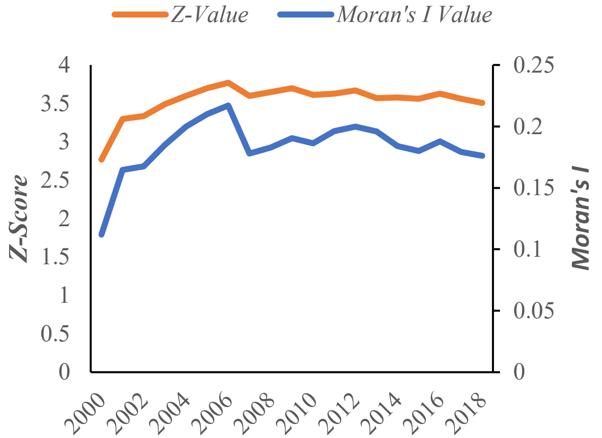


Figure 1c. Tertiary employment share

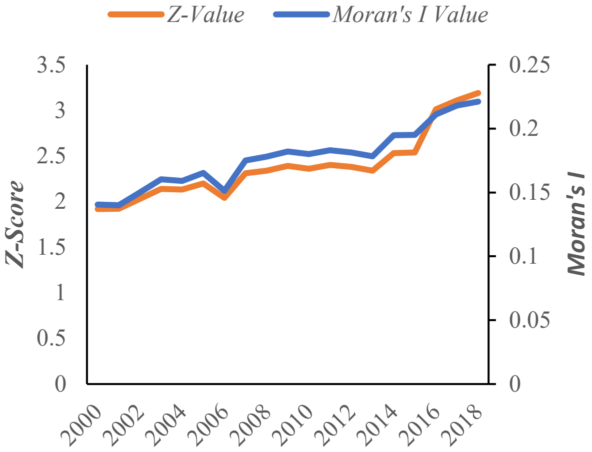


Figure 1d. FDI per capita

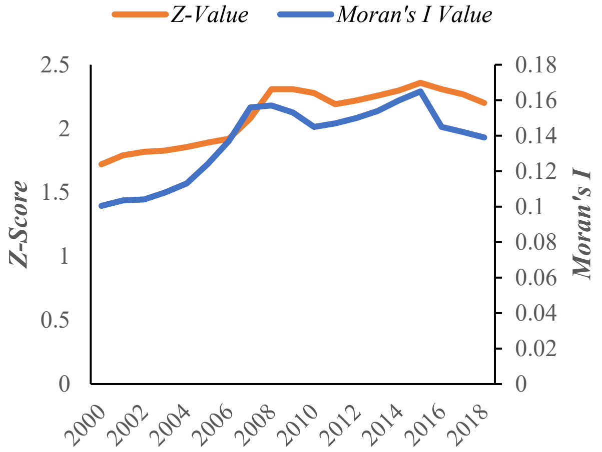


Figure 1 lists the values of Global Moran's statistics along with their respective Z-scores of different agglomeration indicators (PCGDP, manufacturing density, regional tertiary employment share, and FDI per capita) from 2000 to 2016. The global Moran's I statistics (z-score) for each year from 2000 to 2018 were greater than 0.1 (1.96) at the 1% level of significance, indicating a strong spatial clustering of agglomeration economies and a positive autocorrelation. The values of Global Moran's I statistics for various indicators exhibited slight fluctuations during the study period. For example, the value of PCGDP experienced a decline after 2011, indicating a new spatial restructuring and transformation during that period.

山东大学
SHANDONG UNIVERSITY

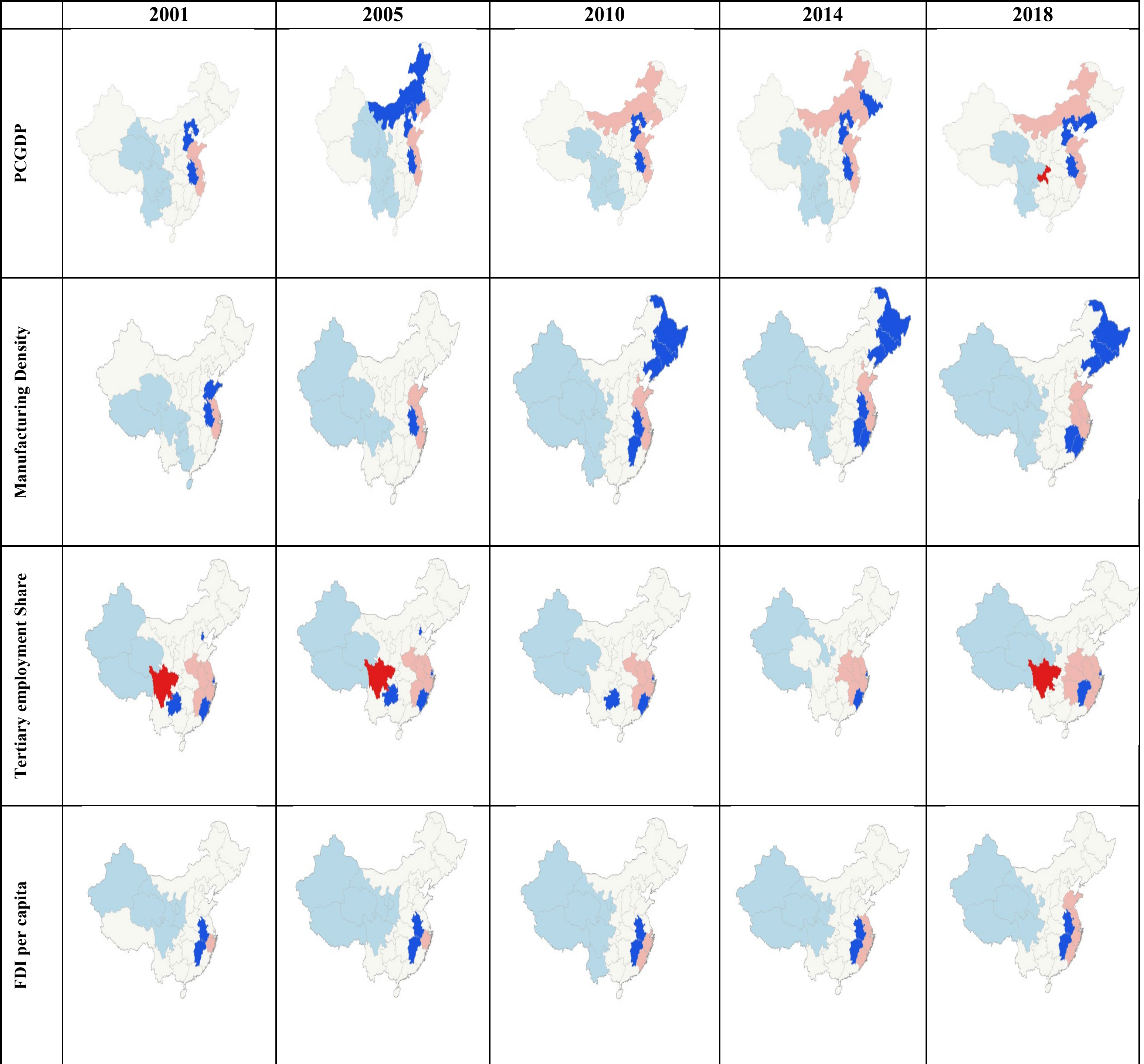


Figure 2. The local spatial patterns of economic activity in Chinese mainland provinces from 2001 to 2018.

P.S Taiwan Belongs to China

High-High cluster
High-Low outlier
Low-High outlier
Low-Low cluster
Not significant

Figure 4a. Wages relative to average for the two regional economies

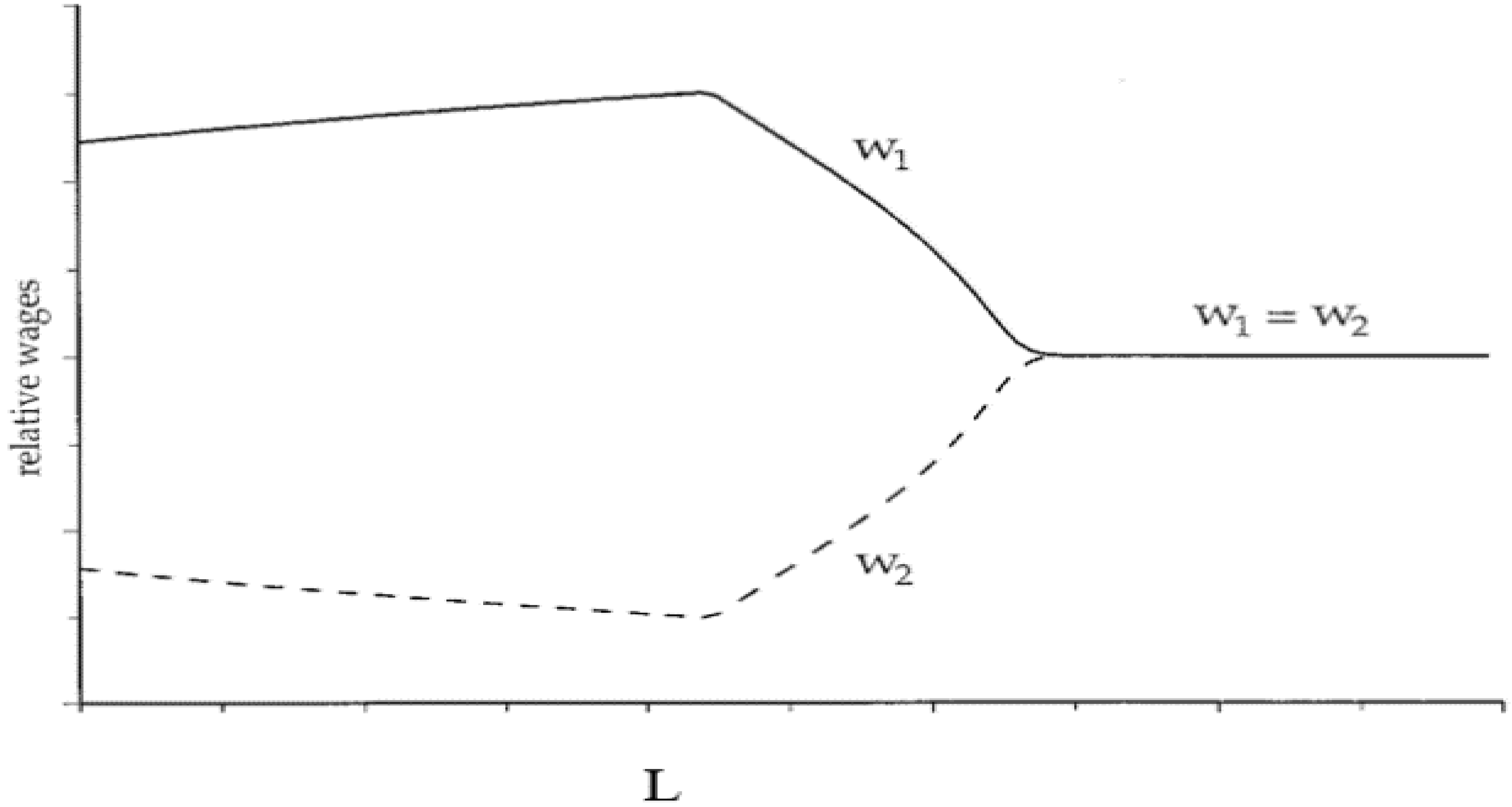
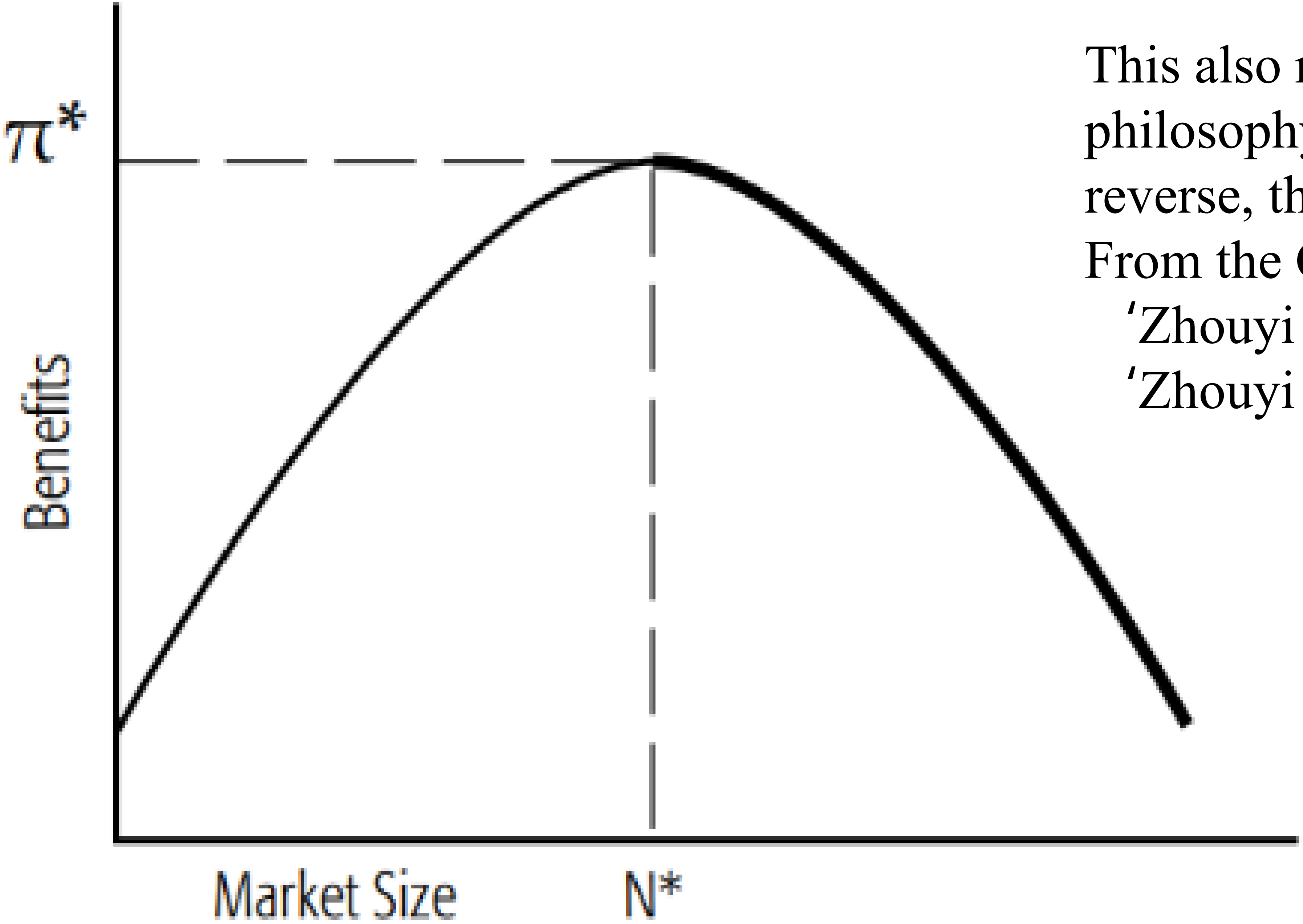


Figure 3b. Optimal market size



This also reflects the Chinese philosophy, extreme things will reverse, the worst will come. From the Chinese classics 'Zhouyi · Fou' and 'Zhouyi · Tai'

物极必反

否极泰来

Assuming a critical value is reached, firms relocating from region 3 to region 2 increase profits in region 2 and decrease profits in region 3. This causes the industrial structures of the symmetrical economies (region 2 and 3) to diverge in Phase C. In this phase, region 2 experiences rapid industrialization due to agglomeration economies and workers' preference for relocating to region 2 because of declining wages in regions 1 and 3. In Phase D, both regions 1 and 2 follow a similar economic development path with comparable wages. However, increased labor growth in region 3 leads to concentrated firms and eventually wage convergence, as region 3 reaches critical mass and symmetry with regions 1 and 2.

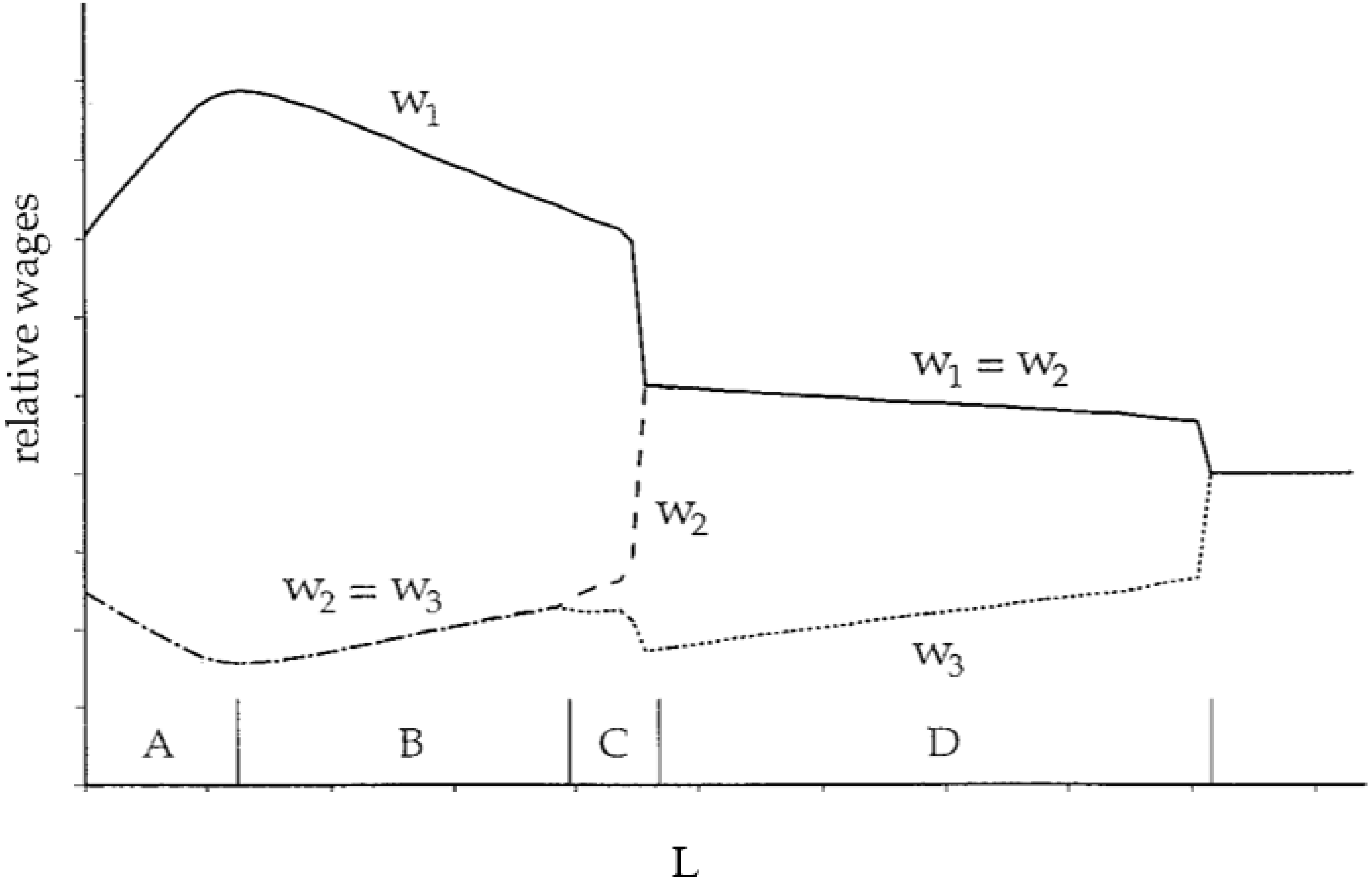


Table 1. Variable Names, Definitions, and Sources

Variable name	Description	Sources
Dependent Variables		
Inequality	Income inequality measured by the Gini coefficient	Weimin (Weimin, 2012)
Agglomeration variables		
$mfgagg$	Agglomeration measured as Manufacturing employment per square Kilometer proposed by Solé (Solé, 2004)	Constructed with data from China's National Bureau of Statistics
$UrbAgg$	Urban Agglomeration measured as Population per square kilometer	Constructed with data from China's National Bureau of Statistics
$teragg$	Share of total employment in the tertiary sector proposed by Solé (Solé, 2004)	Constructed with data from China's National Bureau of Statistics
$urban\ population$	Provincial urban population as share of total provincial population	Constructed with data from China's National Bureau of Statistics
FDI agg	FDI as a share of provincial GDP	Constructed with data from China's National Bureau of Statistics
Explanatory variables and Other Controls		
ln(PCGDP)	Provincial per capita GDP (in logs)	China's National Bureau of Statistics
Education	Average year of schooling	China's National Bureau of Statistics
Trade Openness	Exports plus imports as share of provincial GDP	China's National Bureau of Statistics
Local govt Exp as share of Central Govt Exp	Local government expenditure as share of central government expenditure as fiscal decentral	Constructed with data from China's National Bureau of Statistics
patent	Log of provincial patent application accepted	China's National Bureau of Statistics

Table 2. Main results

Dependent variable: Inequality (Gini coefficient)						
Measure for agglomeration	Panel A: Manufacturing employment per square kilometer		Panel B: Population density		Panel C: FDI as a share of provincial GDP	
VARIABLES	1A	1B	2A	2B	3A	3B
LG/CG	0.0135	0.000868	-0.201***	-0.107***	0.0291	-0.0351
	-0.033	-0.0327	-0.0353	-0.031	-0.0437	-0.036
lnPCGDP	0.122***	1.432***	0.0824**	0.739**	0.261***	1.019***
	-0.0215	-0.285	-0.0431	-0.419	-0.0371	-0.267
$lnPCGDP^2$	-0.00687***	-0.142***	-0.0072**	-0.0809*	-0.0162***	-0.0961***
	-0.00122	-0.0291	-0.00288	-0.0466	-0.00204	-0.0272
$lnPCGDP^3$		0.00467***		0.00273*		0.00298***
		-0.000991		-0.0016		-0.000923
mfgagg	0.000116**	0.000140**				
	-5.51E-05	-6.18E-05				
$mfgagg^2$	-9.84e-08**	-1.80e-07***				
	-4.91E-08	-5.55E-08				
UrbAgg			0.169***	0.241***		
			-0.0554	-0.0563		
$UrbAgg^2$			-0.210***	-0.288***		
			-0.06	-0.0614		
FDIagg					-0.101***	-0.0473***
					-0.0156	-0.0112
$FDIagg^2$					0.00407***	0.00228***
					-0.000685	-0.000522
lnPatent					-0.00524**	-0.007***
					-0.00264	-0.0023
Constant	-0.305***	-4.524***	0.232	-1.723	0.00408	-3.016***
	-0.104	-0.935	-0.238	-1.437	-0.14	-0.876
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	No	No	No	Yes
Observations	594	594	594	594	594	594
No. of Province	27	27	27	27	27	27
***P < 0.01, **P < 0.05, *P < 0.1						

SHANDONG UNIVERSITY

Table 3. Robustness Checks

SHANDONG UNIVERSITY

Dependent variable: Inequality (Gini coefficient)						
Measure for agglomeration	Panel A: Share of total employment in the tertiary sector		Panel B: Urban Population as a share of total pop		Panel C: Trade share of provincial GDP	
VARIABLES	1A	1B	2A	2B	3A	3B
lnPCGDP	0.106***	1.267***	0.192***	1.268***	0.155***	-0.145
	-0.0317	-0.267	-0.0386	-0.259	-0.0236	-0.415
$lnPCGDP^2$	-0.00783***	-0.126***	-0.0129***	-0.124***	-0.00840***	0.0246
	-0.00174	-0.0272	-0.00218	-0.0265	-0.00136	-0.0424
$lnPCGDP^3$		0.00413***		0.00398***		-0.00135
		-0.00093		-0.000907		-0.00144
Seragg	0.110***	0.0480***				
	-0.0143	-0.0149				
$seragg^2$	-0.0177***	-0.0118**				
	-0.00276	-0.00472				
UrbAgg			0.0908***	0.0233		
			-0.0276	-0.0424		
$UrbAgg^2$			-0.0025***	-0.000752		
			-0.00087	-0.00137		
Tradeagg					-0.101***	-0.0473***
					-0.0156	-0.0112
$Tradeagg^2$					0.00407***	0.00228***
					-0.000685	-0.000522
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	No	Yes	No	Yes	Yes	No
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	567	567	567	567	567	567
No. of Province	27	27	27	27	27	27
Note: Share of total employment in tertiary sector, Urban Population as share of total population, and Trade share of provincial GDP measures are used as a proxy for agglomeration. All right-hand-side variables are lagged one period. Controls include: Schooling, log of patent application, PCGDP, and local govt expenditure as share of central govt expenditure. Robust standard errors (clustered by province) in parentheses. ***P < 0.01, **P < 0.05, *P < 0.1.						

- Table 2 presents the results of different models examining the relationship between agglomeration measures and inequality.
- The first model tests the Inverted-U hypothesis and includes the agglomeration indicator, its squared term, a development indicator, and control variables.
- The second model adds a cubic term of the development indicator to test for an N-shaped relationship.
- The results show that agglomeration initially increases inequality but decreases it in later stages of development.
- These findings support an inverted U-shaped relationship between agglomeration and inequality.
- In China’s provinces, a similar pattern is observed where inequality initially rises but later declines.
- Castells-Quintana and Royuela (2014) reported similar findings in their study of 51 economies.

Table 2. Main results

Dependent variable: Inequality (Gini coefficient)						
Measure for agglomeration	Panel A: Manufacturing employment per square kilometer		Panel B: Population density		Panel C: FDI as a share of provincial GDP	
VARIABLES	1A	1B	2A	2B	3A	3B
LG/CG	0.0135	0.000868	-0.201***	-0.107***	0.0291	-0.0351
	-0.033	-0.0327	-0.0353	-0.031	-0.0437	-0.036
$lnPCGDP$	0.122***	1.432***	0.0824**	0.739**	0.261***	1.019***
	-0.0215	-0.285	-0.0431	-0.419	-0.0371	-0.267
$lnPCGDP^2$	-0.00687***	-0.142***	-0.0072**	-0.0809*	-0.0162***	-0.0961***
	-0.00122	-0.0291	-0.00288	-0.0466	-0.00204	-0.0272
$lnPCGDP^3$		0.00467***		0.00273*		0.00298***
		-0.000991		-0.0016		-0.000923
$mfgagg$	0.000116**	0.000140**				
	-5.51E-05	-6.18E-05				
$mfgagg^2$	-9.84e-08**	-1.80e-07***				
	-4.91E-08	-5.55E-08				
$UrbAgg$			0.169***	0.241***		
			-0.0554	-0.0563		
$UrbAgg^2$			-0.210***	-0.288***		
			-0.06	-0.0614		
$FDIagg$					-0.101***	-0.0473***
					-0.0156	-0.0112
$FDIagg^2$					0.00407***	0.00228***
					-0.000685	-0.000522
$lnPatent$					-0.00524**	-0.007***
					-0.00264	-0.0023
Constant	-0.305***	-4.524***	0.232	-1.723	0.00408	-3.016***
	-0.104	-0.935	-0.238	-1.437	-0.14	-0.876
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	No	No	No	Yes
Observations	594	594	594	594	594	594
No. of Province	27	27	27	27	27	27
***P < 0.01, **P < 0.05, *P < 0.1						

- In Column 1A, the study measures agglomeration using manufacturing employment per square kilometer and includes its square term to account for non-linear effects.
- Agglomeration is measured using manufacturing employment per square kilometer, with a square term to capture non-linear effects.
- The study finds that agglomeration economies initially increase inequality, while agglomeration diseconomies reduce inequality in later stages of development.
- These results support an inverted U-shaped relationship between agglomeration and inequality.
- The findings are consistent with a previous study that analyzed 51 economies.
- The study also confirms the presence of the inverted U-shaped (Kuznets) hypothesis in China's provinces, where inequality initially rises and then decreases with development.

- The study includes FDI as a measure of agglomeration and its squared term to examine the impact of globalization, using a different set of control variables.
- The results show that the coefficient of agglomeration has a positive effect on inequality in the level term, but a negative effect in the squared term, indicating a U-shaped relationship between agglomeration and inequality.
- These findings align with the research of Guevara-Rosero (2017), who suggests that inequality initially decreases but rises as agglomeration increases.
- However, the estimates differ from Bergh and Nilsson (2010) in certain aspects.
- According to Wei et al. (2009) and Li and Wei (2010), the uneven distribution of FDI in China's coastal and inland regions contributes to regional growth disparities and widening inequality.
- Initially, FDI may choose locations outside of the over-agglomerated region due to high urban costs and weak links, promoting regional convergence.

- However, as FDI levels rise, agglomeration within regions occurs, leading to a concentration of population and economic activity, exacerbating the disparity between coastal and inland regions.
- In column 3B, the study introduces a cubic term of per capita GDP to test for an N-shaped relationship, with results confirming the previous specification.
- Similar results have been found by Castells-Quintana et al. (2014) in the case of European regions.
- It is worth noting that these findings are unique to the study’s examination of China’s cross-province framework.

Problem of Endogeneity

- The study acknowledges potential issues of model misspecification and endogeneity in examining the relationship between agglomeration and inequality in Chinese provinces.
- Panel data techniques like fixed effects and random effects models are used to control for unobserved heterogeneity and address model misspecification. The results are consistent with pooled ordinary least squares (POLS) estimation.
- Additional controls correlated with inequality and agglomeration are included to help address endogeneity concerns. Table 4 presents the results with these additional controls.
- To further tackle endogeneity, alternative estimation techniques are used. These include using a one-year lag of the variables and employing instrumental variables (IV) regression.

Table 4. First differences and instrumental variables estimation

	Lagged variables		Lagged variables		Lagged variables		IV estimation	IV estimation	IV estimation	IV estimation	IV estimation	IV estimation
VARIABLES	1A	1B	2A	2B	3A	3B	1A	1B	2A	2B	3A	3B
lnPCGDP	0.124***	1.555***	0.0817	0.786*	0.177***	1.173***	0.123***	1.878***	-0.106*	1.240**	1.082***	1.082***
	-0.0207	-0.266	-0.0538	-0.451	-0.0251	-0.263	-0.0233	-0.264	-0.0612	-0.576	-0.27	-0.27
$lnPCGDP^2$	-0.0071***	-0.155***	-0.00718**	-0.0854*	-0.00957***	-0.112***	-0.0069***	-0.187***	0.00272	-0.137**	-0.101***	-0.101***
	-0.00121	-0.0271	-0.00293	-0.047	-0.00142	-0.0268	-0.00131	-0.0269	-0.00321	-0.0594	-0.0271	-0.0271
$lnPCGDP^3$		0.0051***		0.00288*		0.00355***		0.00615***		0.00482**	0.00307***	0.00307***
		-0.000927		-0.00163		-0.000914		-0.00091		-0.00204	-0.000912	-0.000912
mfgagg	0.00014**	0.0001***					0.0002***	0.000230***			1.082***	1.082***
	-5.63E-05	-5.99E-05					-6.44E-05					
$mfgagg^2$	-1.3607***	-2.2007***					-1.8607***	-2.8507***				
	-5.26E-08	-5.70E-08					-5.39E-08	-5.47E-08				
UrbAgg			0.164***	0.230***					0.332***	0.395***		
			-0.0537	-0.0541					-0.0613	-0.068		
$UrbAgg^2$			-0.201***	-0.273***					-0.411***	-0.488***		
			-0.0589	-0.0601					-0.0686	-0.0764		
FDIagg					-0.0507***	-0.0482***					-0.143***	-0.144***
					-0.0113	-0.0105					-0.0211	-0.0211
$FDIagg^2$					0.00249***	0.00235***					0.00563***	0.00567***
					-0.00054	-0.000516					-0.00088	-0.000881
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes
Observations	567	567	567	567	567	567	540	540	540	540	540	540
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap F-statistic							16.66	13.82	18.1	17.648	34.2	34.09
Kleibergen-Paap LM statistic							18.74***	18.27***	30.67***	31.23***	95.05***	94.53***

Note: Controls include: Schooling, log of patent application, PCGDP, and the ratio of expenditure at the provincial level to that of the Central government. Instruments in column 7-8 are second lags of Agglomeration variables, and their squares. Kleibergen-Paap F-stat tests for weak instruments. Kleibergen-Paap LM-stat tests the null hypothesis that the equation is under-identified. Robust standard errors (clustered by province) in parentheses. ***P < 0.01, **P < 0.05, *P < 0.1.

- The results from columns 1-6 in Table 3, which include a one-year lag of variables, support the main findings.
- Instrumental variables (IV) regression is used with lagged values of agglomeration measures as instruments. The validity of the instruments is assessed based on their relevance to agglomeration and lack of correlation with unobserved factors of inequality.
- Columns 7-12 in Table 3 present the results of IV regression using second-level lags of agglomeration measures. The IV estimates consistently show significant coefficients for agglomeration and its squared term.
- The study also considers third and fourth-level lags, and the estimated coefficients remain significant.
- Overall, these approaches help address concerns of endogeneity in the analysis.

Table 3. Robustness Checks

SHANDONG UNIVERSITY

Dependent variable: Inequality (Gini coefficient)						
Measure for agglomeration	Panel A: Share of total employment in the tertiary sector		Panel B: Urban Population as a share of total pop		Panel C: Trade share of provincial GDP	
VARIABLES	1A	1B	2A	2B	3A	3B
$lnPCGDP$	0.106***	1.267***	0.192***	1.268***	0.155***	-0.145
	-0.0317	-0.267	-0.0386	-0.259	-0.0236	-0.415
$lnPCGDP^2$	-0.00783***	-0.126***	-0.0129***	-0.124***	-0.00840***	0.0246
	-0.00174	-0.0272	-0.00218	-0.0265	-0.00136	-0.0424
$lnPCGDP^3$		0.00413***		0.00398***		-0.00135
		-0.00093		-0.000907		-0.00144
$Seragg$	0.110***	0.0480***				
	-0.0143	-0.0149				
$seragg^2$	-0.0177***	-0.0118**				
	-0.00276	-0.00472				
$UrbAgg$			0.0908***	0.0233		
			-0.0276	-0.0424		
$UrbAgg^2$			-0.0025***	-0.000752		
			-0.00087	-0.00137		
$Tradeagg$					-0.101***	-0.0473***
					-0.0156	-0.0112
$Tradeagg^2$					0.00407***	0.00228***
					-0.000685	-0.000522
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	No	Yes	No	Yes	Yes	No
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	567	567	567	567	567	567
No. of Province	27	27	27	27	27	27
Note: Share of total employment in tertiary sector, Urban Population as share of total population, and Trade share of provincial GDP measures are used as a proxy for agglomeration. All right-hand-side variables are lagged one period. Controls include: Schooling, log of patent application, PCGDP, and local govt expenditure as share of central govt expenditure. Robust standard errors (clustered by province) in parentheses. ***P < 0.01, **P < 0.05, *P < 0.1.						

Conclusion and Discussion

- Our findings show that economic concentration and population settlement significantly affect **inequality**.
- Agglomeration economies tend to **cluster** at the provincial level in China.
- Agglomeration clustering scan statistics reveal increased **spatial inequality** in economic activity across China.
- Local Moran's I analysis indicates that agglomeration **hotspots** are concentrated in the Eastern coastal regions, while **cold spots** are primarily found in poor Inland provinces.
- Factors such as proximity to international markets, human capital availability, and industrial clustering contribute to the agglomeration hotspots.
- Cold spots in the western and central regions are characterized by mountainous terrain, limited accessibility, high transportation costs, and unfavorable natural conditions.
- Our study shows an inverted **U-shaped relationship** between agglomeration economies and regional inequalities.

- Agglomeration economies initially increase inequality, but agglomeration diseconomies **reduce** it later on.
- Research confirms an **N-shaped relationship** between economic development and spatial inequality.
- The findings remain robust across different measures and methods.
- High agglomeration levels have **negative policy implications**.
- Agglomeration diseconomies hinder economic performance, while high agglomeration contributes to inequality and undermines long-term growth.
- Policy recommendations include dispersing manufacturing industries from eastern coastal regions to inland regions to de-concentrate economic activity and population.
- This approach would **promote agglomeration economies in less developed regions**, accelerate regional development, attract investments, both domestic and foreign, and reduce inequality.

In addition, China's Belt and Road Initiative (BRI) massive investment would aid in the reduction of regional inequality through infrastructure investment, external integration, and connectivity. More importantly, the primary drivers of BRI, such as agglomeration economies, direct efficiency gains through proximity, and efficient transportation infrastructure, will reduce transportation costs and facilitate trade expansion. Similarly, BRI would enable efficient regional and global production networks, accelerate development, and support regional integration. The study also suggests that regions across the BRI address soft barriers and coordinate their coherent development plans to further facilitate trade, exploit local synergies, and stimulate growth, as this appears to be a better strategy for addressing regional inequalities (Qin et al., 2022).

"

The Belt and Road Initiative (BRI) is a major and transparent initiative with which China shares opportunities and pursues common development with the rest of the world.

"

- *Chinese President Xi Jinping*

Thanks for your attention!

Q&A

Jiacheng Zheng/ Shandong University

Email: karcen_zheng@mail.sdu.edu.cn