Environment, Development and Sustainability
https://doi.org/10.1007/s10668-023-03242-9

REVIEW

The urban population agglomeration capacity and its impact on economic efficiency in the Yangtze River Delta Urban Agglomeration

Jinping Lin[1,2] · Shan Yang[1,2] · Yanghui Liu[1,2] · Yiheng Zhu[1,2] · Anning Cai[3]

Received: 12 August 2022 / Accepted: 4 April 2023

Abstract
This paper constructs and elaborates a theoretical model of urban economic efficiency (UEE) from the perspective of urban scaling law. A framework of urban economic analysis is established with urban population agglomeration capacity (UPAC) as the explanatory factor. Taking the Yangtze River Delta Urban Agglomeration (YRDUA) as a case study, explore the influence of UPAC on UEE. The results show that the gap between the UEE in the YRDUA gradually decreases, the spatial agglomeration characteristic weakens, and the UEE among cities leads to a balanced tendency. However, the spatial agglomeration pattern of UPAC becomes more and more significant. (Mega/super) large cities are mostly advanced types, while small cities are lagging types. The influence of UPAC on UEE gradually decreases and diverges from significant positive influence to insignificant influence in advanced cities and significant positive influence in lagging cities. The framework of UEE research provides a more objective way to understand and compare the economic performance of cities of different scales. The empirical study findings provide a basis for decision-making on developing different types of cities.

Keywords Urban economic efficiency (UEE) · Urban population agglomeration capacity (UPAC) · Urban scaling law · Rank size · Yangtze River Delta Urban Agglomeration (YRDUA)

✉ Shan Yang
yangshan@njnu.edu.cn

1 School of Geographic Science, Nanjing Normal University, No.1, Wenyuan Road, Qixia, Nanjing 210023, China

2 Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China

3 School of Tourism and Administration, Nanjing XiaoZhuang University, Nanjing 211171, China

1 Introduction

The causes of urban economic growth and the spatial patterns of population mobility are among the most complex enigmas in contemporary social science (Storper & Scott, 2009). They are prompting extensive discussion by scholars in China and abroad. For example, the question of whether people move with jobs or jobs move with people emerged in the early literature (Mazek & Chang, 1972), as well as Weber's view of urban growth as a result of industrialization and its accompanying local economic development (Dyos & Warner, 1966; Edgeworth, 1899). Furthermore, later, two influential paradigms—the "growth machine" model (Molotch & Logan, 2014) and the "urban regime" analysis (Stone, 1993; Tretter, 2008)—have conceptually and methodologically attempted to emphasize economic growth as the primary goal of urban politics (Lin, 2002). However, in recent years, many studies have argued that urban (or regional) growth is essentially the result of population flow and spatial agglomeration (Rozenfeld et al., 2011; Wei et al., 2020, Zheng & Du, 2020), which has risen to a prominent position. The argument common to these studies is that people always migrate selectively to cities with good quality attributes, large cities, or economically developed cities (Clark et al., 2002), especially those with convenient services and good facilities (Chung et al., 2020; Glaeser, PB-2005–1). So, population concentration is uneven, and concentration in large cities has become a worldwide social development (Guo et al., 2021; Scott, 2012). Places with high population concentrations always experience rapid development and urban economic growth due to people's creative and innovative energy (Wang et al., 2020).

The theoretical discussion and empirical evidence on urban agglomeration for economic growth began with the interest in regional aggregation in the "new economic geography" of Krugman and others (Fujita & Krugman, 2004; Krugman, 1990). Despite the findings of many Western studies, population agglomeration promotes regional economic growth (Faberman & Freedman, 2016; Martin & Ottaviano, 2001). There is insufficient evidence that population agglomeration promotes cities' economic growth (Brinkman, 2016; Castells-Quintana & Royuela, 2014; Monkkonen et al., 2018). Population agglomeration fosters innovation and technological progress, driving economic growth and urban prosperity while simultaneously creating problems such as congestion, poverty, crime, and disease that hinder urban development (Parnreiter, 2021; Rodriguez-Pose & Storper, 2020; Scott, 2012). In a sense, the former benefits cities, and the latter is the cost of population agglomeration (Glaeser, 1998, 2011), which is characteristic of the Williamson hypothesis (Bruelhart & Sbergami, 2009; Williamson, 1965). Empirical studies in Chinese cities also confirm that population agglomeration has a heterogeneous effect on economic growth (Chen & Partridge, 2013; Zhang, Wei, & Zhang, 2021). These conclusions reveal that population agglomeration only sometimes produces balanced growth paths and convergent growth rates. Instead, uneven, differential growth rates may be a recurring feature of urban development (Cheshire & Malecki, 2004; Czamanski & Broitman, 2018).

In measuring the spatial agglomeration of the urban population, many empirical studies have used indicators such as urban population size (Henderson, 2000), employment, and sectoral density (Bruelhart & Mathys, 2008; Ciccone, 2002; Ciccone & Hall, 1996), urbanization (Henderson, 2003), and industrial agglomeration (Duranton & Puga, 2000; He & Pan, 2010). These indicators reflect the results and extent of urban population agglomeration. Nevertheless, it needs to reflect the agglomeration status of a city due to population migration and explore the impact of population (re)distribution on urban economic development. In addition, the per capita Gross Domestic Production (GDP) growth rate is mainly used when measuring urban economic growth. It undeniably reflects the speed of urban economic growth and the dynamics of economic development in a certain period. However, per capita GDP is more applicable to cities' self-comparison in time series or comparison between cities of similar size. Because the implicit assumption of per capita economic indicators is that the population size and economic level are linearly related, ignoring the agglomeration effect caused by the nonlinear interaction between the two in social dynamics and spatial organization (Lei et al., 2021).

More importantly, economists' theoretical assumptions and logical, empirical evidence about urban economic growth are attempts to address economic issues outside of geography and do not reflect the local character (Martin, 1999). Cities are not isolated and self-contained entities but an integral part of geographic systems and territorial functions (Sigler & Martinus, 2017; Taylor & Derudder, 2015). Moreover, an urban organization is constantly changing in response to changes in global capital accumulation, national political strategies, and regional environmental reorganization (Ahani & Dadashpoor, 2021; Ye & Liu, 2020). The economic fate of a city depends not only on the participating elements within the city but also on the growth and functioning of the national and regional economies in which the city is located (Lin, 1999; Hens et al., 2018; Nguyen et al., 2019). Therefore, cities should be placed in national or region-specific contexts when discussing and evaluating urban economic growth. Especially in China, the national urban system under the socialist urbanization model consists of several juxtaposed and complex regional systems that have developed in different historical contexts at other times (Chen et al., 2019). In the context of globalization and urban networks, the close functional ties and vital competing roles among cities increasingly highlight the status and role of urban agglomerations in regional development and competition. Therefore, urban agglomerations are the appropriate spatial scale to explore urban economic performance in a regional context (Ye et al., 2019). At this stage, urban agglomerations are the center of gravity of China's regional development spatial strategy and play the role of regional economic growth poles (Fang, 2019; Fang & Yu, 2017). The coordinated and sustainable development of urban agglomerations is one of the most critical factors affecting China's regional economic pattern (Guan et al., 2018; Hui et al., 2020). The balance between urban agglomerations' population and economic development directly affects the development level and pattern enhancement of China's regional economy (Liu et al., 2022; Lyu & Jiang, 2022). It is an essential aspect of narrowing the regional development gap and balancing the economic development pattern in China in the coming period.

This paper constructs a conceptual and theoretical framework of urban population agglomeration capacity (UPAC) and urban economic efficiency (UEE). The object of this empirical study is the Yangtze River Delta Urban Agglomerations (YRDUA)—one of the fastest-growing urban agglomerations in China and currently at a critical stage of the urban growth rates shifting to high-quality, coordinated development. We first evaluate the UEE from the perspective of urban scaling law, analyze the heterogeneity of UPAC caused by population flow and migration by applying Zipf's law and explore the explanatory mechanism of UPAC affecting UEE.

2 Conceptual definition and theoretical framework

2.1 Economic efficiency concept based on urban scaling law

The urban scaling law becomes a mathematical theory that explores the problem of spatial complexity and complex mechanisms of urban systems in a geographic sense (Li et al., 2017). It reflects the generalized equation of universal geographic laws at the macro-level of urban geography (Chen, 2012; Chen & Jiang, 2018), with the expression (Bettencourt et al., 2010),

$$Y = Y_0 N^{\beta} \tag{1}$$

where Y denotes the element value of the city; Y_0 is the standardization factor; N denotes the city population size; β is the scaling factor, which indicates the scaling relationship between the element and the population in the urban system.

The scaling relationship's magnitude varies with the study area's scale. It reflects an idealized relationship between a city's factor and the population in a particular area scale (Bettencourt, 2013). Then, under this theoretical scaling relationship, if a city's population is known, there is a theoretically expected value of the economic level. However, there may be a specific difference between the expected economic value and the actual one of a city, which is defined in this paper as the "urban economic efficiency" to measure the economic performance of a city, i.e., the extent to which the actual economic performance of a city exceeds the expected level, which can be expressed in double logarithmic coordinates as

$$E_i = \log \frac{Y_i}{Y(N_i)} = \log \frac{Y_i}{Y_0 N_i^{\beta_e}} = \log Y_i - \log Y_0 N_i^{\beta_e} \tag{2}$$

where E_i refers to the UEE of the city i; Y_i is the actual value of the city economy (GDP); $Y\left(N_i\right)$ is the expected value of the city economy at population size N_i and can be obtained by having Eq. (1). β_e refers to the scaling relationship between urban economic factors and population size.

The higher the value of UEE, the better the city's economic performance and vice versa. Three economic development patterns can be classified based on UEE: when $E > 0$, the actual performance is better than expected; $E = 0$, the actual performance is equal to expected; $E < 0$, the actual performance is inferior to anticipated.

2.2 Rank size of urban migrant population and agglomeration capacity

The rank-size law of total population (TP) size distribution is widely proven (Xu & Harriss, 2010). The Zipf model is the most classical, referring to the city size distribution showing a straight line with a slope of 1 on the double logarithmic rank-size plot. In the empirical analysis, the regression is usually performed in the form of

$$\ln\left(\mathrm{TR}_i - 1/2\right) = \ln k_T - \lambda_T \ln \mathrm{TP}_i + \varepsilon \tag{3}$$

where $\mathrm{TP}_i$and TR_i refer to the TP size and its rank order of city i, respectively. The rank order $TR_i - 1/2$ is a common practice referred to Gabaix and Ibragimov (2011) to make the equation fit well. k_T is a constant, ε is a perturbation term. λ_T is the Pareto index of the TP model. An enormous value indicates a more dispersed city size distribution and a smaller population gap between cities. Conversely, its smaller value indicates a more clustered city size distribution and a more significant population gap between cities. The magnitude and variation of the Pareto index, the slope presented on the bilogarithmic plot of rank size, reflect important geographical significance (Li & Sui, 2013). First, the magnitude of the absolute value of the slope, if $|\lambda_T| > 1$, indicates more major urban development in the lower order; if $|\lambda_T| < 1$, more major urban development in the higher order. The second is the change in the absolute value of the slope. If the change is significant, it indicates that the agglomeration force influencing the city size distribution is larger than the dispersion force and vice versa.

Studies have shown that China's migrant population (MP) is concentrated in a few large cities. The TP size of a city positively affects its attraction and concentration of MP. Accordingly, drawing on the Zipf model, this paper constructs the form that the MP obeying the rank-size power function as

$$\ln\left(\mathrm{MR}_i - 1/2\right) = \ln k_M - \lambda_M \ln \mathrm{MP}_i + \varepsilon \tag{4}$$

where $\mathrm{MP}_i, \mathrm{MR}_i$ refer to the size and rank order of the MP in city i, respectively; λ_M is the Pareto index in the MP model with the same meaning as the TP model. k_M is a constant, and ε is a disturbance term.

Based on the coupling relationship between the rank size of the MP to the TP in a city, we measure the MP absorption capacity of a city relative to its TP size. Ideally, large cities concentrate more on MP, and small cities concentrate less. In other words, the MP concentrated in each city is linearly related to its own TP size. As shown in Fig. 1a, ideally, the double logarithmic curve of TP rank size and the double logarithm of MP rank size are parallel. The MP order of any city i is the same as the TP order, and the difference of vertical coordinate $\Delta R_i = 0$. Meanwhile, the ratio of MP to TP is constant, equal to the ratio of MP to TP in the city cluster. The difference of horizontal coordinate ΔP_i is a constant C. The specific formula is expressed as

$$\Delta R_i = \ln\left(\mathrm{MR}_i - 1/2\right) - \ln\left(\mathrm{TR}_i - 1/2\right) = \ln\left({}^{\mathrm{MR}_i - 1/2}\big/_{\mathrm{TR}_i - 1/2}\right) \tag{5}$$

$$\Delta P_i = \ln \mathrm{MP}_i - \ln \mathrm{TP}_i = \ln\left(\mathrm{MP}_i/\mathrm{TP}_i\right) \tag{6}$$

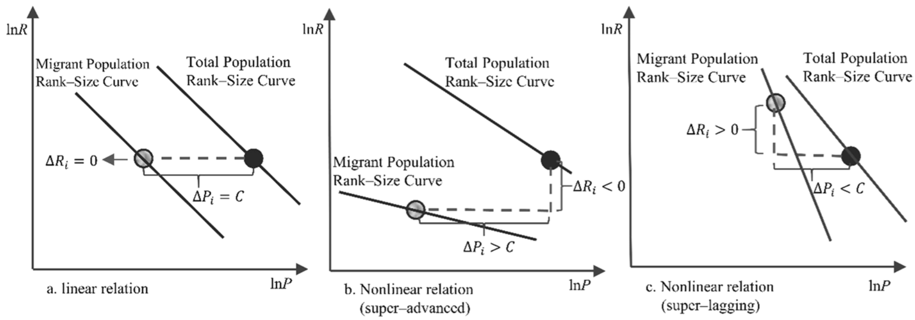

Fig. 1 Coupling relationship among the rank-size distribution curves of total urban population (TP) and mobile population (MP)

$$C = \ln\left(\sum_{i=1}^{n} \mathrm{MP}_i / \sum_{i=1}^{n} \mathrm{TP}_i\right) \tag{7}$$

However, non-ideally, the magnitude and combination of ΔR_i, ΔP_i will show four coupling patterns.

① When $\Delta R_i \leq 0, \Delta P_i \geq C$, it is the super-advanced type, indicating that the MP of the city i is advanced in terms of rank size (Fig. 1b). When $\Delta R_i = 0, \Delta P_i = C$, it is the ideal state of population distribution.
② When $\Delta R_i > 0, \Delta P_i \geq C$, it is the sub-advanced type, indicating that the MP of the city i lags in the rank order but is advanced in size.
③ When $\Delta R_i \leq 0, \Delta P_i < C$, it is a sub-lagging type, indicating that the MP of the city i is ahead in rank order but lagging in size.
④ When $\Delta R_i > 0, \Delta P_i < C$, it is a super-lagging type, indicating that the MP of the city i lags in both rank order and size (Fig. 1c).

2.3 Theoretical model of urban population agglomeration capacity affecting economical efficiency

The literature needs to provide a complete explanatory framework on whether UPAC affects UEE. Therefore, this paper attempts to review empirical studies related to population agglomeration based on new economic geography and explore the explanatory mechanism of UPAC on UEE from the perspective of uneven population distribution and its nonlinear relationship with economic growth efficiency. The new economic geography sees the long-term growth of urban economies due to population agglomeration facilitated by the returns to scale of "knowledge" (Polese, 2005; Romer, 1994; Sunley et al., 2020). The theoretical framework of this paper can draw on the basic ideas of urban economic growth theory. However, UEE focuses more on regional equilibrium and inter-city growth differences than urban economic growth.

The assumptions of UEE include: (1) urban population agglomeration is uneven and is the result of social mobility based on people's autonomous choice preferences; (2) cities have a positive growth trend, and urban economic growth curves are in logarithmic form; and (3) a scaling relationship between TP and economy is assumed. Based on these theoretical assumptions, there is some similarity in the spatial distribution of UPAC and UEE at a specific regional scale, and geographic detectors can identify this relationship. The rate and size of population agglomeration embodied in each UPAC will have different positive or negative effects on UEE. This process can be tested by UEE models (regression models). Consequently, the theory of UEE can be briefly expressed as the varying concentration of population among cities is the geospatial process and manifestation of UEE differences, which implies the uneven growth relationship of the urban economy at a specific regional scale. In this theoretical perspective, the impact of UPAC on UEE can be understood as: the population agglomeration in cities is a process of urban economic growth. Population agglomeration usually promotes economic growth. Nevertheless, the positive role of UPAC on UEE will decline or even be reversed when the rate and scale of population exceed a certain threshold of urban carrying capacity.

The geographic detection model is a statistical model that identifies spatial heterogeneity and reveals the driving forces behind it. In this paper, we first apply the geographic detection model to identify the extent and variation of the influence of UPAC on UEE. The core idea is that the factors of UPAC that affect UEE are spatially heterogeneous. Suppose there is significant consistency or similarity in the spatial distribution of UPAC and UEE. In that case, it indicates that the spatially distinguishing factors of UPAC have an essential impact on UEE. The model expression (Wang & Hu, 2012; Wang et al., 2010),

$$P_U = 1 - \frac{1}{n\sigma_U^2} \sum_{j=1}^{m} n_j \sigma_{U_j}^2 \tag{8}$$

where P_U is an indicator of the explanatory index of UPAC on UEE; n is the number of cities in the study area; m is the number of UPAC types; n_j is the number of cities with UPAC in type j; σ_U^2 is the variance of UEE coefficients of all cities in the study area; $\sigma_{U_j}^2$ is the variance of UEE coefficients of cities in type j. The range of P_U is [0,1], and the larger its value, the greater the influence of UPAC on UEE.

An ordinary least squares (OLS) linear model, based on the logic and characteristics of UEE, is used to study the linear relationship between UEE (dependent variable) and influencing factors (independent variable). The premise of the OLS model is assumed that the variables are independent of each other. The spatial information of the variables is ignored in the model. The formula is expressed as

$$E_{it} = \alpha_1 X_{1it} + a_2 X_{2it} + a_3 X_{3it} + \varepsilon_{it} \tag{9}$$

where $X_{1_{it}} = (\mu_1 \Delta R_{it} + \mu_2 \Delta P_{it})$, indicating the UPAC, the meanings of $\Delta R_{it}, \Delta P_{it}$ are the same as Eqs. (5–7), μ_1, μ_2 are the weight coefficients of $\Delta R_{it}, \Delta P_{it}$, respectively.$X_{2it}$ is the total population size of the city;X_{3it} is the city's economic size.$\alpha_1, \alpha_2, \alpha_3$ are the coefficients of the fitted relationship; ε_{it} is the error term of the model. The cities' total population and economic size are taken in the logarithmic form to eliminate the influence of the difference in magnitudes on the results.

Equation (9) is the base model for this paper to explore the effect of UPAC on UEE, where $X_{1_{it}}$ is the core variable. In order to overcome the interference of causality and endogeneity issues between UPAC and UEE, finding appropriate instrumental variables becomes an essential element in the measurement process. The generalized matrix estimation method (GMM) provides an excellent analytical method for the study. In this paper, $X_{2_{it}}$ and X_{3it} are set as endogenous variables expressed through lnPOP and lnGDP, respectively. The endogenous variables are determined by instrumental variables such as human capital, technological progress, and industrial structure. The share of urban college students in the total population expresses human capital. Technological progress is a percentage of GDP invested in science and technology. The industrial structure includes the share of the secondary and tertiary industries.

Based on the basic linear model described above, the Spatial Lag Model (SLM) considers the effect of a city's economic efficiency on other cities, i.e., the spatial spillover effect, and is a form of the spatial regression model. The expression is that

$$E_{it} = \rho \mathrm{WE}_{it} + \alpha_1 X_{1it} + a_2 X_{2it} + a_3 X_{3it} + \varepsilon_{it} \tag{10}$$

where ρ is the coefficient value of spatial autoregression, and W denotes the spatial weight matrix.

The Spatial Durbin Model (SDM), based on the SLM, takes into account the influence of population concentration and economic scale factors of a city on the economic effectiveness of other cities with the expression that

$$E_{it} = \rho \mathrm{WE}_{it} + \alpha_1 X_{1_{it}} + a_2 X_{2it} + a_3 X_{3it} + b_1 X_{1_{it}} + b_2 X_{2jt} + b_3 X_{3it} + \varepsilon_{it} \tag{11}$$

In spatial regression analysis, there may be spatial autocorrelation in the independent error terms of the model. The Spatial Error Model (SEM) can consider the spatial spillover effect of the independent error terms. The basic form of SEM is that

$$E_{it} = \eta W \varphi + \alpha_1 X_{1it} + a_2 X_{2it} + a_3 X_{3it} + \varepsilon_{it} \tag{12}$$

where η is the spatial autocorrelation coefficient of the error term, φ represents the spatial autocorrelation error term.

3 Study area and data sources

3.1 Study area

Regarding the scope of the YRDUA as defined in the "Yangtze River Delta Urban Agglomeration Development Plan" promulgated in 2016, 62 cities were identified for the study (Fig. 2). They were taking into account the latest administrative division adjustment. Although the cities in Anhui Province were included in the "Plan" only in 2016, a Hefei-centered metropolitan area and close economic and social interactions with other cities had been formed. The population size classes of cities in the YRDUA refer to the "Plan" classification criteria (Table 1).

Fig. 2 Location of the study area

Table 1 Population size classes of cities in the YRDUA

City size		Classification criteria (Resident population in urban areas)	Cities
Mega		> 10 million	Shanghai
Super		5–10 million	Nanjing
Large	Type I	3–5 million	Hangzhou, Hefei, Suzhou
	Type II	1–3 million	Wuxi, Ningbo, Nantong, Changzhou, Shaoxing, Wuhu, Yancheng, Yangzhou, Taizhou(Su), Taizhou(Zhe)
Medium		0.5–1 million	Zhenjiang, Huzhou, Jiaxing, Maanshan, Anqing, Jinhua, Zhoushan, Yiwu, Cisi
Small	Type I	0.2–0.5 million	Tongling, Chuzhou, Xuancheng, Chizhou, Yixing, Yuyao, Changshu, Kunshan, Dongyang, Zhangjiagang, Jiangyin, Danyang, Zhuji, Rugao, Dongtai, Shengzhou, Wenling, Taixing, Lanxi, Tongxiang, Taicang, Jingjiang, Yongkang, Gaoyou, Haining, Qidong, Yizheng, Xinghua, Liyang
	Type II	< 0.2 million	Tianchang, Ningguo, Tongcheng, Pinghu, Yangzhong, Jurong, Mingguang, Jiande

Many studies suggested that the municipal district or built-up area is more suitable for heterogeneous scale analysis of Chinese cities (Chen, 2010). To more accurately reflect the economic development pattern of an urbanized area in the true sense of "urban," the unit of measurement in this paper is the municipal district, where population and non-agricultural activities are densely distributed, rather than the administrative city area, which includes non-urbanized areas.

3.2 Data sources

The TP is the total resident population in the city's municipal area. The MP is defined as the separated population from outside the city's administrative area into the city's municipal area. The population data are obtained from the county and city scale data of the fifth, sixth and seventh censuses. MP is the sum of the "population moving in from other cities (counties) and urban areas within the province" and "population moving in from other provinces" in the census data. Table 2 shows the TP and MP descriptive statistical analysis.

The "urban economy" is represented by the gross regional product of urban municipalities, and the original data are obtained from the China Urban Statistical Yearbook in 2000, 2010, and 2020.

2000, 2010, and 2020 are selected as the study time points based on two primary considerations. On the one hand, China's national census data are conducted once every 10 years. These 3 years are chosen because the data for these years are more comprehensive than any previous years. On the other hand, since the reform and opening up, China's urban economy's operation and growth have shown a clear "decade-by-decade cycle." Although the financial crisis in 2008 and the economic transition in 2012 have slightly changed the characteristics of cyclical fluctuations. Both growth and cyclical fluctuations are deeply rooted in China's institutional arrangements and the regulatory mechanisms of a constructive government (Chang et al., 2016). In such economies, where the political system is well established, the division of power is generally recognized, social conditions are stable, and the effects of cyclical fluctuations tend to be mitigated (Smirnov, Ozyildirim, & Picchetti, 2019). Therefore, it is reasonable to use 10 years as the stage division of the study (Zhao, Jia, & Chang, 2019).

4 Results and analysis

4.1 Urban economic efficiency and evolution

Before measuring the UEE, it is necessary to verify the applicability of the "scaling law" of cities. Then carries on the scaling law fitting to the urban economy and the population. The scaling relation is judged according to the parameters of the fitted equation, and the urban system's characteristics are analyzed.

The fitted determination coefficients R^2 of population and economy for 2000, 2010, and 2020 are 0.7145, 0.8298, and 0.9074, respectively, indicating that the urban economy shows a power-law relationship with population size and is increasingly significant. The significant power-law relationship shows that it conforms to the urban scaling law, so it is feasible to measure the urban economic performance from the scaling law perspective (Fig. 3). The fitted equation β values for the 3 years are 1.2134, 1.2496, and 1.1704, respectively. They are all more remarkable than the general threshold of 1.15 for quantitative

Table 2 Descriptive statistical analysis of the TP and MP

Year	Sample name	Sample size	Minimum value	Maximum value	Average value	Standard deviation	Median value
2000	TP	62	30.167	1434.853	114.852	180.336	75.158
	MP	62	1.301	493.919	26.329	64.092	11.970
2010	TP	62	33.498	2231.547	283.677	355.244	125.016
	MP	62	4.162	1249.076	66.519	164.073	27.084
2020	TP	62	31.546	2270.815	214.301	326.173	117.314
	MP	62	−39.93	1012.730	57.384	143.530	20.274

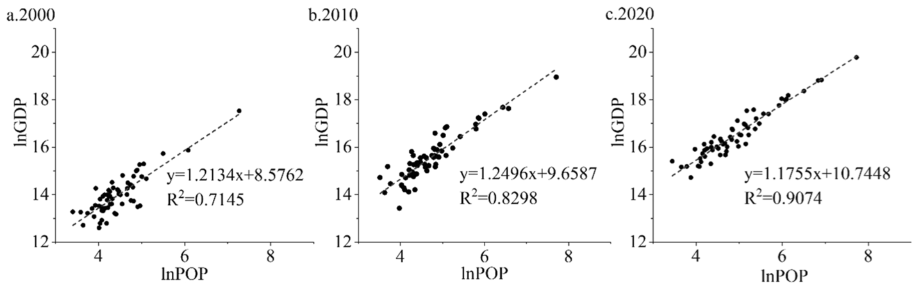

Fig. 3 Fitting result of scaling relationship between urban population and economy

socioeconomic indicators (Bettencourt et al., 2008), indicating that the economy and population have a superlinear relationship. This scaling relationship reflects the cumulative payoff effect of population agglomeration. The change of β value also indicates that the payoff of scale effect of population agglomeration increases first and then decreases.

According to the formula of UEE, the UEE values in the YRDUA in 2000, 2010, and 2020 are calculated. As shown in Table 3, it can be seen that the mean and median UEE tend to be “0” in the 3 years. It indicates that the positive and negative UEE is roughly in a “half-and-half” city distribution. Roughly half of the cities performed better than expected to match their population concentrations, while the other half underperformed. The standard deviation of UEE in the 3 years gradually decreases, indicating that the gap between UEE decreases. Cities with significant positive UEE in 2000, such as Taicang, Wuxi, Ningbo, Suzhou, Hangzhou, and Wuhu, generally showed a gradual decline in 2010 and 2020. On the contrary, cities with significant negative UEE in 2000, such as Hefei, Liyang, Yizheng, Dongtai, Chizhou, and Xinghua, are characterized by a gradual increase in 2010 and 2020. In addition, cities with balanced population–economy development, such as Taizhou, Jinhua, Wenling, and Dongyang, whose UEE value was around “0” in 2000, also showed a gradual decrease in 2010 and 2020.

In terms of spatial distribution (Fig. 4), in 2000, the negative UEE was mainly distributed in the northern and western parts of the YRDUA. In comparison, the positive UEE was mainly distributed in the central part, with a clear “high–high” agglomeration feature (*Local Moran’s I* of 0.203). In 2010, the number of cities with negative UEE increased significantly, mainly in the peripheral cities. While the cities with positive UEE were still mainly distributed in the central part, forming an evident “high–high” agglomeration and “high–low” agglomeration (*Local Moran’s I* is 0.117). Most of the cities in the central

Table 3 Statistical descriptive analysis of UEE in YRDUA

Year	Sample size	Minimum value	Maximum value	Mean value	Standard deviation	Median value
2000	62	− 1.084	0.903	0.000	0.452	0.059
2010	62	− 1.198	0.907	0.000	0.423	− 0.032
2020	62	− 0.688	0.691	0.007	0.311	− 0.027

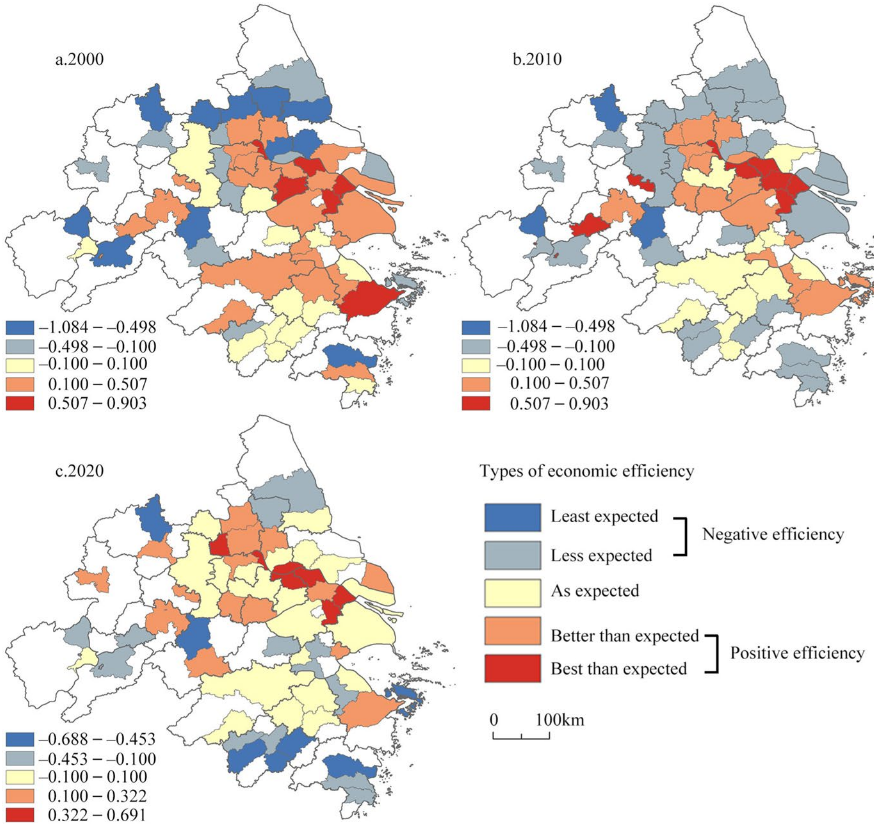

Fig. 4 Spatial distribution and evolution of UEE

part of the region, especially those around the value of "0," still have the characteristics of "high–high" agglomeration in 2020, but their significance decreases (*Local Moran's I* is 0.381). The change of *Moran's I* indicate that the spatial agglomeration of UEE in the YRDUA shows a change of increasing and then decreasing, reflecting the trend of regional economic development toward equilibrium.

4.2 Urban population agglomeration capacity and its influence mechanism

With "size" as the independent variable and "rank" as the dependent variable, the rank-size scatter plots and fitted curves are plotted for the TP and the MP in 2000, 2010, and 2020, respectively (Fig. 5). The results show that the double logarithmic curves of the rank-size distribution of the TP and the MP are well-fitted (the fitted values of R^2 are 0.9291, 0.8871, and 0.9676 for the TP and 0.9289, 0.9335, and 0.8143 for the MP in 2000, 2010 and 2020, respectively). Moreover, they all obey the power function distribution law, and the smaller the scale, the larger the order, showing the decreasing characteristic. The Pareto index $|a|$ of the TP in 2000, 2010, and 2020 is 1.5977, 0.9331, and 1.1662, respectively. 2010 and 2020 converge to the expected value of 1, indicating that the city size distribution is more in line with Zipf's law and the population distribution

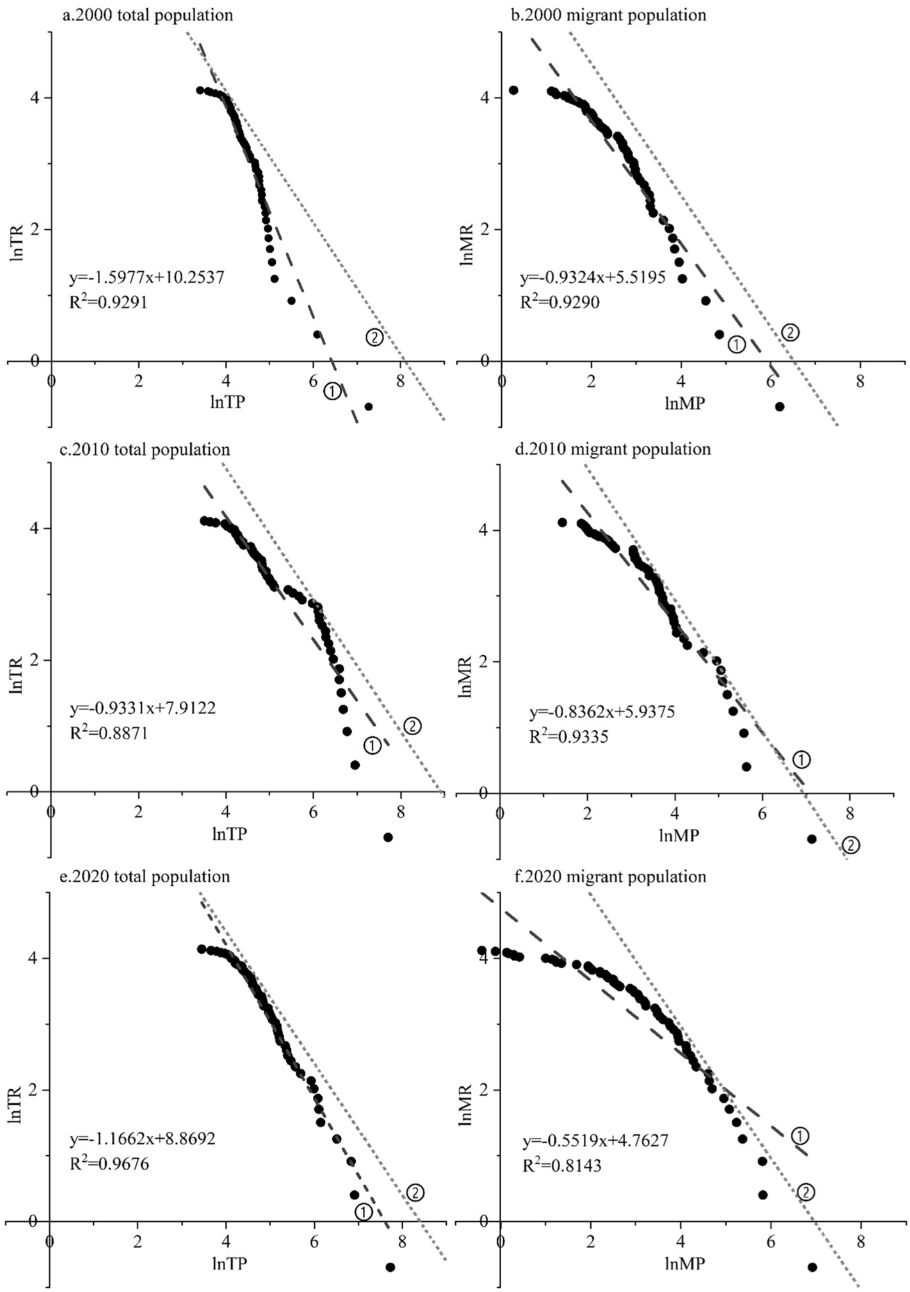

Fig. 5 Rank-size fitting of urban TP and MP. ① is the fitted curve, ② is the reference line with index 1

among cities is relatively balanced. The city rank's two ends deviate from the main body's linear distribution in the rank-size scatter plot. It indicates that the actual population scale of high- and low-rank cities is smaller than the expected value of Zipf's law. They reflect the characteristics of population-scale collapse in the mega/super/large cities and small cities in the YRDUA.

The Pareto index |a| of the MP in 2000, 2010, and 2020 is 0.9324, 0.8362, and 0.5519, respectively, which keeps decreasing, indicating that the concentration of the MP is decreasing. The distribution is becoming more and more dispersed. In the rank-size scatter plot, the two ends of the city rank also have the characteristic of deviating from the linear distribution of the main body. It reflects the significant characteristics of the collapse of the MP scale in the mega/super-cities and small cities in the YRDUA. The actual size of the MP in the middle-rank cities is larger than the expected value of Zipf's law. It indicates that these cities have the advantage of absorbing the MP, which leads to a significant polarization of the MP distribution in the YRDUA.

Obviously, from the above analysis results, it is clear that the ideal state of "a linear relationship between the size of the MP and the TP of a city" is difficult to exist. The relationship between the two needs to be more aligned. In the YRDUA, very few cities have the same rank of MP as their TP (such as Shanghai, Nanjing, Hangzhou, Hefei, and Ningbo in 2000, Shanghai and Ningbo in 2010, and Shanghai, Hefei, and Tianchang in 2020). No city has the same proportion of MP to TP as the average of the urban agglomeration. According to the previous coupling relationship between the MP and the TP, the four UPACs are calculated. The proportion of MP to TP is 22.92%, with a *C* value of − 1.4730 in 2000, 23.45% and − 1.4504 in 2010, and 26.78% and -1.3176 in 2020. The related results are shown in Table 4. There is no statistically significant and strictly positive correlation between UPAC and size. In general, the UPAC of (mega/super) large cities is mostly super-advanced, while that of medium and small cities is mostly lagging. However, it also varies depending on the urban function, development policy, and strategic location. In 2010 and 2020, some medium and small cities, such as Jiangyin, Taicang, Changshu, and Zhangjiagang, were super-advanced, and some super/large cities, such as Nanjing, Yangzhou, Shaoxing, and Wuhu, as super-lagging type in 2020. The proportion of cities with super-advanced and super-lagging UPAC is more significant, followed by cities with sub-lagging and most minor sub-advanced types.

UPAC's clustering distribution is becoming more significant (Fig. 6). In 2000, the UPAC of advanced and lagging types was randomly distributed (with Local Moran's I being -0.083 and P value being 0.235). While in 2010, the agglomeration characteristics of super-advanced and super-lagging types were prominent (with *Local Moran's I* = 0.118, *P value* = 0.235), concentrated in the middle, south, and north of the region, respectively. In 2020, the spatial agglomeration characteristics of UPAC were more significant (with *Local Moran's I* being 0.240 and *P value* being 0.000). The super-advanced types are concentrated in the region's central part, the super-lagging types are in the northern and western parts, and the sub-lagging types are in the northern and southern periphery.

A geographic detector was used to identify the influence of UPAC on UEE. The results show (Table 5) that the influence coefficients in 2000, 2010, and 2020 are 0.2500, 0.1191, and 0.0225, respectively, with gradual decrease and significance. They indicate that the influence of UPAC on UEE gradually decreases. In other words, the scale effect of population agglomeration has yet to play a decisive role in urban economic performance.

The detector method identified the influence degree of UPAC on UEE. However, the effect direction of different UPACs on UEE could not be judged, so regression methods conducted further analysis. Before the regression analysis, the UPAC needs to be

Table 4 Statistical analysis of UPAC

Type	Year	Number of cities	Proportion (%)	Cities	City size class
Super-advanced	2000	18	29.03	Shanghai, Hefei, Wuxi, Kunshan, Yiwu, Nanjing, Hangzhou, Ningbo, Changzhou, Suzhou, Nantong, Zhenjiang, Yangzhou, Wuhu, Jinhua, Shaoxing, Maanshan, Anqing	Mega, Large (I, II)
	2010	26	41.94	Shanghai, Hefei, Wuxi, Kunshan, Yiwu, Nanjing, Hangzhou, Changzhou, Ningbo, Jiangyin, Taicang, Changshu, Zhangjiagang, Haining, Yongkang, Cisi, Tongxiang, Pinghu, Yuyao, Wenling, Tongcheng, Zhuji, Zhoushan, Yixing, Dongyang, Danyang	Mega, Large(I, II), Small(I)
	2020	20	32.26	Shanghai, Hefei, Wuxi, Kunshan, Yiwu, Suzhou, Jiaxing, Jinhua, Huzhou, Chuzhou, Jiangyin, Taicang, Changshu, Zhangjiagang, Haining, Yongkang, Cisi, Tongxiang, Pinghu, Yuyao	Mega, Large(I, II), Medium, Small(I)
Sub-advanced	2000	0	0	–	
	2010	0	0	–	
	2020	3	4.84	Hangzhou, Nanjing, Changzhou	Large(I, II)
Sub-lagging	2000	16	25.81	Yangzhong, Jiande, Jurong, Ningguo, Jingjiang, Taicang, Yongkang, Cisi, Pinghu, Wenling, Taizhou, Yancheng, Jiaxing, Zhoushan, Chuzhou, Tongling	Medium, Small(I, II)
	2010	9	14.51	Yangzhong, Jiande, Jurong, Ningguo, Jingjiang, Shengzhou, Yizheng, Jiaxing, Wuhu	
	2020	15	24.19	Yangzhong, Jiande, Jurong, Mingguang, Tianchang, Shengzhou, Yizheng, Lanxi, Liyang, Dongtai, Danyang, Linhai, Dongyang, Anqing, Maanshan	
Super-lagging	2000	28	45.16	Chizhou, Xuancheng, Taizhou(Zhe), Xinghua, Rugao, Taixing, Gaoyou, Qidong, Huzhou, Mingguang, Dongtai, Tianchang, Linhai, Liyang, Lanxi, Danyang, Changshu, Zhangjiagang, Zhuji, Dongyang, Yizheng, Shengzhou, Yixing, Haining, Tongxiang, Yuyao, Jiangyin, Tongcheng	Small(I)
	2010	27	43.55	Chizhou, Xuancheng, Taizhou(Zhe), Xinghua, Rugao, Taixing, Gaoyou, Qidong, Huzhou, Mingguang, Dongtai, Tianchang, Linhai, Liyang, Lanxi, Suzhou, Yangzhou, Zhenjiang, Nantong, Taizhou, Yancheng, Shaoxing, Jinhua, Maanshan, Tongling, Anqing, Chuzhou	Large(I), Medium, Small(I)
	2020	24	38.71	Chizhou, Xuancheng, Taizhou(Zhe), Xinghua, Rugao, Taixing, Gaoyou, Qidong, Nanjing, Zhenjiang, Yangzhou, Nantong, Taizhou, Yancheng, Wuhu, Shaoxing, Zhoushan, Tongling, Tongcheng, Ningguo, Jingjiang, Zhuji, Wenling, Yixing	Super, Large(II), Medium, Small(I)

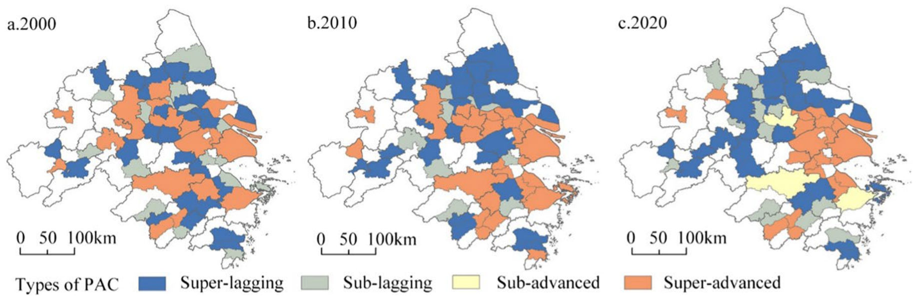

Fig. 6 Spatial distribution and evolution of UPAC

Table 5 The influence and change of UPAC

	2000	2010	2020
q value	0.2500***	0.1191**	0.0225*
p value	0.000	0.0378	0.093

Table 6 Weighting coefficients of UPAC

Year	2000	2010	2020
ΔR_i	0.6759	0.4657	0.5030
ΔP_i	0.3241	0.5343	0.4970

Table 7 Spatial dependence test for each array (LM test)

Year	Array	UPAC	LMlag	R-LMlag	LMerr	R-LMerr	LM(SARMA)
2000	A	Super-advanced	0.556	1.326	0.090	0.861	1.417
	B	Sub-lagging	2.378	2.126	0.580	0.328	2.706
	C	Super-lagging	0.596	0.168	0.430	0.002	0.598
2010	D	Super-advanced	4.24**	0.019	5.324**	1.103	5.343
	E	Sub-lagging	1.246	0.08	1.282	0.116	1.361
	F	Super-lagging	2.505*	0.304	3.741**	1.539	4.044
2020	G	Super-advanced	8.795***	1.715*	7.089***	0.009	8.804
	H	Sub-advanced	0.054	0.079	0.003	0.028	0.082
	I	Sub-lagging	0.499	7.276***	5.327**	12.104***	12.603
	J	Super-lagging	0.219	2.257	0.411	2.448	2.667

***, **, * correspond to significance levels of 0.01, 0.05, and 0.1, respectively

transformed from qualitative to quantitative data. The qualitative data are jointly determined by the rank difference ΔR_i and the size difference ΔP_i between the TP and the MP, so the weight values are calculated by the entropy value method (Table 6). Then, the quantitative data of UPAC are calculated by weighting and grouped for regressions.

The results of the spatial dependence test (LM test) to select the appropriate spatial econometric model are shown in Table 7. 2000 LMlag and LMerr for arrays A, B, and C failed the significance test. 2010 LMlag and LMerr for arrays D and F passed the significance test, but array E failed the significance test. 2020 LMlag and LMerr for arrays G passed the significance test, and array I passed the significance test by LMerr. However, arrays H and J failed the significance test.

For the arrays that passed the spatial dependence significance test, OLS, GMM, and the corresponding spatial regression models (SLM, SEM, SDM) were used to analyze and screen the optimal models (see the Appendix for a comparison of the main parameters of the various models). Table 8 shows the regression results of the optimal models for each array.

The best fit of the OLS model in 2000 for all three arrays of A, B, and C indicates that the spatial spillover effect of UEE in this period is so weak as to be negligible. At the same time, it also indicates that the effect of endogenous factors in the economy can be neglected. The regression coefficients of X_1 for three arrays are all positive, indicating a positive effect of UPAC on UEE. However, X_2 is negative, indicating that the population does not form a scale benefit, which verifies the conclusion that the endogenous growth effect does not exist. Comparing the three arrays A, B, and C, it is found that the more lagging the cities, the greater the negative effect of population size (X_2) on UEE. So population size has a constraining effect. Economic size (X_3) positively affects UEE. The more lagging the cities, the greater its effect, highlighting the importance of urban economic development.

The best fit of both D and F in the SDM model in 2010 indicates that the spatial spillover effect of UEE is evident for both super-advanced and super-lagging cities. The best fit of E in the GMM model indicates that the endogenous effect of UEE is noticeable for sub-lagging cities. The regression coefficients of X_1 and WX_1 in group D do not pass the significance test ($P=0.8874$, $P=0.9176$), indicating that UPAC does not significantly affect the UEE of super-advanced cities. The regression coefficients of X_2 and WX_2 are -0.4646 and -0.3874, respectively. They pass the significance test, indicating that population size plays a significant adverse effect and has a spatial spillover effect, which is a constraint to the UEE of super-advanced cities. The regression coefficients of X_3 and WX_3 are 0.1377 and 0.0515, respectively, and pass the significance test. Thus the economic scale exerts a significant favorable influence and has a spatial spillover effect, which is a critical factor in improving the UEE of super-advanced cities. The regression coefficients of WE are 1.0063 and significant, indicating that the UEE has a spatial agglomeration effect. The regression coefficient of X_1 in array E is 0.1280, which does not pass the significance test, but indicates that UPAC begins to exert a positive influence on the UEE of sub-lagging cities. X_2 and X_3 still exert a significant negative and positive influence, respectively, as in array D. The regression coefficient of X_1 in array F is 0.4143 and significant, indicating that it has a positive effect on the UEE of super-lagging cities, which in contrast to super-advanced cities. The effects of X_2 and X_3 on UEE and their spatial spillover effects are also in contrast to super-advanced cities.

The 2020 G, H, I, and J arrays fit best in the SDM, OLS, SEM, and GMM models. The regression coefficient of X_1 in array G is 0.7001, which does not pass the significance test but indicates the positive effect of UPAC on UEE. The regression coefficient of X_1 in array H is -0.0738, which does not pass the significance test but indicates the negative effect of UPAC on UEE. The regression coefficient of X_1 in array I is 0.6423 and significant,

Table 8 Regression coefficients of the effect of UPAC on the UEE for the optimal model of each array

Array	Sample size	Model	R^2	Independent variable	Regression coefficient	Standard error	t/z statistical values	*P* value
A	18	OLS	0.6473	X_1	0.5994***	0.1619	3.7000	0.0005
				X_2	– 0.4431***	0.0871	– 5.0842	0.0000
				X_3	0.1931***	0.0349	5.5272	0.0000
B	16	OLS	0.5510	X_1	1.5113***	0.3670	4.0853	0.0001
				X_2	– 1.0078***	0.1252	– 8.0460	0.0000
				X_3	0.3914***	0.0521	7.5073	0.0000
C	28	OLS	0.4510	X_1	0.5323**	0.2546	2.0910	0.0409
				X_2	– 1.8216***	0.3401	– 5.3564	0.0000
				X_3	0.5949***	0.1170	5.0865	0.0000
D	26	SDM	0.2970	X_1	– 0.0366	0.2581	– 0.1417	0.8874
				X_2	– 0.4646***	0.1305	– 3.5607	0.0004
				X_3	0.1377***	0.0393	3.5063	0.0005
				WE	1.0063***	0.2001	5.0301	0.0000
				WX_1	– 0.0260	0.2512	– 0.1034	0.9176
				WX_2	– 0.3874***	0.1142	-4.2670	0.0000
				WX_3	0.0515***	0.0358	4.2313	0.0000
E	9	GMM	0.3229	X_1	0.1280	0.1529	0.8373	0.4024
				X_2	– 0.2388***	0.0824	– 2.8970	0.0038
				X_3	0.0907***	0.0338	2.6865	0.0072
F	27	SDM	0.2896	X_1	0.4143**	0.2211	1.8743	0.0609
				X_2	0.2037	0.1355	1.5037	0.1326
				X_3	– 0.0478	0.0493	– 0.9684	0.3328
				WE	– 1.0561**	0.3922	– 0.6505	0.0071
				WX_1	0.4582**	0.2527	1.8136	0.0697
				WX_2	0.1007	0.1340	0.7518	0.4522
				WX_3	– 0.0119	0.0529	– 0.2243	0.8225
G	20	SDM	0.3596	X_1	0.7001	0.4890	1.4317	0.1522

Table 8 (continued)

Array	Sample size	Model	R^2	Independent variable	Regression coefficient	Standard error	t/z statistical values	P value
				X_2	− 0.2826**	0.1220	− 2.3158	0.0206
				X_3	0.1163**	0.0545	2.1330	0.0329
				WE	1.2220***	0.1776	6.8825	0.0000
				WX_1	1.0291**	0.4702	2.1884	0.0286
				WX_2	− 0.3165***	0.1179	− 2.6853	0.0072
				WX_3	0.1416***	0.0540	2.6223	0.0087
H	3	OLS	0.8754	X_1	− 0.0738	0.0590	− 1.2506	0.2161
				X_2	− 0.3265***	0.0236	− 13.8348	0.0000
				X_3	0.1174***	0.0089	13.1680	0.0000
I	15	SEM	0.6096	X_1	0.6423***	0.1132	5.6761	0.0000
				X_2	− 1.0594***	0.1149	− 9.2190	0.0000
				X_3	0.3301***	0.0367	8.9881	0.0000
				LAMBDA	0.9304***	0.3325	2.7980	0.0051
J	24	GMM	0.5432	X_1	0.5994***	0.0770	7.7841	0.0000
				X_2	− 0.8648***	0.1113	− 7.7724	0.0000
				X_3	− 0.0021***	0.0377	− 0.1133	0.0000

***, **, * correspond to significance levels of 0.01, 0.05, and 0.1, respectively. The sample observations of two types, "E" and "H," are too small. Observations are only single digits. The effect of UPAC in these two arrays failed the significance test. For E, comparing it with the other types of the same year and the same type of the other years, it is found that the failure to pass the significance test for UPAC may be due to the small sample size of observations. For H, the reason may be more of an influence on the selection of the best-fit model indicating a positive effect of UPAC on UEE. The regression coefficients of X_2 and X_3 in arrays H and I indicate that population and economic size factors significantly negatively and positively affect UEE, respectively. The regression coefficient of X_1 in group J is 0.5994 and significant, indicating a positive effect on UEE. However, the population and economic size factors have a significant adverse effect.

In conclusion, there is significant spatial and temporal heterogeneity in the effects of UPAC on UEE of different types. In 2000, UPAC significantly contributed to the UEE of advanced and lagging types. In 2010, UPAC had no significant effect on the UEE of advanced types but still significantly contributed to the lagging types. By 2020, the effects of UPAC on the UEE of both advanced and lagging types have diverged significantly.

5 Conclusion and discussion

This paper constructs a theoretical model of UPAC and UEE. The model defines the difference in the rank-size distribution of urban MP and TP as the UPAC and the fluctuation of the actual output value of the urban economy and the expected economic level based on the population size as the UEE from the perspective of the nonlinear scaling relationship between the population and the economy in the urban system. The heterogeneity of UPAC is thus considered the primary test variable to explain the differences in UEE. It forms a framework for analyzing the impact of population distribution on economic performance from the perspective of scaling law and empirically analyzing the impact of UPAC on UEE using the YRDUA as an example.

The YRDUA has a significant superlinear scaling relationship between urban population and economy. However, the agglomeration distribution pattern of UPAC becomes more and more significant. The UPAC of (mega/super) large cities mostly exceeds that of medium and small cities. The influence of UPAC on UEE gradually weakens and diverges from significant positive influence to insignificant influence in advanced cities and significant positive influence in lagging cities. It reveals that the economic development dynamics of the advanced cities have changed. It is not easy to rely on the scale effect to improve the UEE now and in the future but to seek the new dynamics of science and technology innovation or the new location effect of network development. On the other hand, the economic efficiency of lagging cities can still rely on the scale effect at this stage, which is the concern of this type of city to improve their economic efficiency.

This paper presents a theoretical framework of UEE. Compared with the traditional urban economic growth theory, the UEE theory is more applicable to explaining and analyzing the differences in the economic development of cities of different sizes. The concept of UEE puts the evaluation of urban economic growth in its urban system (urban agglomeration), compensating for the lack of comparability or significance of absolute urban growth due to city size and volume differences. It is more in line with the characteristics of significant differences in city size in China and the solid regional context in developing Chinese cities under the socialist market economy. UEE provides an alternative result different from the urban economic scale (or per capita economy) ranking. Its significance is to help more objectively understand and compare the economic performance of cities of different scales and provide a basis for decision-making for coordinated urban development.

UEE is a regional average measure of urban economic performance from a scaling-law perspective. From a single-sample perspective, the results of this paper are counterintuitive and differ from related research results (Ren et al., 2019). For example, Shanghai is a metropolitan city with a UEE of about 0.1326 in 2000, -0.3424 in 2010, and -0.0478 in 2020. Why is Shanghai's economic efficiency so low, even negative? Nanjing is similar, and its economic efficiency is lower than many smaller cities. The reason may be, on the one hand, that the concept of UEE, as defined in this paper, is different from the efficiency defined by the input–output concept, but the deviation between the expected economy based on the population size and the actual economy. When the deviation tends to zero, it indicates that the UEE meets the expectation. On the other hand, the economic efficiency of mega- and super-cities is lower than that of small cities, which can be found during the analysis of the influencing factors. The regression model sets population size as an endogenous variable and shows its significant adverse effect on UEE in each model. For large cities, the constraining effect of population size on UEE is more pronounced, implying a larger required economic size. It is also the scaling relationship between population and economy expressed in the premise assumption of the UEE concept. For every 1 unit of population growth, the economy should grow by 1^{β} units accordingly ($\beta > 1$, as discussed in 4.1 proof).

Under the guidance of the urban scaling relationship and the regional "endogenous" orientation, the UPAC is considered the main influencing factor of the heterogeneity of UEE. The theoretical framework of UEE and the endogenous scaling perspective of UPAC proposed in this paper can be used to study the economic growth performance of Chinese cities. Urban economic attributes vary with UPAC, but cities are more than simple population agglomeration. Rather than being a causal influence on UEE, UPAC is a proxy variable for different socioeconomic mechanisms reflected in people's co-locations and close interactions (Bettencourt et al., 2010). The empirical study's findings reveal UPAC's impact on UEE and its changes in the YRDUA, providing decision support for urban population management and agglomeration policies.

The results of this paper show that high UEE indicates that cities are performing above expectations for their current urban population size and that there is still room for future economic development. These are primarily small cities that lag behind and still have a positive effect of UPAC on UEE. Therefore, the future development of these cities lies in attracting more people and moving from "efficiency" to "quantity." For example, Yangzhong City, which ranks second in terms of UEE in 2020, has a resident population of 315,460. Its UPAC is relatively lagging, suggesting it may be a "small but excellent" city. It should strive to absorb more people and expand its scale in the future. On the contrary, low UEE indicates that the city's economy has yet to develop to a level that matches the size of the existing urban population. These cities are mostly mega- and super-cities, and the impact of UPAC on UEE is insignificant; even too much population restricts the improvement in UEE. In the future, we should improve the efficiency of economic development from the level of technology and innovation while controlling the size of the population and transforming from "large" to "efficient." In Shanghai, for example, the city has the highest population size and a super-advanced UPAC. However, the UEE will be negative in 2010 and 2020, so the population growth should be controlled. The economic potential should be fully exploited to improve efficiency.

The findings of this paper provide a unique perspective for the study of urban economic growth in China, as well as a reference for development and population control measures in cities of different class sizes. It should be noted that the choice of cross-sectional data at decadal intervals for the analysis inevitably needs some improvement. One is that there may be a lag in the impact of UPAC on UEE, which is difficult to be reflected by the analysis of cross-sectional data. This study technically overcomes this problem by constructing endogenous and instrumental variables in the regression model and considering the lagged type to reduce the regression error caused by the cross-sectional data. The regression results also show a good fit, indicating the reasonableness of the model construction. In addition, administrative forces, public policy, and environmental factors have a non-negligible role in UEE (Luo et al., 2021; Zhang, Zhang, & Bin, 2022). However, due to the limitations of quantitative indicators and data acquisition, these factors should be included in the empirical study. Further additions should be made to enhance understanding of these factors influencing UEE and its formation mechanism.

Appendix 1

Comparison of the main parameters of the multivariate model for each array.

Array	Model	R^2	AIC	Log-L	Sargan test p value
A	OLS	0.6473	86.2560	47.1280	
	GMM	0.6274			0.2194
B	OLS	0.5510	− 81.9940	45.9970	
	GMM	0.3973	14.6190	− 3.3090	0.2350
C	OLS	0.4510	− 87.3990	47.7000	
	GMM	0.4506			0.2216
D	OLS	0.1896	1.4450	3.2780	
	GMM	0.2293			0.1224
	SLM	0.2898	− 0.4250	5.2130	
	SEM	0.2278	− 3.7100	5.8550	
	SDM	0.2970	− 0.3150	6.2900	
E	OLS	0.2898	− 122.1030	65.0510	
	GMM	0.3229			0.2481
F	OLS	0.1373	21.1880	-6.5940	
	GMM	0.1753			0.1878
	SLM	0.2219	20.5170	-5.2590	
	SEM	0.1770	16.8720	-4.4360	
	SDM	0.2896	12.3810	− 3.2560	
G	OLS	0.2716	− 45.7700	26.8850	
	GMM	0.2998			0.1275
	SLM	0.3081	− 50.1790	30.0900	
	SEM	0.2964	− 52.4970	30.2490	
	SDM	0.3596	− 60.9890	31.6400	
H	OLS	0.8754	− 378.8420	193.4210	
	GMM	0.5547			0.3050
I	OLS	0.5684	− 90.5390	49.2690	
	GMM	0.5892			0.1967
	SEM	0.6096	− 97.5980	52.7990	
J	OLS	0.5196	− 83.7430	45.8710	
	GMM	0.5432			0.2357

Appendix 1 shows the main parameters of each array under each model, and R^2 is used as the basis for judging the degree of fit; the more significant the R^2 value, the better the fit. In addition, the higher the Log-likelihood (Log-L) value and the lower the Akaike info criterion (AIC) value, the better the model fit.

The Sargan test P value judges the GMM model to determine the fitting effect of the model. The larger the P value of the Sargan test, the better, and generally, more than 0.1 can indicate that the null hypothesis that the instrumental variables are valid cannot be rejected.

Appendix 2

Regression results of each array under different models.

Array	Sample size	Model	R^2	Independent variable	Regression coefficient	Standard error	t/z statistical values	*P* value
A	18	**OLS**	0.6473	X_1	0.5994***	0.1619	3.7000	0.0005
				X_2	− 0.4431***	0.0871	− 5.0842	0.0000
				X_3	0.1931***	0.0349	5.5272	0.0000
		GMM	0.6274	X_1	0.3249*	0.1701	1.9099	0.0561
				X_2	− 0.2745***	0.0929	− 2.9566	0.0031
				X_3	0.1248***	0.0373	3.3486	0.0008
B	16	**OLS**	0.5510	X_1	1.5113***	0.3670	4.0853	0.0001
				X_2	− 1.0078***	0.1252	− 8.0460	0.0000
				X_3	0.3914***	0.0521	7.5073	0.0000
		GMM	0.3973	X_1	− 0.0425	0.5927	− 0.0718	0.9428
				X_2	− 0.3913*	0.2174	− 1.8000	0.0719
				X_3	0.1205*	0.0933	1.2915	0.0965
C	28	**OLS**	0.4510	X_1	0.5323**	0.2546	2.0910	0.0409
				X_2	− 1.8216***	0.3401	− 5.3564	0.0000
				X_3	0.5949***	0.1170	5.0865	0.0000
		GMM	0.4506	X_1	0.4759	0.3144	1.5136	0.1301
				X_2	− 1.7273***	0.4636	− 3.7260	0.0002
				X_3	0.5625***	0.1596	3.5245	0.0004
D	26	**OLS**	0.1896	X_1	− 0.0421	0.2853	− 0.1475	0.8833
				X_2	− 0.4683***	0.1311	− 3.5727	0.0007
				X_3	0.1461***	0.0403	3.6258	0.0006
		GMM	0.2293	X_1	− 0.0225	0.2767	0.0814	0.9351
				X_2	− 0.4997***	0.1297	− 3.8526	0.0001
				X_3	0.1568***	0.0401	3.9131	0.0000
		SLM	0.2898	X_1	− 0.0098	0.2670	0.0367	0.9707
				X_2	− 0.5124***	0.1247	− 4.1091	0.0000
				X_3	0.1573***	0.0384	4.0981	0.0000
				W-E	0.8411***	0.2736	3.0745	0.0021
		SEM	0.2278	X_1	-0.0260	0.2512	− 0.1034	0.9176
				X_2	− 0.4874***	0.1142	− 4.2669	0.0000
				X_3	0.1515***	0.0358	4.2313	0.0000
				LAMBDA	0.5118**	0.2250	2.2748	0.0229
		SDM	0.2970	X_1	− 0.0366	0.2581	− 0.1417	0.8874
				X_2	− 0.4646***	0.1305	− 3.5607	0.0004
				X_3	0.1377***	0.0393	3.5063	0.0005
				WE	1.0063***	0.2001	5.0301	0.0000

Array	Sample size	Model	R^2	Independent variable	Regression coefficient	Standard error	t/z statistical values	*P* value
				WX_1	− 0.0260	0.2512	− 0.1034	0.9176
				WX_2	− 0.3874***	0.1142	− 4.2670	0.0000
				WX_3	0.0515***	0.0358	4.2313	0.0000
E	9	OLS	0.2898	X_1	0.1934	0.1555	1.2440	0.2185
				X_2	− 0.2782***	0.0835	− 3.3313	0.0015
				X_3	0.1070***	0.0342	3.1301	0.0027
		GMM	0.3229	X_1	0.1280	0.1529	0.8373	0.4024
				X_2	− 0.2388***	0.0824	− 2.8970	0.0038
				X_3	0.0907***	0.0338	2.6865	0.0072
F	27	OLS	0.1373	X_1	0.3670	0.2551	1.4387	0.1556
				X_2	0.1494	0.1369	1.0917	0.2795
				X_3	− 0.0039	0.0527	− 0.6294	0.5316
		GMM	0.1753	X_1	0.4754	0.2530	1.8789	0.0603
				X_2	0.0757	0.1376	0.5504	0.5821
				X_3	− 0.0038	0.0531	− 0.0718	0.9428
		SLM	0.2219	X_1	0.4025	0.2470	1.6295	0.1032
				X_2	0.2144	0.1494	1.4348	0.1513
				X_3	− 0.0517	0.0563	− 0.9192	0.3580
				WE	− 0.7195	0.3926	− 1.8329	0.0668
		SEM	0.1770	X_1	0.4582	0.2527	1.8136	0.0697
				X_2	0.1007	0.1340	0.7518	0.4522
				X_3	− 0.0119	0.0529	− 0.2243	0.8225
				LAMBDA	− 0.4854	0.2131	− 2.2781	0.0227
		SDM	0.2896	X_1	0.4143**	0.2211	1.8743	0.0609
				X_2	0.2037	0.1355	1.5037	0.1326
				X_3	− 0.0478	0.0493	− 0.9684	0.3328
				WE	− 1.0561**	0.3922	− 0.6505	0.0071
				WX_1	0.4582**	0.2527	1.8136	0.0697
				WX_2	0.1007	0.1340	0.7518	0.4522
				WX_3	− 0.0119	0.0529	− 0.2243	0.8225
G	20	OLS	0.2716	X_1	2.1817***	0.4792	4.5527	0.0000
				X_2	− 0.5852***	0.1215	− 4.8160	0.0000
				X_3	0.2727***	0.0548	4.9766	0.0000
		GMM	0.2998	X_1	1.1644**	0.5129	2.2702	0.0232
				X_2	− 0.3108**	0.1310	− 2.3729	0.0176
				X_3	0.1463**	0.0592	2.4698	0.0135
		SLM	0.3081	X_1	0.9700*	0.5104	1.8997	0.0575
				X_2	− 0.3908***	0.1248	− 3.1301	0.0017
				X_3	0.1606***	0.0570	2.8196	0.0048
				WE	1.1412***	0.2829	4.0333	0.0000
		SEM	0.2964	X_1	1.0291***	0.4702	2.1884	0.0286
				X_2	− 0.3165***	0.1179	− 2.6853	0.0072
				X_3	0.1416***	0.0540	2.6223	0.0087
				LAMBDA	0.5099***	0.1659	3.0738	0.0021
		SDM	0.3596	X_1	0.7001	0.4890	1.4317	0.1522

Array	Sample size	Model	R^2	Independent variable	Regression coefficient	Standard error	t/z statistical values	*P* value
				X_2	− 0.2826**	0.1220	− 2.3158	0.0206
				X_3	0.1163**	0.0545	2.1330	0.0329
				WE	1.2220***	0.1776	6.8825	0.0000
				WX_1	1.0291**	0.4702	2.1884	0.0286
				WX_2	− 0.3165***	0.1179	− 2.6853	0.0072
				WX_3	0.1416***	0.0540	2.6223	0.0087
H	3	**OLS**	0.8754	X_1	− 0.0738	0.0590	− 1.2506	0.2161
				X_2	− 0.3265***	0.0236	− 13.8348	0.0000
				X_3	0.1174***	0.0089	13.1680	0.0000
		GMM	0.5547	X_1	0.1171	0.3139	0.3730	0.7091
				X_2	− 0.1607***	0.0500	− 3.2160	0.0013
				X_3	0.0685***	0.0121	5.6651	0.0000
I	15	OLS	0.5684	X_1	0.6448***	0.1232	5.2348	0.0000
				X_2	− 1.0722***	0.1182	− 9.0727	0.0000
				X_3	0.3338***	0.0372	8.9684	0.0000
		GMM	0.5892	X_1	0.5509***	0.1243	4.4335	0.0000
				X_2	− 0.9417***	0.1231	− 7.6472	0.0000
				X_3	0.2918***	0.0389	7.5013	0.0000
		SEM	0.6096	X_1	0.6423***	0.1132	5.6761	0.0000
				X_2	− 1.0594***	0.1149	− 9.2190	0.0000
				X_3	0.3301***	0.0367	8.9881	0.0000
				LAMBDA	0.9304***	0.3325	2.7980	0.0051
J	24	OLS	0.5196	X_1	0.5866***	0.0761	7.7100	0.0000
				X_2	− 0.8436***	0.1082	− 7.7948	0.0000
				X_3	0.2892***	0.0366	7.8973	0.0000
		GMM	0.5432	X_1	0.5994***	0.0770	7.7841	0.0000
				X_2	− 0.8648***	0.1113	− 7.7724	0.0000
				X_3	− 0.0021***	0.0377	− 0.1133	0.0000

Robustness tests are usually conducted using variable substitution, variable supplementation, changing the sample size, or changing the measurement method. In this paper, endogenous variable supplementation and changing measures are used to test the robustness of the research models. Appendix 2 shows the regression results of each model for all the array groups. The results show that although the regression coefficients of the UPAC for each array fluctuate under different models, they are stable at a certain level in each array, indicating the robustness of the model. The model shown in bold is used in the results analysis section.

Author contributions JL and SY contributed to conceptualization; JL contributed to methodology, formal analysis and investigation, writing—original draft preparation, and review; YL, YZ contributed to material preparation; SY and AC contributed to funding acquisition.

Funding This research was supported by the National Natural Science Foundation of China (No. 42171200, No. 52078237); Graduate Research and Innovation Projects of Jiangsu Province (KYCX21_1284).

Data availability The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors declare that they have no competing interests.

Consent to participate Not applicable.

Consent for publication Not applicable.

Ethics approval Not applicable.

References

Ahani, S., & Dadashpoor, H. (2021). Urban growth containment policies for the guidance and control of peri-urbanization: A review and proposed framework. *Environment, Development, and Sustainability, 23*(10), 14215–14244. https://doi.org/10.1007/s10668-021-01268-5

Bettencourt, L. M. A., Lobo, J., Strumsky, D., et al. (2010). Urban scaling and its deviations: Revealing the structure of wealth, innovation and crime across cities. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0013541

Bettencourt, L. M. A. (2013). The origins of scaling in cities. *Science, 340*(6139), 1438–1441. https://doi.org/10.1126/science.1235823

Bettencourt, L. M. A., Lobo, J., & West, G. B. (2008). Why are large cities faster? Universal scaling and self-similarity in urban organization and dynamics. *European Physical Journal B, 63*(3), 285–293. https://doi.org/10.1140/epjb/e2008-00250-6

Brinkman, J. C. (2016). Congestion, agglomeration, and the structure of cities. *Journal of Urban Economics, 94*, 13–31. https://doi.org/10.1016/j.jue.2016.05.002

Bruelhart, M., & Mathys, N. A. (2008). Sectoral agglomeration economies in a panel of European regions. *Regional Science and Urban Economics, 38*(4), 348–362. https://doi.org/10.1016/j.regsciurbeco.2008.03.003

Bruelhart, M., & Sbergami, F. (2009). Agglomeration and growth: Cross-country evidence. *Journal of Urban Economics, 65*(1), 48–63. https://doi.org/10.1016/j.jue.2008.08.003

Castells-Quintana, D., & Royuela, V. (2014). Agglomeration, inequality and economic growth. *Annals of Regional Science, 52*(2), 343–366. https://doi.org/10.1007/s00168-014-0589-1

Chang, C., Chen, K. J., Waggoner, D. F., et al. (2016). Trends and cycles in China's macroeconomy. *Nber Macroeconomics Annual, 30*(1), 184.

Chen, Y. G. (2012). Zipf's law, hierarchical structure, and cards-shuffling model for urban development. *Discrete Dynamics in Nature and Society*. https://doi.org/10.1155/2012/480196

Chen, Y. G., & Jiang, B. (2018). Hierarchical scaling in systems of natural cities. *Entropy*. https://doi.org/10.3390/e20060432

Chen, A. P., & Partridge, M. D. (2013). When are cities engines of growth in China? Spread and backwash effects across the urban hierarchy. *Regional Studies, 47*(8), 1313–1331. https://doi.org/10.1080/00343404.2011.589831

Chen, M. X., Ye, C., Lu, D., et al. (2019). Cognition and construction of the theoretical connotations of new urbanization with Chinese characteristics. *Journal of Geographical Sciences, 29*(10), 1681–1698. https://doi.org/10.1007/s11442-019-1685-z

Chen, Y. G. (2010). Characterizing growth and form of fractal cities with allometric scaling exponents. *Discrete Dynamics in Nature and Society*. https://doi.org/10.1155/2010/194715

Cheshire, P. C., & Malecki, E. J. (2004). Growth, development, and innovation: A look backward and forward. *Regional Science, 83*(1), 249–267. https://doi.org/10.1007/s10110-003-0185-8

Chung, S. H., Zhang, M., & Partridge, M. D. (2020). Positive feedback in skill aggregation across Chinese cities. *Regional Studies, 54*(11), 1596–1611. https://doi.org/10.1080/00343404.2020.1743823
Ciccone, A., & Hall, R. (1996). Productivity and the density of economic activity. *American Economic Review, 86*(1), 54–70.
Ciccone, A. (2002). Agglomeration effects in Europe. *European Economic Review, 46*(2), 213–227. https://doi.org/10.1016/S0014-2921(00)00099-4
Clark, T. N., Lloyd, R., Wong, K. K., et al. (2002). Amenities drive urban growth. *Journal of Urban Affairs, 24*(5), 493–515. https://doi.org/10.1111/1467-9906.00134
Czamanski, D., & Broitman, D. (2018). The life cycle of cities. *Habitat International, 72*, 100–108. https://doi.org/10.1016/j.habitatint.2016.09.002
Duranton, G., & Puga, D. (2000). Diversity and specialisation in cities: Why, where and when does it matter? *Urban Studies, 37*(3), 533–555. https://doi.org/10.1080/0042098002104
Dyos, H. J., & Warner, S. B. (1966). The growth of cities in the nineteenth century: A review of some recent writing. *Victorian Studies, 9*(3), 225–237.
Edgeworth, F. Y. (1899). Adna ferrin weber. The growth of cities in the nineteenth century. A study in statistics. *The Economic Journal*, *9*(36), 558–560. https://doi.org/10.2307/2956573
Faberman, R. J., & Freedman, M. (2016). The urban density premium across establishments. *Journal of Urban Economics, 93*, 71–84. https://doi.org/10.1016/j.jue.2016.03.006
Fang, C. L. (2019). The basic law of the formation and expansion in urban agglomerations. *Journal of Geographical Sciences, 29*(10), 1699–1712. https://doi.org/10.1007/s11442-019-1686-y
Fang, C. L., & Yu, D. (2017). Urban agglomeration: An evolving concept of an emerging phenomenon. *Landscape and Urban Planning, 162*, 126–136. https://doi.org/10.1016/j.landurbplan.2017.02.014
Fujita, M., & Krugman, P. (2004). The new economic geography: Past, present and the future. *Papers in Regional Science, 83*(1), 139–164. https://doi.org/10.1007/s10110-003-0180-0
Gabaix, X., & Ibragimov, R. (2011). Rank-1/2: A simple way to improve the OLS estimation of tail exponents. *Journal of Business & Economic Statistics, 29*(1), 24–39. https://doi.org/10.1198/jbes.2009.06157
Glaeser, E. L. (1998). Are cities dying? *Journal of Economic Perspectives, 12*(2), 139–160. https://doi.org/10.1257/jep.12.2.139
Glaeser, E. L. (2011). Cities, productivity, and quality of life. *Science, 333*(6042), 592–594. https://doi.org/10.1126/science.1209264
Glaeser, E. L. (PB-2005–1). *Smart growth: Education, skilled workers and the future of cold-weather cities*. Harvard University, Kennedy School, Policy Brief.
Guan, X., Wei, H., Lu, S., et al. (2018). Assessment on the urbanization strategy in China: Achievements, challenges and reflections. *Habitat International, 71*, 97–109. https://doi.org/10.1016/j.habitatint.2017.11.009
Guo, F. Y., Qu, X. Q., Ma, Y. Y., et al. (2021). Spatiotemporal pattern evolution and influencing factors of shrinking cities: Evidence from China. *Cities*. https://doi.org/10.1016/j.cities.2021.103391
He, C., & Pan, F. (2010). Economic transition, dynamic externalities and city-industry growth in China. *Urban Studies, 47*(1), 121–144. https://doi.org/10.1177/0042098009346865
Henderson, J. V. (2000). *The effects of urban concentration on economic growth* National Bureau of Economic Research, Inc, Issue.
Henderson, V. (2003). The urbanization process and economic growth: The so-what question. *Journal of Economic Growth, 8*(1), 47–71. https://doi.org/10.1023/A:1022860800744
Hens, L., Karintseva, O., Kharchenko, M., et al. (2018). The state's structural policy innovations influenced by the ecological transformations. *Marketing and Management of Innovations, 3*, 290–301. https://doi.org/10.21272/mmi.2018.3-26
Hui, E. C. M., Li, X., Chen, T., et al. (2020). Deciphering the spatial structure of China's megacity region: A new bay area-the Guangdong-Hong kong-Macao greater bay area in the making. *Cities*. https://doi.org/10.1016/j.cities.2018.10.011
Krugman, P. (1990). Increasing returns and economic geography. National Bureau of Economic Research Working Paper Series, No. 3275. https://doi.org/10.3386/w3275
Lei, W. Q., Jiao, L. M., Xu, G., et al. (2021). Urban scaling in rapidly urbanizing China. *Urban Studies*. https://doi.org/10.1177/00420980211017817
Li, R. Q., Dong, L., Zhang, J., et al. (2017). Simple spatial scaling rules behind complex cities. *Nature Communications*. https://doi.org/10.1038/s41467-017-01882-w
Li, S., & Sui, D. (2013). Pareto's law and sample size: A case study of China's urban system 1984–2008. *GeoJournal, 78*(4), 615–626. https://doi.org/10.1007/s10708-012-9455-9
Lin, G. C. S. (1999). State policy and spatial restructuring in post-reform China, 1978–95. *International Journal of Urban Regional Research, 23*(4), 670–696.

Lin, G. C. S. (2002). The growth and structural change of Chinese cities: A contextual and geographic analysis. *Cities, 19*(5), 299–316. https://doi.org/10.1016/S0264-2751(02)00039-2
Liu, K., Xue, Y. T., Chen, Z. F., et al. (2022). Economic spatial structure of China's urban agglomerations: Regional differences, distribution dynamics, and convergence. *Sustainable Cities and Society*. https://doi.org/10.1016/j.scs.2022.104253
Luo, D., Liang, L. W., Wang, Z. B., et al. (2021). Exploration of coupling effects in the economy-society-environment system in urban areas: Case study of the Yangtze River Delta urban agglomeration. *Ecological Indicators*. https://doi.org/10.1016/j.ecolind.2021.107858
Lyu, Y. M., & Jiang, F. Y. (2022). Spatial and temporal distribution of population in urban agglomerations changes in China. *Scientific Reports*. https://doi.org/10.1038/s41598-022-12274-6
Martin, P., & Ottaviano, G. I. P. (2001). Growth and agglomeration. *International Economic Review, 42*(4), 947–968. https://doi.org/10.1111/1468-2354.00141
Martin, R. (1999). The new "geographical turn" in economics: Some critical reflections. *Cambridge Journal of Economics, 23*(1), 65–91. https://doi.org/10.1093/cje/23.1.65
Mazek, W. F., & Chang, J. (1972). The chicken or egg fowl-up in migration: Comment. *Southern Economic Journal, 39*(1), 133–139. https://doi.org/10.2307/1056236
Molotch, H., & Logan, J. (2014). Tensions in the growth machine: Overcoming resistance to value-free development. *Social Problems, 31*(5), 483–499. https://doi.org/10.2307/800236
Monkkonen, P., Comandon, A., Escamilla, J. A. M., et al. (2018). Urban sprawl and the growing geographic scale of segregation in Mexico, 1990–2010. *Habitat International, 73*, 89–95. https://doi.org/10.1016/j.habitatint.2017.12.003
Nguyen, T. A., Le, P. M. T., Pham, T. M., et al. (2019). Toward a sustainable city of tomorrow: A hybrid Markov–cellular automata modeling for urban landscape evolution in the Hanoi city (Vietnam) during 1990–2030. *Environment, Development and Sustainability, 21*(1), 429–446. https://doi.org/10.1007/s10668-017-0046-2
Parnreiter, C. (2021). The janus-faced genius of cities. *Urban Studies*. https://doi.org/10.1177/00420980211007718
Polese, M. (2005). Cities and national economic growth: A reappraisal. *Urban Studies, 42*(8), 1429–1451. https://doi.org/10.1080/00420980500150839
Ren, Y. F., Fang, C. L., Lin, X. Q., et al. (2019). Evaluation of the eco-efficiency of four major urban agglomerations in coastal eastern China. *Journal of Geographical Sciences, 29*(8), 1315–1330. https://doi.org/10.1007/s11442-019-1661-7
Rodriguez-Pose, A., & Storper, M. (2020). Housing, urban growth and inequalities: The limits to deregulation and upzoning in reducing economic and spatial inequality. *Urban Studies, 57*(2), 223–248. https://doi.org/10.1177/0042098019859458
Romer, P. M. (1994). The origins of endogenous growth. *Journal of Economic Perspectives, 8*(1), 3–22. https://doi.org/10.1257/jep.8.1.3
Rozenfeld, H. D., Rybski, D., Gabaix, X., et al. (2011). The area and population of cities: New insights from a different perspective on cities. *American Economic Review, 101*(5), 2205–2225. https://doi.org/10.1257/aer.101.5.2205
Scott, A. J. (2012). Review of triumph of the city: How our greatest invention makes us richer, smarter, greener, healthier, and happier, by e. Glaeser [Triumph of the City: How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier, Edward Glaeser]. *Economic Geography*, *88*(1), 97–100.
Sigler, T. J., & Martinus, K. (2017). Extending beyond "world cities' in world city network (WCN) research: Urban positionality and economic linkages through the Australia-based corporate network. *Environment and Planning A-Economy and Space, 49*(12), 2916–2937. https://doi.org/10.1177/0308518x16659478
Smirnov, S., Ozyildirim, A. & Picchetti, P. (2019). Introduction. In S. Smirnov et al. (Eds.), *Business cycles in* BRICS (pp. 1–6). Springer International Publishing, USA. https://doi.org/10.1007/978-3-319-90017-9_1
Stone, C. N. (1993). Urban regimes and the capacity to govern: A political economy approach. *Journal of Urban Affairs, 15*(1), 1–28. https://doi.org/10.1111/j.1467-9906.1993.tb00300.x
Storper, M., & Scott, A. J. (2009). Rethinking human capital, creativity and urban growth. *Journal of Economic Geography, 9*(2), 147–167. https://doi.org/10.1093/jeg/lbn052
Sunley, P., Martin, R., Gardiner, B., et al. (2020). In search of the skilled city: Skills and the occupational evolution of British cities. *Urban Studies, 57*(1), 109–133. https://doi.org/10.1177/0042098019834249
Taylor, P. J., & Derudder, B. (2015). *World city network: A global urban analysis*. USA: Routledge.
Tretter, E. M. (2008). Scales, regimes, and the urban governance of Glasgow. *Journal of Urban Affairs, 30*(1), 87–102. https://doi.org/10.1111/j.1467-9906.2007.00375.x

Wang, L., Xue, Y. B., Chang, M., et al. (2020). Macroeconomic determinants of high-tech migration in China: The case of Yangtze River Delta urban agglomeration. *Cities*. https://doi.org/10.1016/j.cities.2020.102888

Wang, J. F., & Hu, Y. (2012). Environmental health risk detection with geogdetector. *Environmental Modelling & Software, 33*, 114–115. https://doi.org/10.1016/j.envsoft.2012.01.015

Wang, J. F., Li, X. H., Christakos, G., et al. (2010). Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun region, China. *International Journal of Geographical Information Science, 24*(1), 107–127. https://doi.org/10.1080/13658810802443457

Wei, Y. D., Wu, Y., Liao, F. H., et al. (2020). Regional inequality, spatial polarization and place mobility in provincial China: A case study of Jiangsu province. *Applied Geography*. https://doi.org/10.1016/j.apgeog.2020.102296

Williamson, J. G. (1965). Regional inequality and the process of national development: A description of the patterns. *Economic Development and Cultural Change, 13*(42), 1–84. https://doi.org/10.1086/450136

Xu, Z., & Harriss, R. (2010). A spatial and temporal autocorrelated growth model for city rank-size distribution. *Urban Studies, 47*(2), 321–335. https://doi.org/10.1177/0042098009348326

Ye, C., Liu, Z., Cai, W., et al. (2019). Spatial production and governance of urban agglomeration in China 2000–2015: Yangtze River Delta as a case. *Sustainability, 11*(5), 1343.

Ye, C., & Liu, Z. (2020). Rural-urban co-governance: Multi-scale practice. *Science Bulletin, 65*(10), 778–780. https://doi.org/10.1016/j.scib.2020.02.021

Zhang, Z. H., Zhang, G. X., & Bin, S. (2022). The spatial impacts of air pollution and socio-economic status on public health: Empirical evidence from China. *Socio-Economic Planning Sciences*. https://doi.org/10.1016/j.seps.2021.101167

Zhang, Y., Wei, Y., & Zhang, J. (2021). Overpopulation and urban sustainable development—population carrying capacity in Shanghai based on probability-satisfaction evaluation method. *Environment, Development and Sustainability, 23*(3), 3318–3337. https://doi.org/10.1007/s10668-020-00720-2

Zhao, J., Jia, D. & Chang, W. (2019). China's economic cycles: Characteristics and determinant factors. In S. Smirnov et al. (Eds.), *Business cycles in* BRICS (pp. 175–184). Springer International Publishing. https://doi.org/10.1007/978-3-319-90017-9_9

Zheng, S. Q., & Du, R. (2020). How does urban agglomeration integration promote entrepreneurship in China? Evidence from regional human capital spillovers and market integration: *Cities*. https://doi.org/10.1016/j.cities.2019.102529