

Article

# Urban Competitiveness Measurement of Chinese Cities Based on a Structural Equation Model

Zhiyuan Yuan, Xinqi Zheng \*, Lulu Zhang and Guoliang Zhao

School of Information Engineering, China University of Geosciences, Beijing 100083, China; zyyuan@cugb.edu.cn (Z.Y.); Luluzhang@Cugb.edu.cn (L.Z.); zhaogl@cugb.edu.cn (G.Z.)

\* Correspondence: zxqsd@126.com; Tel.: +86-10-8232-2116

Academic Editor: Yehua Dennis Wei

Received: 12 January 2017; Accepted: 12 April 2017; Published: 22 April 2017

**Abstract:** In the current era, competition among countries and regions is in fact among cities. Thus, how to measure urban competitiveness precisely is a basic and important question. The two main approaches to this are comprehensive evaluation based on a set of indicators and network analysis based on inter-city relations. However, both have shortcomings. In this study, we introduced structural equation model (SEM) into urban competitiveness measurement to integrate the two approaches. We built a partial least squares structural equation model (PLS-SEM) according to the analysis of causal relationship among urban attribute indicators → urban functions → urban competitiveness → urban flow intensities. Following the processes of algorithm selection, model building, fitting and assessment, and modification in PLS-SEM modeling, we measured the urban competitiveness of Chinese cities in 2010 and analyzed its distribution quantitatively and spatially. The results revealed relationships between factors contained in the model and urban competitiveness and proved that the PLS-SEM urban competitiveness measurement approach we proposed is theoretically reliable and statistically valid.

**Keywords:** urban competitiveness; structural equation model (SEM); partial least squares (PLS); China; cities

## 1. Introduction

Cities have always been among the most important objects of geography research. In this era of globalization, the competitiveness of a city in the urban system or network decides how many resources, funds and talents it can abstract and how it should develop according to its comparative advantages [1,2]. Urban competitiveness measurement offers a basic evaluation for a city's development statue, potential and influence. There are two main approaches: one is comprehensive evaluation based on indicators reflecting cities' internal attributes, and the other is judging in a city network by network analysis.

Comprehensive evaluation based on urban attribute indicators is the traditional approach to urban competitiveness measurement. Population and GDP are the most used indicators in the early studies, and are still widely used in empirical studies. As research continues, scholars have realized that what affects cities' functioning and resource-flow control includes not only economic spatial agglomeration, but also the synthetic action of the local social, ecological, and political environment. Therefore, selecting indicators that can truly reflect urban characters and organizing an indicator system have become an important approach to the comprehensive measurement of urban competitiveness. Scholars have attempted to use different angles in indicator selection [1,3–6]. However, there are shortcomings to this approach: the selection of indicators and the decision of relationships among indicators are inevitably subjective [3].

On the other hand, evaluating urban competitiveness in an urban network has become a popular approach since Friedmann [7] proposed the world city hierarchy in 1986 and Sassen's writing on global city [8]. From this perspective, urban competitiveness is viewed as the capacity of a city to dominate and control resources. Stronger cities have bigger outer influence in both agglomeration and radiation forms, and, therefore, can attract and expose more resource flows in the urban network [9]. GaWC research team has done significant researches on urban competitiveness (mainly on economic success) in global scale by considering it as a networked phenomena and measuring the quantity and quality of the connections a city has with other (world) cities [10–12]. Other scholars have measured urban competitiveness through calculating and analyzing the network, which usually consists of cities as nodes and urban flows as connections, such as population migration [13], traffic flow [14], and economic flow [15]. Currently acknowledged urban flows representing inter-city connections mainly include population migration flow, logistics flow, fund flow, information flow, and technology flow [16]. However, this approach has defects. It is mostly used in global context studies, which are supported by relatively large-scale flow data. When the study scale is smaller, the urban network becomes more elaborate, and therefore, the requirements for the inter-city flow data are refined and the data are more difficult to acquire. Taking China's urban level as an example, because of the comprehensive traffic situation and abstraction of information flow and technology flow, it is difficult to acquire enough necessary flow data for an urban network, in which inter-city connections are measured by interactions among cities. When the necessary flow data for the whole network is difficult or even impossible to acquire, the connections a city has with other cities in a city network should be considered as the reflection of urban competitiveness rather than the cause.

Therefore, although both main approaches to urban competitiveness measurement have been developed fully in theory and practice, they still have defects. Scholars have proved there is a certain relationship between indicators presenting urban attributes and urban flows reflecting inter-city connections. According to Martin [17], for example, observers and analysts have claimed population mobility is affected by business and job opportunities, the promise of wages and fortunes, the scope for consumption, and the array of cultural and leisure amenities. At the meantime, urban competitiveness is affected by both internal and external factors, which correspond to comprehensive evaluation and urban network analysis respectively. Therefore, taking urban competitiveness as an intermediate context, we proposed hypotheses of two sets of causality relationships: one is between the internal attributes of cities presented by indicators and the results of urban competitiveness, and the other is between urban competitiveness and urban flow intensities. Based on these hypotheses, this study attempted to find a new approach that can combine the perspectives of comprehensive evaluation and urban flow calculation for urban competitiveness measurement.

Structural equation model (SEM) was first proposed in 1970 by Jöreskog [18]. It is theoretically based on mathematical statistics. It can describe and measure complex causality correlations between latent variables as well as between each latent variable and corresponding observed variables. Therefore, we introduced SEM into this research for a new approach to urban competitiveness measurement, combining and cross-corroborating the two abovementioned perspectives. This enabled us to verify our hypotheses by building a mathematical matrix SEM and undertaking statistical testing. We attempt to offer a new perspective and approach to urban competitiveness measurement that is theoretically feasible and statistically reliable.

We conducted our research on urban scale in China in 2010, and focus on municipal districts of cities, which to some extent follow the typical downtowns of Western cities. Cities in our research include the four municipalities directly under the central government (Beijing, Tianjin, Shanghai and Chongqing) and prefecture-level cities which were included in the state-generated statistics, namely 286 municipal districts of China in 2010. To investigate the urban competitiveness of these cities, we built an SEM to measure urban competitiveness and analyzed the results.

The rest of this paper is organized as follows. Section 2 outlines the data used. Section 3 presents our approach and explains the rationale of SEM as well as the model process in the order of algorithm

selection, model building, model fitting and assessment, and model modification. The measurement results are described and analyzed in Section 4. Section 5 discusses the advantages and disadvantages of this approach compared with other measurement results, and expounds the statistical reliability of the proposed approach. Section 6 concludes.

## 2. Data Resource

The data for this research include statistical data reflecting socio-economic status and urban flow intensity. Urban flow data contain all available statistical data relating to population migration flow, logistics flow, fund flow, information flow, and technology flow. The decennial population census is the only official statistical data for population migration in China. The last national population census was organized in 2010 and, as all data collected for measurement should be consistent in statistical time points, the other data should also be for 2010. Therefore, population-related data in our research were collected from the *Tabulation on the 2010 Population Census of the People's Republic of China by County* [19], while other data are from the *2011 Urban Statistical Yearbook of China* [20]. Only municipal district data are considered. All 123 indicators presenting urban attributes offered in the statistics are initially collected, while only 20 representative ones are used in the final model. The detailed steps are described in Section 3.2.2. There are 286 Chinese cities in the 2010 statistics that are included in our sample. Some cities or districts not included in the statistics are also not included in our research, such as autonomous prefectures as well as Hong Kong, Macao and Taiwan.

## 3. Model Building

### 3.1. Rationale for SEM

SEM is widely applied in business, psychology, management, and social sciences to reveal actual correlation by estimating and testing relationships between model variables [21]. SEM comprehensively combines and improves traditional statistical methods, including exploratory factor analysis, confirmatory factor analysis, path analysis, multiple regression, and variance analysis. SEM has advantages in many aspects, such as handling variables simultaneously, estimating factor structure and factor correlation simultaneously, considering measure errors of variables in model estimating, and testing goodness-of-fit of the whole model. Scholars have developed several kinds of software for model calculation.

Variables in SEM include latent and observed variables. Latent variables cannot be measured directly, but can be estimated by corresponding observed ones. SEM is formally defined by two sets of models: inner model and outer model. The outer model specifies relationships between latent variable and corresponding observed variables, while the inner model specifies relationships between latent variables. The outer model includes reflective and formative modes. The causal relationship in the reflective mode is from the latent variable to observed variables, whereas that in the formative mode is the inverse. Assessment and testing methods differ for the two modes. Figure 1 below depicts the structure and rationale of SEM, in which the exogenous outer model is formative and the endogenous outer model is reflective.

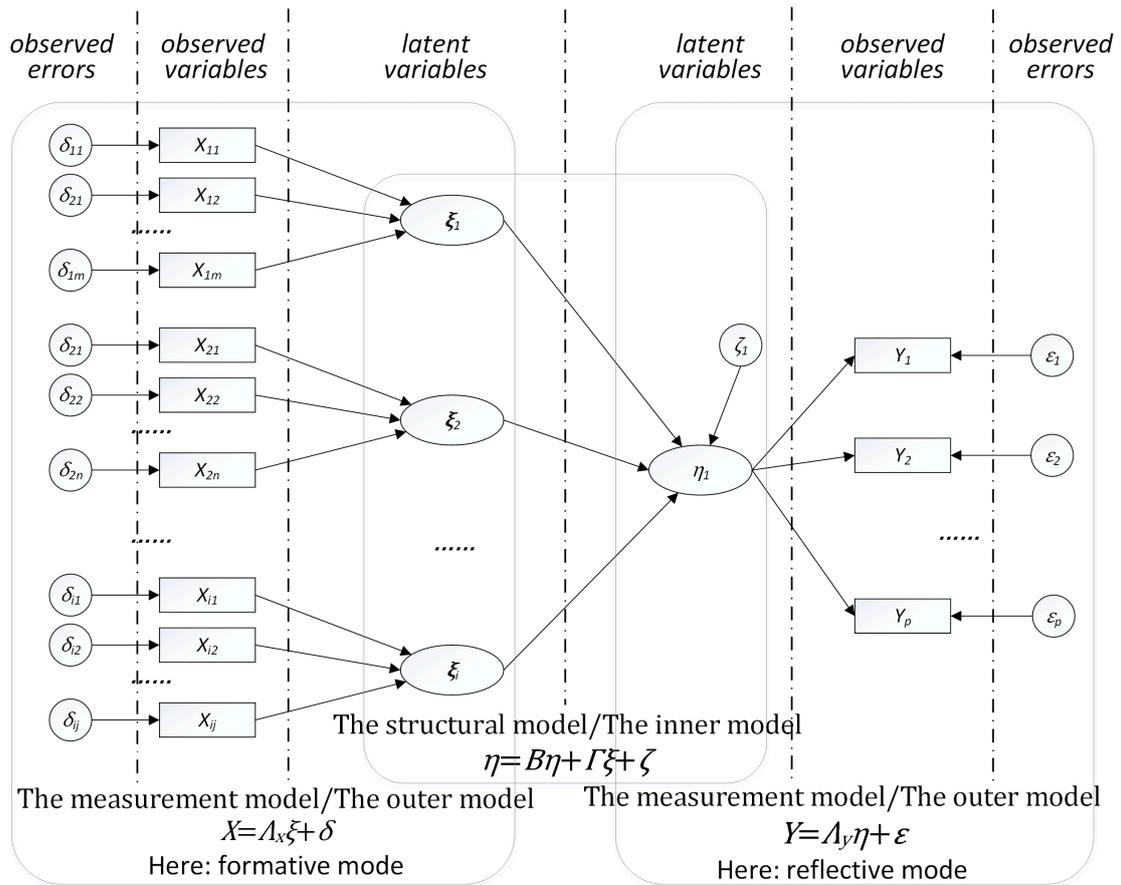
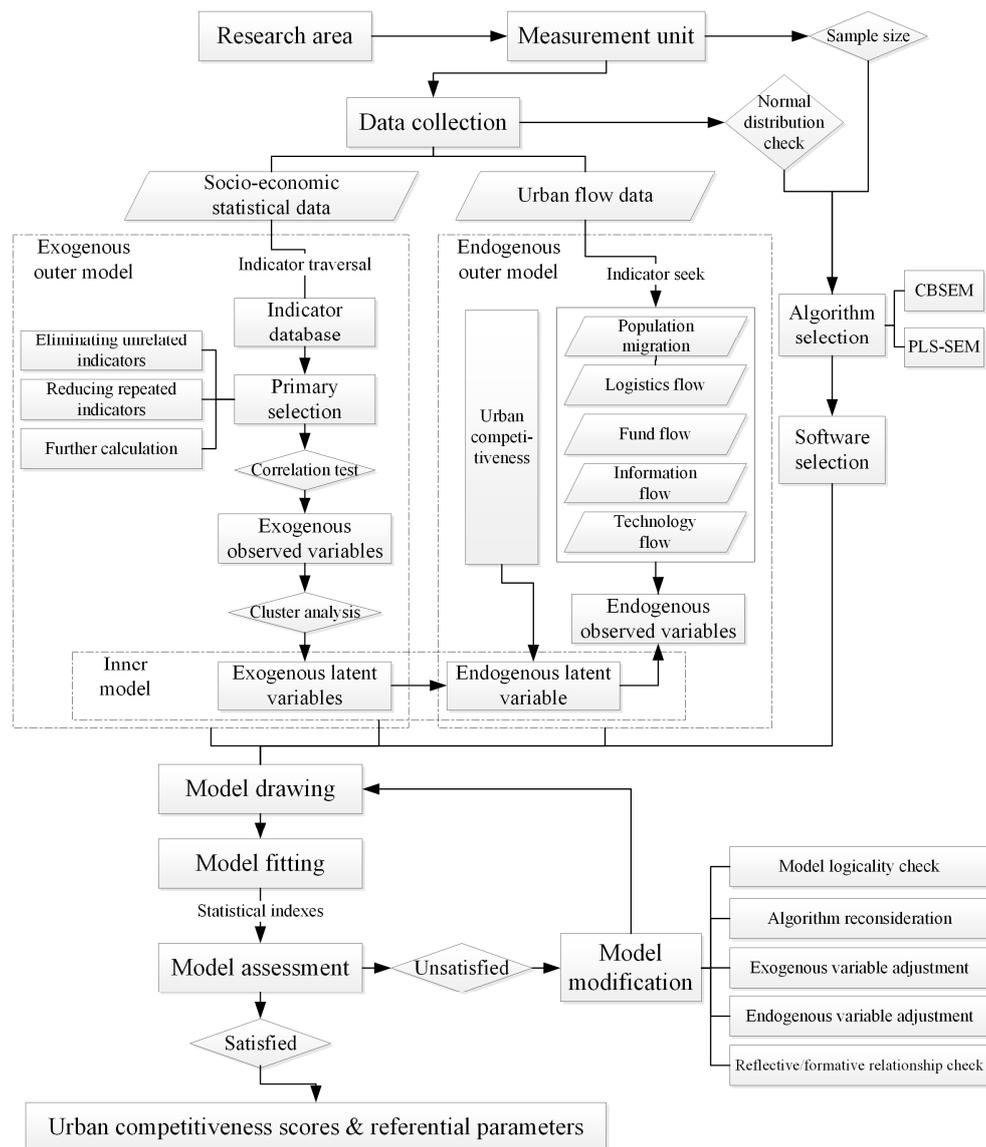


Figure 1. Structure and rationale of structural equation model (SEM).

According to algorithms of parameter estimation, SEM can be categorized into two groups: one is covariance-based SEM (CBSEM) [18], and the other is variance-based SEM, of which the most prominent example is partial least squares (PLS) path modeling [22–24]. The CBSEM focuses on the total variation of coefficients of the sample matrix and model expectation value, and has higher requirements for data distribution and sample volume. On the other hand, variance-based SEM focuses on the explanatory ability of the model for endogenous variables, but the data requirements are lower [25–27]. PLS path modeling is a prediction-oriented SEM technique based on the PLS algorithm. It includes two types according to different algorithms developed by Wold [23] and Lohmöller [28]. The calculation for both algorithms is a process of loop iteration for gradually approximating true parameter values.

### 3.2. Model Building for Urban Competitiveness Measurement

As shown in Figure 2, there are four main procedures in our approach. First, the algorithm of SEM has to be selected according to data characteristics, which determine the subsequent procedures. The second and most important part is building the model in the software tool, that is, the model design, confirmation of each part, and model drawing. Third, model fitting and assessment test statistical rationality to check whether the original model is statistically satisfied. Last, in model modification, the model is adjusted to be satisfactory. We provide a detailed explanation for each part in subsections hereafter.



**Figure 2.** Modeling procedure for urban competitiveness measurement SEM.

### 3.2.1. Algorithm Selection

First, a proper algorithm should be selected from the two families of SEM mentioned in Section 3.1 according to their suitable conditions. The sample size for our research is relatively small and not all variable data necessarily accord with normal distribution, and, thus, we selected PLS path modeling for our SEM. Meanwhile, a strength of PLS is its latent variable explanatory power, which would be helpful for revealing influences of each indicator to urban competitiveness. Based on our algorithm selection, we used SmartPLS software tool [29] for the modeling procedure, which was developed according to Lohmöller's PLS algorithm.

### 3.2.2. Model Building

#### Model Design

Our research aims to provide an SEM approach to urban competitiveness measurement that can cross-corroborate comprehensive evaluation of urban internal attributes and inter-city connections perspectives. Therefore, we considered urban competitiveness as the only endogenous latent variable

of the whole model, while integrating the comprehensive evaluation perspective into the exogenous outer model part and the inter-city connection perspective into the endogenous outer model part. For the exogenous outer model, we considered representative urban attribute indicators as observed variables and urban characteristics clustered by those indicators as latent variables, and therefore, formed an evaluating system composed of exogenous observed variables, exogenous latent variables, and the endogenous variable. For the endogenous outer model, we considered urban flow indicators reflecting inter-city connections as observed variables, which represented the correlation between urban competitiveness and urban flow.

### Exogenous Outer Model

We want to integrate the comprehensive evaluation approach into our model by the exogenous outer model part. Therefore, we need a reasonable approach to decide observed and latent variables through urban attribute indicators' processing. The indicators selected should be targeted, systematic, independent, representative, available, and suitable for the development status of the samples. They should cover all aspects of urban change fully, as they act as the causes of the exogenous latent variables, and thereby, as the causes of the endogenous latent variable (i.e., urban competitiveness). Therefore, we collected all the available official statistical indicators, selected and grouped them in the order of indicator traversal, primary selection, reselection, and cluster, and then confirmed the final components and structure of the exogenous outer model. The following provides a detailed description of the process.

- Step 1 *2011 Urban Statistical Yearbook of China* offers urban attribute indicators for our research. We collected 123 indicators for municipal districts of cities by traversing the yearbook. This is the primary indicator database for our research.
- Step 2 We considered the indicators collected one by one and made our primary selection. Indicators that obviously have nothing to do with urban competitiveness, such as number of primary schools and middle-school student enrollment, should be eliminated. Some indicators need to be calculated further as per capita data or percentiles to reflect urban competitiveness better, such as changing registered unemployed people to the unemployment rate. Moreover, when more than one indicator reflects the same or similar meaning, such as population at year-end and average population, they should be reduced to one. In total, 113 indicators were left in our research after primary selection.
- Step 3 We reselected indicators according to correlation test. This test was processed in SPSS after data standardization. Significant correlation indicators should be eliminated. There were 20 indicators left after reselection, which were defined as observed variables in our exogenous outer model.
- Step 4 We clustered the reselected indicators through principal component analysis. This can also be achieved in SPSS. The cluster results should be checked: indicators reflecting the same aspects of urban characteristics or urban competitiveness should be clustered into one dimension; the indicator number of each dimension should meet the basic requirement of SEM. Necessary replacement or adjustment on reselected indicators can be helpful for cluster outcome. Dimensions obtained from the analysis should present a certain perspective of urban characteristics corresponding to several observed indicators, thereby defining latent variables of the exogenous outer model. We identified five major dimensions of the 20 indicators, which are expressions of economic strength, living standard, space support, social security, and environmental governance.

The exogenous outer models of our SEM are of formative mode. Causal relationships from observed variables to latent variables are sufficient but not necessary. The more positively these urban attributes behave, the stronger are the latent variables, which might not be valid the other way round. The following assessment and modification procedures are decided by the model mode.

### Endogenous Outer Model

The endogenous outer model part reveals the relationship between the connections a city has with other cities in the network and urban competitiveness in our model design. Urban flow data representing inter-city connections are difficult to collect completely, especially for small-scale city networks. Therefore, urban flow data, which are not all-inclusive, are more appropriate to act as results of urban competitiveness instead of causes. The stronger are urban competitiveness, the more intense are urban flows, which might not be valid in the reverse. This causal relationship determines that the endogenous outer model is of reflective mode. Indicators reflecting all kinds of urban flows are considered observed variables of the endogenous outer model.

According to statistical data, we collected the domestic in-migratory population proportion and the out-migratory population proportion for population migration; total freight traffic and business volume of postal services for logistics flow; and number of Internet users and total business volume of telecommunication services for information flow. The model does not specify indicators for fund flow and technology flow because we could not find appropriate indicators for them. However, the reflective mode allows insufficiency of observed variables, thereby avoiding the influence of the urban flow data for the whole model.

### Inner Model

According to the two outer models, the five urban characteristics clustered from urban attribute indicators are exogenous latent variables (represented as  $\zeta_1$ – $\zeta_5$ , respectively), whereas urban competitiveness is the endogenous latent variable. In our design, exogenous latent variables act as the causes of urban competitiveness, which determines that the arrow direction in the model should be from exogenous latent variables (i.e., urban characteristics) to urban competitiveness.

We modeled our urban competitiveness measurement SEM in SmartPLS as per the following Figure 3.

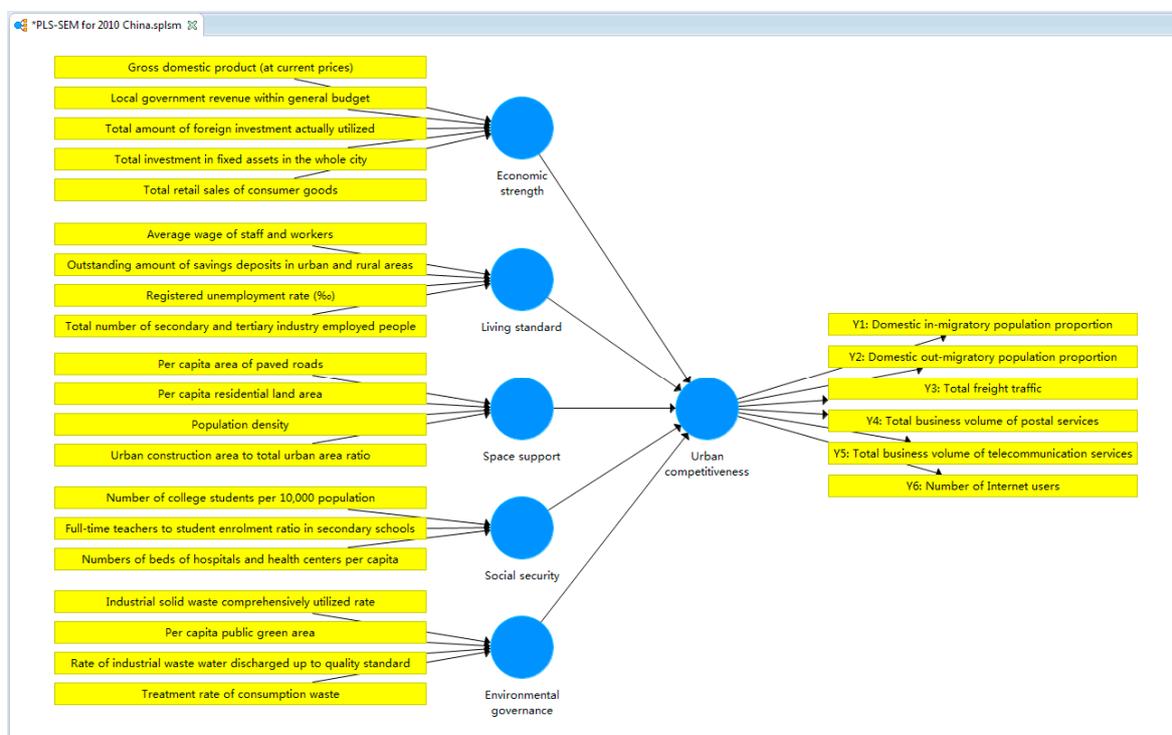


Figure 3. PLS-SEM urban competitiveness measurement for cities in 2010 China (original model).

### 3.2.3. Model Fitting and Assessment

SmartPLS contains bootstrapping and PLS algorithm modules for fitting and assessing the built model. Indexes obtained from the modules, which should be compared to the required statistical extent one by one, offer criteria for the model. If all indexes confirm the requirements, the model is satisfied statistically; otherwise, it should be modified.

PLS-SEM does not provide any global goodness-of-fit criterion. Scholars have proposed a two-step process for model assessment. The outer model should be assessed first, and then the inner model. The assessment context for each part is different, and, therefore, different indexes are employed. Meanwhile, formative and reflective modes of the outer model also consider different indexes. As mentioned in Section 3.2.2, we built an urban competitiveness measurement PLS-SEM consisting of five exogenous outer models in formative mode, an endogenous outer model in reflective mode, and an inner model between five urban characteristic aspects and urban competitiveness. Table 1 shows the criteria requirement and indexes obtained from our original model. According to the assessed indexes, our original model essentially meets the statistical requirement of PLS-SEM, except that indicator reliability of domestic in-migratory population proportion and out-migratory population proportion in the reflective endogenous outer model were lower than the required standard. Thus, their latent variable, urban competitiveness, could not explain more than 50% of the two indicators' variance in the original model, meaning the model needs to be modified.

**Table 1.** Assessment criteria and corresponding results for PLS-SEM.

| Assessed Part              | Context                                     | Criterion  | Description  | Suggest Extent  | Model Fitting Result                               |         |         |       |
|----------------------------|---|--|--|---|--|---------|---------|-------|
| Reflective outer model     | Internal consistency reliability            | Cronbach's $\alpha$ [30]   | Assume that all indicators are equally reliable, and then estimate reliability based on the indicator inter-correlations.  | >0.7 [31]   | Cronbach's $\alpha = 0.74$ , fits well.            |         |         |       |
|                            |   | Composite reliability $\rho_c$ [32]  | As above, but taking differences between indicator loadings into account.  | >0 [31]   | Composite reliability $\rho_c = 0.84$ , fits well. |         |         |       |
|                            | Indicator reliability                       | Absolute standard outer loadings   | A latent variable should explain a substantial part of each indicator's variance (usually at least 50%).   | >0.7 ( $\approx \sqrt{50\%}$ )  | Variables  | $Y_1$   | $Y_2$   | $Y_3$ |
|                            |   |  |  |   | Loading  | 0.582   | -0.194  | 0.732 |
|                            |   |  |  |   | Fitness  | Below   | Below   | Fit   |
| Variables                  | $Y_4$                                       | $Y_5$  | $Y_6$  |   |  |         |         |       |
| Loading                    | 0.917                                       | 0.906  | 0.816  |   |  |         |         |       |
| Fitness                    | Fit   | Fit  | Fit  |   |  |         |         |       |
| Convergent validity        | Average variance extracted (AVE)            | Measuring how much a latent variable is able to explain the variance of its indicators on average.   | >0.5 [33]  | AVE = 0.54, fits well.  |  |         |         |       |
| Discriminant validity      | Fornell-Larcker criterion or cross-loadings | Two conceptually different concepts should exhibit sufficient difference.  | —  | Not necessary for this part because there is only one latent variable in our model. |  |         |         |       |
| Formative outer model      | Nomological validity                        | Hypotheses check   | Assessing whether the formative index behaves within a net of hypotheses as expected, and whether those relationships between the formative index and other constructs in the path model that are sufficiently referred to in prior research are strong and significant. | Compare gradually   | Relationship                                       | Loading | Fitness |       |
|                            |   |  |  |   | $\zeta_1 \rightarrow \eta$                         | 0.74    | Fit     |       |
|                            |   |  |  |   | $\zeta_2 \rightarrow \eta$                         | 0.17    | Fit     |       |
| $\zeta_3 \rightarrow \eta$ | 0.11  | Fit  |  |   |  |         |         |       |
| $\zeta_4 \rightarrow \eta$ | 0.07  | Fit  |  |   |  |         |         |       |
| $\zeta_5 \rightarrow \eta$ | -0.03                                       | Fit  |  |   |  |         |         |       |
| External validity          | 1-Var (v)                                   | Measuring how much of the construct is not captured by any indicator by means of regressing the formative index on a reflective measure of the same construct. | >0.8 [22]  | Fits well.  |  |         |         |       |
| Multicollinearity          | Variance inflation factor (VIF)             | Assessing the degree of multicollinearity among manifest variables in a formative block.   | <10 [22]   | Each fits well.   |  |         |         |       |
| Inner model                | Determination coefficient                   | $R^2$  | Evaluating the fitting degree of the endogenous latent variables.  | >0.67 [34]  | $R^2 = 0.89$ , fits well.                          |         |         |       |
| Bootstrapping              | Weights and path coefficients               | Students' T-test   | Revealing the significance of path model relationships by creating a large, prespecified number of bootstrap samples.  | Student's <i>t</i> -distribution table  | Each fits well.                                    |         |         |       |

### 3.2.4. Model Modification

Model modification is the adjustment of the statistically unsatisfied model and includes the modification for both outer and inner models. Modification for the outer model is usually on observed variables by adding to, discarding, or replacing them. For the inner model, model modification includes altering latent variables or connecting paths between them [35]. Because the model is causally connected and acts as a whole, any changes would affect the algorithm environment and assessment result. Therefore, whatever modification method is performed, particular caution should be applied. Only one change should be made at a time, and the next change should be considered after the fitting and assessment result of the previous action. Finally, the confirmed model should be both theoretically reasonable and statistically satisfied, and can provide a strong explanation for reality.

As mentioned in Section 3.2.3, indicator reliabilities of domestic in-migratory population proportion and out-migratory population proportion in our original model were statistically unsatisfactory and needed to be modified. Considering the necessity and importance of population migration for urban flow, we attempted to find a substitution representing the population part instead of discarding these indicators. Therefore, we calculated overall population migratory proportion, which equals the in-migratory population minus the out-migratory population as a proportion of the total population, and then substituted it for the two while keeping the other parts of the model unchanged. Fitting and assessment of the modified model proved that the replacement worked well. All the indexes were statistically satisfied, which demonstrated the reasonableness of our proposal to integrate the two main perspectives for urban competitiveness measurement into one by PLS-SEM for a new approach. Thus, the modified model is able to provide proof for urban competitiveness measurement.

## 4. Results and Discussion

SmartPLS provided latent variable scores directly, as well as latent variable correlations, and loadings and weights of observed variables for corresponding latent variables. We analyzed urban competitiveness conditions in China in 2010 according to the results.

### 4.1. Urban Competitiveness Distribution Characteristics in 2010 China

#### 4.1.1. Quantitative Characteristic

According to the results, the score range of urban competitiveness in China in 2010 is 13.35–601.22. Shanghai ranks first, while Beijing, Shenzhen, Chongqing, and Tianjin rank in the top five. We analyzed the quantitative characteristic of urban competitiveness scores with clustering methodology. Cities were clustered into five classes, as shown in Table 2, and represented an uneven distribution. Shanghai is the only city in the first class, and is far ahead. The second class contains only six cities, while the score span is relatively large (Table 3). There are only about 50 cities altogether in the first three classes, accounting for just 17% of all samples. Thus, the results show that powerful cities occupy only a small part of the whole, and there are quite great gaps between most cities and the preponderant ones.

**Table 2.** Urban competitiveness cluster of 2010 China.

| Class | City Number | Accumulated Number | Score Extent  |
|-------|-------------|--------------------|---------------|
| 1     | 1           | 1                  | 601.22        |
| 2     | 6           | 7                  | 178.34–354.12 |
| 3     | 45          | 52                 | 72.56–140.50  |
| 4     | 152         | 204                | 40.71–71.54   |
| 5     | 82          | 286                | 13.35–39.88   |

Table 3. City lists of the first two clustered classes in 2010 China.

| Class | Rank | City      | Score  | Rank | City      | Score  |
|-------|------|-----------|--------|------|-----------|--------|
| 1     | 1    | Shanghai  | 601.22 |      |           |        |
| 2     | 2    | Beijing   | 354.12 | 5    | Tianjin   | 224.45 |
|       | 3    | Shenzhen  | 316.78 | 6    | Guangzhou | 209.74 |
|       | 4    | Chongqing | 231.57 | 7    | Chengdu   | 178.34 |
|       |      |           |        |      |           |        |

4.1.2. Spatial Distribution Characteristic

China is usually divided into eastern, central, western, and northeastern parts, which is influential in developing status and policymaking. The spatial distribution of the urban competitiveness shows an obvious regional difference. Thus, we mapped the spatial distribution of urban competitiveness (Figure 4) and analyzed and compared in these four regions. According to the five classes clustered in Section 4.1, we obtained the following results shown in Figure 5.

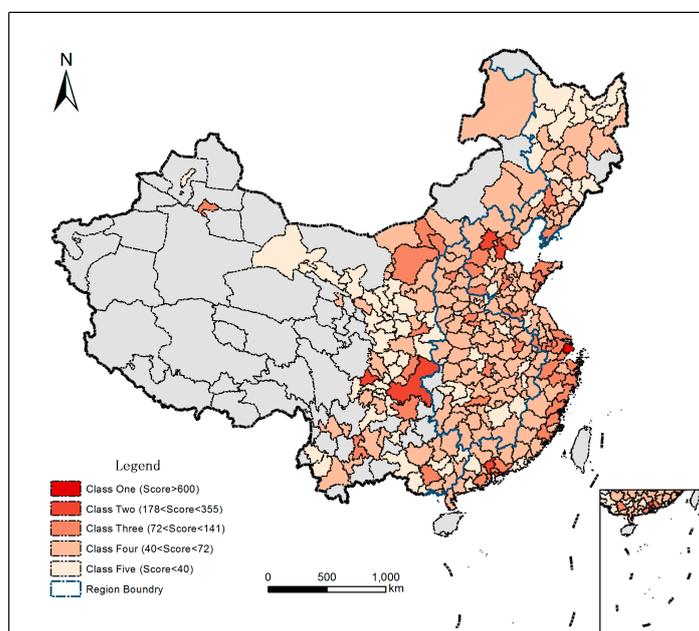


Figure 4. Spatial distribution of urban competitiveness of 2010 China.

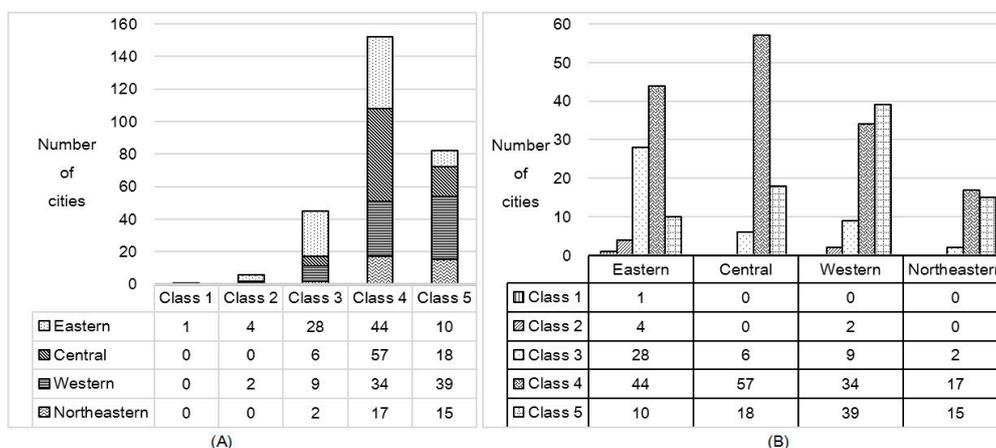
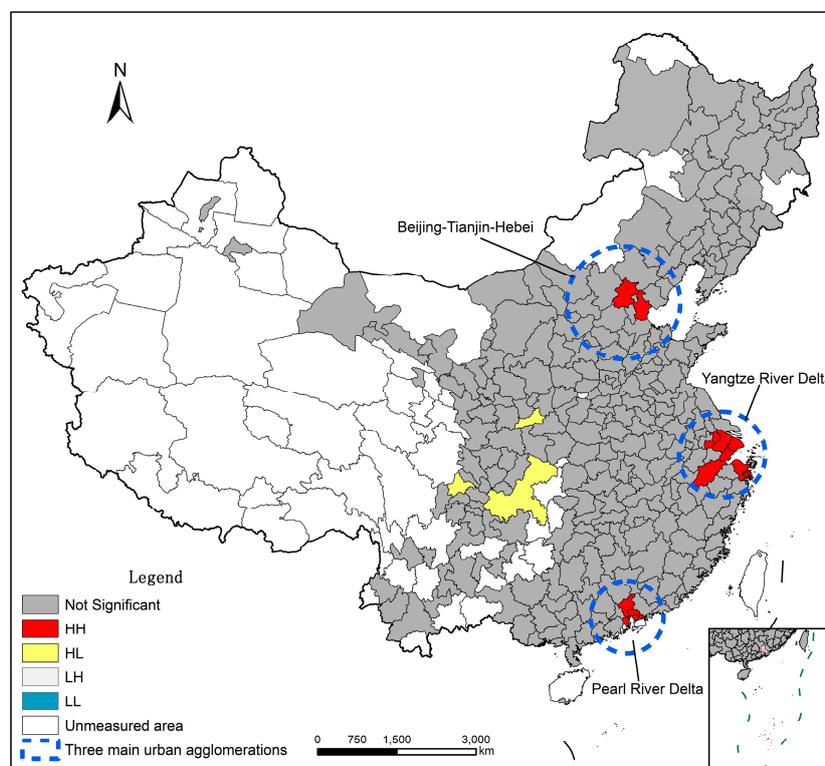


Figure 5. Distribution of clustered classes in the four regions. (A) Number of cities of each region in the five classes; (B) Number of cities of each class in the four regions.

According to statistics, eastern China is still the most prosperous region with most top three-class cities and a relatively balanced quantitative constitution compared to the national condition. The western part has always been the weakest region of China, limited by geographical and historical reasons, and remains the region with most class-five cities. However, several superior cities have been developed under the policy bias of the Grand Western Development Program, and they even occupy a notable proportion, to promote development and drive the whole region. Central and northeastern China seem to have mediocre performance, but the large proportion of common cities and the lack of superior leading ones may be a forewarning of underpowered and negative development perspectives in the two regions.

#### 4.1.3. Spatial Correlation Characteristic

Moran's I index was applied in our spatial agglomeration characteristic analysis of urban competitiveness at both global and local scales. Global Moran's I of urban competitiveness in China in 2010 was 0.11 measured by Euclidean distance, which means weak but positive spatial autocorrelation globally. Local Moran's I is helpful for recognizing urban agglomeration regions based on urban competitiveness local spatial autocorrelation (Figure 6). Three highlighted HH areas exist, which refer to a high-high urban competitiveness agglomeration: Beijing-Tianjin-Shanghai, Shanghai-Suzhou-Hangzhou-Wuxi-Ningbo-Jiaxing, and Guangzhou-Shenzhen-Zhongshan-Dongguan. In addition, three cities form HL correlation with surrounding cities, which refer to a high-low urban competitiveness agglomeration with the high score located in the center: Chongqing, Chengdu, and Xi'an. The recognized results of HH are coincident with the three most powerful urban agglomerations in China: Beijing-Tianjin-Hebei, Yangtze River Delta, and Pearl River Delta. In addition, the number of the most competitive cities fits the developing reality very well. Meanwhile, the results of HL conform quite well to the situation in western China.



**Figure 6.** Local spatial autocorrelation of urban competitiveness in China in 2010.

## 4.2. Influencing Factors of Urban Competitiveness in China in 2010

### 4.2.1. Correlation between Urban Attributes and Urban Competitiveness in China in 2010

According to the path coefficients, loadings, and weights calculated in PLS-SEM, we could analyze the influence of each variable in the model on urban competitiveness. The reliability and reasonableness of the analysis can be ensured by the aforementioned statistical assessment. We undertook our brief analysis on urban characteristically clustered latent variables (Table 4). Based on the path coefficients between exogenous latent variables and urban competitiveness in our modified model, economic strength, living standard, space support, and social security are positively correlated with urban competitiveness, whereas environmental governance have a negative impact, which reveals an unhealthy development situation that environmental treatment had a negative influence on urban competitiveness for a local government in 2010 China, especially compared with economy development. Economic strength is the most positively influencing factor, followed by living standard; the effects of the three other influencing factors are relatively weak. Further analysis can be undertaken on the contribution of the urban attribute indicators on their corresponding latent variables based on loadings or weights. We do not report this detailed analysis for space reasons, but the results are available upon request to the authors.

**Table 4.** Path coefficients of latent variables (modified model).

|                            | Variables  | Correlation | Path Coefficient |
|----------------------------|--|-------------|------------------|
| $\zeta_1 \rightarrow \eta$ | Economic strength $\rightarrow$ Urban competitiveness        | Positive    | 0.71             |
| $\zeta_2 \rightarrow \eta$ | Living standard $\rightarrow$ Urban competitiveness          | Positive    | 0.21             |
| $\zeta_3 \rightarrow \eta$ | Space support $\rightarrow$ Urban competitiveness            | Positive    | 0.06             |
| $\zeta_4 \rightarrow \eta$ | Social security $\rightarrow$ Urban competitiveness          | Positive    | 0.06             |
| $\zeta_5 \rightarrow \eta$ | Environmental governance $\rightarrow$ Urban competitiveness | Negative    | -0.02            |

### 4.2.2. Correlation between Urban Competitiveness and Urban Flows in China in 2010

Similar to Section 4.2.1, loadings or weights between urban competitiveness and observed urban flow variables reveal the correlation between them (Table 5). Urban competitiveness shows a direct and obvious influence, according to the modified model, especially for logistics flow and information flow. Population migration, however, is influenced to a limited extent. Possible reasons might be that migration decisions would relay not only on objective factors, such as urban competitiveness, but also on some more complicated ones, such as genetic relationship, family members, and personal development opportunities. The mechanism might need a new complex model for exploration.

**Table 5.** Outer loadings of urban flow variables (modified model).

| Urban Flow           |    | Variables   | Loading |
|----------------------|----|---|---------|
| Population migration | Y1 | Overall population migratory proportion             | 0.65    |
| Logistic flow        | Y2 | Total freight traffic                               | 0.74    |
|                      | Y3 | Total business volume of postal services            | 0.93    |
| Information flow     | Y4 | Total business volume of telecommunication services | 0.91    |
|                      | Y5 | Number of Internet users                            | 0.82    |

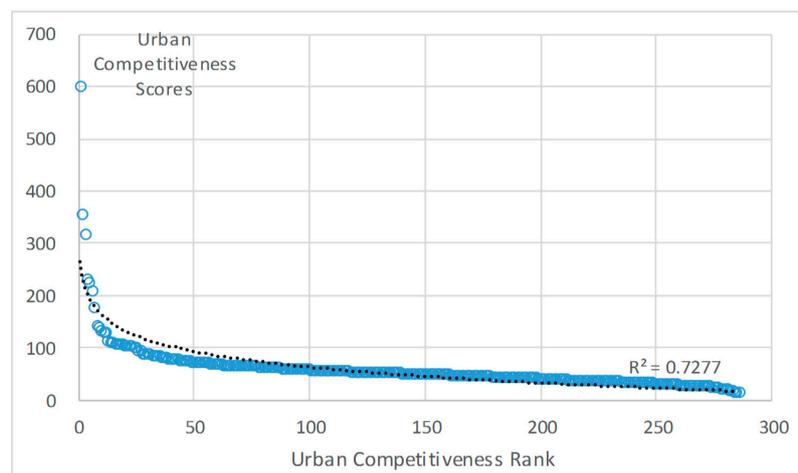
## 5. Discussion on PLS-SEM Approach

### 5.1. Reliability of Results

#### 5.1.1. Result Testing Based on Rank-Size Rule

The rank-size rule is a classical theory on measuring relationships between city size and rank in an urban system, in which city size can be understood as urban competitiveness. It was first proposed by Auerbach in 1913, and developed by scholars into more patterns later [36,37]. The rule has been proven applicable for cities in China [38–40], and, thus, we take it as a reference for testing the reliability of our result on quantitative distribution.

We fitted urban competitiveness scores and their ranks with the basic rank-size formula. The result is basically satisfied (as shown in Figure 7). Multiplicative relationship between the score and rank could be found obviously, and the  $R^2$  obtained in double-log model regression is 0.73, which demonstrates the reliability of our result on quantitative distribution.



**Figure 7.** Fitting test of PLS-SEM urban competitiveness measurement result with rank-size rule.

#### 5.1.2. Result Comparison with Other Approaches

We test the reliability of the PLS-SEM result by comparing it with traditional approaches. Chinese scholars usually take population, GDP, or area of built urban district to substitute for urban competitiveness, giving us comparable results. In addition, the *Blue Book of Urban Competitiveness*, which is annually released by the Chinese Academy of Social Sciences and measures all cities in China based on a comprehensive evaluation of substantial amount of urban attribute data, is supposed to be an authoritative version of urban competitiveness measurement.

We compared PLS-SEM urban competitiveness scores with scores from the *2010 Blue Book of Urban Competitiveness* and single index measurements on population, GDP, and area of built urban district by correlation test, and compared the top 50 and bottom 50 cities of our results with the Blue Book to observe how much they match. As shown in the correlation matrix (Table 6), our urban competitiveness scores are obviously correlated with the others. Moreover, considering the overall correlation, the result of PLS-SEM has slight advantages. On the other hand, according to the match test, the numbers of coincident cities in the top 50 and bottom 50 are 38 and 30, respectively, which are relatively high. Overall, we proved the reliability of the results measured by our PLS-SEM approach.

**Table 6.** Comparison of correlations between different results.

| Correlation Coefficient      | PLS-SEM Measurement | Competitiv-Eness in Blue Book | Average Population | GDP      | Area of Built Urban District |
|------------------------------|---------------------|-------------------------------|--------------------|----------|------------------------------|
| PLS-SEM measurement          | 1                   |                               |                    |          |                              |
| Competitiveness in Blue Book | 0.645 **            | 1                             |                    |          |                              |
| Average population           | 0.779 **            | 0.572 **                      | 1                  |          |                              |
| GDP                          | 0.913 **            | 0.687 **                      | 0.853 **           | 1        |                              |
| Area of built urban district | 0.825 **            | 0.699 **                      | 0.882 **           | 0.913 ** | 1                            |

Note: \*\* Correlation is significant at the 0.01 level (2-tailed).

The similarities prove the reliability of our approach in reflecting the urban competitiveness situation. However, our result would be more comprehensive and considerable than reliance on just a single indicator. At the same time, as we have taken inter-city connections in the urban system as verification of urban attribute indicators, and determined the weights according to data and the SEM matrix, our result might be more concise and valid than that of the *Blue Book*.

### 5.2. Theoretical Reliability

PLS-SEM itself fits the topic and our theoretical basis well, therefore, helps to make the approach theoretically reliable for urban competitiveness measurement. The theoretical basis of our approach is causal relationship hypothesis between urban attributes and urban competitiveness, as well as urban competitiveness and inter-city connections. The statistical principle of SEM supports a causal relationship, and thus, matches well our intention and offers a good framework and foundation for testing our theoretical hypothesis. Moreover, the requirement for urban flow data was reduced by their positions in the causal relationship, thereby avoiding the limitations of smaller-scale research caused by data collection.

Meanwhile, statistical indexes used in this approach testify and ensure the theoretical reliability to a large degree. They test our expectations and hypotheses proposed in the model by verifying all sorts of indexes, path coefficients, loadings, and weights required statistically. When the theoretical hypotheses are not satisfied, the modification process helps in reconsidering the model. Thus, the indexes contained in our approach offer strong support for theoretical reliability.

### 5.3. Statistical Reliability

Embedded statistical skills in PLS-SEM are one of the advantages of our approach. The final model, including all of the data and relationships, can be tested statistically in the model assessment and modification. The large amount of research on PLS-SEM by statisticians provides sufficient foundations on assessment criteria and extent, which guarantee the statistical reliability of the approach. The model can be tested statistically in different aspects according to their structure and position, as shown in Table 1.

At the same time, the influence of subjectivity has been avoided to the greatest possible extent. We have filtered each item of statistical data based on statistical analysis, such as correlation test and cluster, rather than artificial sift. Moreover, the assessment, modification, and result analysis of the approach have been performed based on objective data information instead of human opinion. Although there is still human participation in the primary selection of indicators and judgment of cluster results, the statistical data have been used and mined as much as possible.

## 6. Conclusions

The main purpose of this paper has been to propose a new perspective and approach to urban competitiveness measurement that is theoretically feasible and statistically reliable. Our approach could cross-corroborate the perspectives of measuring urban competitiveness by comprehensive evaluation of urban attributes and network analysis of inter-city connections, and covers their

shortcomings. It avoids subjective intervention by relying on data largely in indicator selection, determining weights through SEM matrix, and adding urban flow data into traditional comprehensive evaluation verification. Meanwhile, it decreases the requirement for urban flow data by taking them as reflective rather than causes of urban competitiveness, and therefore, makes it possible to measure small-scale objectives in urban flow perspectives under the limitation of data collection. The reliability of the result has been demonstrated through comparison, and the advantages of this approach have been proven, both theoretically and statistically. We consider our approach reliable for urban competitiveness measurement.

As for the 286 municipal districts of cities which were included in the state-generated statistics in China in 2010, their urban competitiveness showed an uneven distribution. Quantitatively, cities could be clustered into five classes: Shanghai is the only city in the first class and is far ahead, while, altogether, there are only about 50 cities in the first three classes with relatively large score span. Spatially, the urban competitiveness shows an obvious regional difference: eastern China is the most prosperous region with most top three-class cities and a relatively balanced quantitative constitution, while the western part has most class-five cities. It is worth noting that, although central and north-eastern China seem to have mediocre performance, the large proportion of common cities and the lack of superior leading ones may be a forewarning of underpowered and negative development perspectives. From the spatial correlation perspective, the HH regions are coincident with the three most powerful urban agglomerations in China, while the HL regions conform quite well to the situation of Western China. Thus, the government should pay more attention to the urban competitive constitution structures of each region besides the overall development status.

We built our model and measured all 286 cities with the same set of urban attributes. Considering the differences in functions and inter-city connections between different urban hierarchies, we think much more work needs to be done through building more specific models in each hierarchy with more suitable and representative urban attributes and network linkages after having a good understanding of the hierarchy. In addition, our model could be developed further if more data were collected and added. As urban attribute statistics are usually limited for small-scale cities in China, we aim to enrich our model through big data collection in future work.

**Acknowledgments:** We would like to thank Jingli Xue for her primary trial in this work. This work was supported by the National Natural Science Foundation of China (No. 41601432), the Fundamental Research Funds for the Central Universities (No. 53200859187), the Beijing Natural Science Foundation (No. 8174074), the National Key Research and Development Program of China (No. 2016YFA0602500), and the Fundamental Research Funds for the Central Universities (No.2652015339). We appreciate Editage (<http://www.editage.cn/>) for its English language editing in this paper.

**Author Contributions:** Zhiyuan Yuan and Xinqi Zheng conceived and designed this research; Zhiyuan Yuan and Lulu Zhang conducted the model building and results discussion; Zhiyuan Yuan and Guoliang Zhao performed the data collection and calculation; Zhiyuan Yuan and Xinqi Zheng contributed the comparison and discussion, and wrote the paper; and Lulu Zhang and Guoliang Zhao provided important advices for paper promotion.

**Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

## References

1. Sáez, L.; Periañez, I. Benchmarking urban competitiveness in europe to attract investment. *Cities* **2015**, *48*, 76–85. [[CrossRef](#)]
2. Shen, J.; Yang, X. Analyzing urban competitiveness changes in major chinese cities 1995–2008. *Appl. Spat. Anal. Policy* **2014**, *7*, 361–379. [[CrossRef](#)]
3. Jiang, Y.; Shen, J. Weighting for what? A comparison of two weighting methods for measuring urban competitiveness. *Habitat Int.* **2013**, *38*, 167–174. [[CrossRef](#)]
4. Ni, P.F.; Kresl, P.; Li, X.J. China urban competitiveness in industrialization: Based on the panel data of 25 cities in china from 1990 to 2009. *Urban Stud.* **2014**, *51*, 2787–2805. [[CrossRef](#)]
5. Reed, H.C. *The Pre-Eminence of International Financial Centres*; Praeger: New York, NY, USA, 1981.

6. Ding, L.; Shao, Z.F.; Zhang, H.C.; Xu, C.; Wu, D.W. A comprehensive evaluation of urban sustainable development in china based on the topsis-entropy method. *Sustainability* **2016**, *8*, 23. [[CrossRef](#)]
7. Friedmann, J. The world city hypothesis. *Dev. Change* **1986**, *17*, 69–83. [[CrossRef](#)]
8. Sassen, S. *The Global City: New York, London, Tokyo*; Princeton University Press: Princeton, NJ, USA, 1991.
9. Pal, A. The connected city: How networks are shaping the modern metropolis. *Int. J. Urban Reg. Res.* **2014**, *38*, 2.
10. Beaverstock, J.V.; Doel, M.A.; Hubbard, P.J.; Taylor, P.J. Attending to the world: Competition, cooperation and connectivity in the world city network. *Glob. Netw.* **2002**, *2*, 111–132. [[CrossRef](#)]
11. Burger, M.J.; van Oort, F.G.; Wall, R.S.; Thissen, M.J.P.M. Analysing the competitive advantage of cities in the dutch randstad by urban market overlap. In *Metropolitan Regions: Knowledge Infrastructures of the Global Economy*; Klaesson, J., Johansson, B., Karlsson, C., Eds.; Springer: Berlin, Germany, 2013; pp. 375–391.
12. Gedikli, B. *Interurban Competition within the Circumstances Created by Global Capitalism*; GaWC Research Network: Leicestershire, UK, 2001.
13. Benton-Short, L.; Price, M.D.; Friedman, S. Globalization from below: The ranking of global immigrant cities. *Int. J. Urban Reg. Res.* **2005**, *29*, 945–959. [[CrossRef](#)]
14. Mahutga, M.C.; Ma, X.; Smith, D.A.; Timberlake, M. Economic globalisation and the structure of the world city system: The case of airline passenger data. *Urban Stud.* **2010**, *47*, 1925–1947. [[CrossRef](#)]
15. Alderson, A.S.; Beckfield, J.; Sprague-Jones, J. Intercity relations and globalisation: The evolution of the global urban hierarchy, 1981–2007. *Urban Stud.* **2010**, *47*, 1899–1923. [[CrossRef](#)]
16. Bowen, J. Network change, deregulation, and access in the global airline industry. *Econ. Geogr.* **2002**, *78*, 425–439. [[CrossRef](#)]
17. Martin, R. Keys to the city: How economics, institutions, social interaction, and politics shape development. *Econ. Geogr.* **2014**, *90*, 341–344. [[CrossRef](#)]
18. Jöreskog, K.G. A general method for analysis of covariance structures. *Biometrika* **1970**, *57*, 239–251. [[CrossRef](#)]
19. Population Census Office under the State Council; Department of Population and Employment Statistics National Bureau of Statistics. *Tabulation on the 2010 Population Census of the People's Republic of China by County*; China Statistics Press: Beijing, China, 2011.
20. City Socio-economic Survey Office of National Bureau of Statistics. *2011 Urban Statistical Yearbook of China*; China Statistics Press: Beijing, China, 2012.
21. Gefen, D.; Straub, D.W.; Boudreau, M.-C. Structural equation modelling and regression: Guidelines for research practice. *Commun. Assoc. Inf. Syst.* **2000**, *4*, 1–78.
22. Henseler, J.; Ringle, C.M.; Sinkovics, R.R. The use of partial least squares path modeling in international marketing. *Adv. Int. Mark.* **2009**, *20*, 277–319.
23. Wold, H. Soft modelling: The basic design and some extensions. In *System under Indirect Observation: Causality, Structure, Prediction, Part II*; Academic Press: Cambridge, MA, USA, 1982; pp. 36–37.
24. Wold, H. Partial least squares. In *Encyclopedia of Statistical Sciences*; John Wiley & Sons, Inc.: Somerset, NJ, USA, 2004.
25. Jöreskog, K.G.; Wold, H. The ml and pls techniques for modelling with latent variables: Historical and comparative aspects. In *Systems under Indirect Observation*; North-Holland: Amsterdam, The Netherlands, 1982; Volume 1, pp. 263–270.
26. Wold, H. Model construction and evaluation when theoretical knowledge is scarce. In *Evaluation Economic Models*; Academic Press: Cambridge, MA, USA, 1980; pp. 47–74.
27. Tenenhaus, M.; Vinzi, V.E.; Chatelin, Y.-M.; Lauro, C. Pls path modeling. *Comput. Stat. Data Anal.* **2005**, *48*, 159–205. [[CrossRef](#)]
28. Lohmöller, J.-B. *Latent Variable Path Modeling with Partial Least Squares*; Physica-Verlag HD: Heidelberg, Germany, 1989.
29. Ringle, C.M.; Wende, S.; Becker, J.-M. *Smartpls 3*; SmartPLS: Hamburg, Germany, 2005; Available online: <http://www.smartpls.de> (accessed on 26 July 2014).
30. Cronbach, L. Coefficient alpha and the internal structure of tests. *Psychom* **1951**, *16*, 297–334. [[CrossRef](#)]
31. Nunnally, J. *Psychometric Theory*; McGraw-Hill: New York, NY, USA, 1967.
32. Werts, C.E.; Linn, R.L.; Jöreskog, K.G. Intraclass reliability estimates: Testing structural assumptions. *Educ. Psychol. Meas.* **1974**, *34*, 25–33. [[CrossRef](#)]

33. Fornell, C.; Larcker, D.F. Structural equation models with unobservable variables and measurement error: Algebra and statistics. *J. Mark. Res.* **1981**, *18*, 382–388. [[CrossRef](#)]
34. Chin, W.W.; Marcoulides, G.A.; Press, P. The partial least squares approach to structural equation modeling. In *Modern Methods for Business Research*; Lawrence Erlbaum Associates Publishers: Mahwah, NJ, USA, 1998; p. 323.
35. Zhao, F. Research of Customer Satisfaction Measurement Based on PLS Path Modeling. Ph.D. Thesis, Tianjin University, Tianjin, China, 2010.
36. Rosen, K.T.; Resnick, M. The size distribution of cities: An examination of the pareto law and primacy. *J. Urban Econ.* **1980**, *8*, 165–186. [[CrossRef](#)]
37. Singer, H.W. The “courbe des populations”. A parallel to pareto’s law. *Econ. J.* **1936**, *46*, 254–263. [[CrossRef](#)]
38. Dai, H. Study on the type of urban scale distribution and its developing mechanism in China. *Hum. Geogr.* **2001**, *16*, 40–43,57.
39. Li, Z.; Yang, Y. Hierarchy flatening or enlarging: Research on the changing trend of chinese urban hieararchy based on gdp size distribution. *Plan. Stud.* **2010**, *34*, 28–32.
40. Tan, M.; Lv, C. Distrubution of china city size expressed by urban built-up area. *Acta Geogr. Sin.* **2003**, *58*, 285–293.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).