

## Article

# Research on China's Risk of Housing Price Contagion Based on Multilayer Networks

Lu Qiu <sup>\*</sup>, Rongpei Su and Zhouwei Wang

School of Finance and Business, Shanghai Normal University, Shanghai 200234, China

\* Correspondence: nuaaqiulu@shnu.edu.cn

**Abstract:** The major issue in the evolution of housing prices is risk of housing price contagion. To model this issue, we constructed housing multilayer networks using transfer entropy, generalized variance decomposition, directed minimum spanning trees, and directed planar maximally filtered graph methods, as well as China's comprehensive indices of housing price and urban real housing prices from 2012 to 2021. The results of our housing multilayer networks show that the topological indices (degree, PageRank, eigenvector, etc.) of new first-tier cities (Tianjin, Qingdao, and Shenyang) rank higher than those of conventional first-tier cities (Beijing, Shanghai, Guangzhou, and Shenzheng).

**Keywords:** housing price contagion risk; multilayer networks; transfer entropy; generalized variance decomposition

## 1. Introduction

Owing to the rapid development of China's real estate market, housing is a source of concern for ordinary residents and investors, whether for living or investment purposes. Moreover, the factors that influence and cause changes in housing prices are a hot topic. As information exchange becomes increasingly common, obtaining housing price information from various regions has become easier, potentially leading to a correlation effect between housing prices in different regions. Changes in housing prices in one region may cause changes in housing prices in other regions. Therefore, it is necessary to study the relationship between housing prices in different regions to avoid the risk of contagion between regions when housing prices in one place collapse and cause real estate in most regions to experience a downturn through the aforementioned linkage. Moreover, such investigations can improve the effectiveness of housing price control measures and provide useful advice with respect to reducing the systemic risk of real estate.

Research on the correlation of housing prices between cities can be divided into two categories: studies related to the spillover correlation of housing prices using spatial measurement methods and those that investigate the correlation between sequences using time series (Li Zheng et al., 2021) [1]. In general, for spatial econometric models, researchers use Moran's I index to test the spatial correlation between housing prices before using spatial panel models to measure the spatial spillover effect of housing prices between cities. Ding et al. (2015) [2] used spatial econometrics to compute spatial correlations between 288 cities in China at the prefecture level and above. Based on spatial adjacency, spatial distance, economic adjacency, and economic distance matrices, Chen et al. (2012) [3] investigated the regional interaction of housing prices in China. Gong et al. (2020) [4] constructed an adjacency matrix based on distances and used the spatial lag of an X model (SLX), a spatial Durbin model, and a spatial Durbin error model to assess the spillover effect of housing prices in the Chinese regions of Jiangsu, Zhejiang, Shanghai, and Anhui.

The aforementioned studies revealed a positive spatial network spillover effect related to housing prices. Time-series analysis methods are more prevalent in housing price linkage research than spatial econometric models. Early research focused primarily on



**Citation:** Qiu, L.; Su, R.; Wang, Z. Research on China's Risk of Housing Price Contagion Based on Multilayer Networks. *Entropy* **2022**, *24*, 1305. <https://doi.org/10.3390/e24091305>

Academic Editors: Xiangyun Gao, Feng An and Qingru Sun

Received: 28 June 2022

Accepted: 5 September 2022

Published: 15 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

the “ripple effect”, which posits that housing prices have the characteristics of continuous transmission in space. Using data on British urban housing prices, MacDonald et al. (1993) [5] and Alexander (1994) [6] confirmed the existence of a ripple effect. As time-series analysis has become increasingly sophisticated, various methods for analyzing the correlation between series have enriched research on housing price correlations. To test the spatial relationship of urban housing prices, researchers have used the gravity model, the Granger causality test, the cointegration test, impulse response analysis, the generalized autoregressive conditional heteroskedasticity (GARCH) model, and other methods. Wang et al. (2015) [7] used the generalized impulse response of the vector error correction model to investigate the spillover effects of housing price bubbles in Beijing, Tianjin, Shanghai, and Chongqing and discovered that the degree of spillover effects and the direction of overflow differs among the four cities. Some researchers have combined the vector autoregression (VAR) models with the generalized variance decomposition (GVD) technique to study the spillover relationship of housing prices. For example, Yang et al. (2018) [8] combined principal component analysis with the GVD of the VAR models to study the effects of housing price spillovers in 69 large- and medium-sized Chinese cities. Moreover, Lv et al. (2019) [9] investigated the spillover effect of housing prices in 35 large- and medium-sized Chinese cities using the GVD of the thick-tailed VAR models.

In addition to using the VAR model, additional research has been conducted on the level of systemic risk in the real estate industry and the associated risk contagion following the subprime mortgage crisis. Li et al. (2019) [10] used the Tail Event driven NETWORK (TENET) method to construct a systemic risk spillover network across industries. Liu et al. (2014) [11] used the AR-GARCH-CoVaR method to assess systemic risk spillovers in the real estate industry. The GARCH model has also been effectively used to evaluate the spillover contagion and correlation between housing prices and interest rates, exchange rates, and the stock market. Liu et al. (2016) [12] investigated the spillover and nonlinear correlation between housing prices, exchange rates, and stock prices using a smooth transition vector error correction GARCH (STVEC-GARCH) model. Yamaka et al. (2022) [13] investigated the nonlinear causality and dynamic correlation between exchange rates and housing prices in the boom-and-business market using the panel quantile Granger causality and dynamic conditional correlation (DCC) copula GARCH method. Some scholars have also applied the GARCH model to study the linkage effect of housing prices between regions. Zeng et al. (2015) [14] used the DCC-GARCH model to examine the relationship between housing price fluctuations in the three urban agglomerations of Beijing–Tianjin–Hebei, the Yangtze River Delta, and the Pearl River Delta. In research on the Granger causality test in housing price correlation networks, Chen et al. (2016) [15] evaluated the network structure characteristics of housing price linkage in 69 large- and medium-sized cities in China. They also extended the Granger causality test by including a nonlinear relationship to evaluate the correlation effect of housing prices in 70 large- and medium-sized Chinese cities [16].

The multilayer network method is rarely used to study the relationship between housing prices in the literature reviewed above. However, owing to the shortcomings caused by incomplete information from a single-layer network, we first constructed a multilayer network based on the comprehensive indices of housing prices (CIHP) and real housing prices of 31 major Chinese cities using the GVD and transfer entropy (TE) methods. Subsequently, we use the directed minimum spanning tree (DMST) and directed planar maximally filtered graph (DPMFG) methods to simplify each layer of the network and analyze the source and central cities in the housing price correlation network. The Multiplex Infomap method was used to conduct a community analysis of the multilayer network and study the geographical location and economic characteristics of cities with similar housing price fluctuation characteristics.

The remainder of this paper is structured as follows. In Section 2, we introduce the theory of the TE and GVD networks, the calculation of multilayer network centralities, and Multiplex Infomap. In Section 3, we introduce the datasets and the CIHP construction,

followed by presentation of the empirical results in Section 4. Finally, Section 5 comprises the conclusion and discussion.

## 2. Methodologies

### 2.1. Construction of Multilayer Networks

The TE method measures the transfer of housing price information between two cities, whereas the GVD method measures the impact of housing prices in other cities on housing prices in a given city. The correlation network created by these two methods is asymmetric, which corresponds to the asymmetric phenomenon of housing price fluctuation contagion in reality. Therefore, these two methods can describe the relationship between housing prices while also considerably reducing the bias that occurs in single-layer networks owing to the use of a multilayer network. Because the housing price correlation networks constructed by these two methods are all fully connected networks (except for the self-loop), it is difficult to grasp the key information. Consequently, we used the DMST and DPMFG methods to simplify the networks and build simplified multilayer housing price networks.

#### 2.1.1. Calculation of Transfer Entropy

According to Schreiber (2000) [17] and Chen et al. (2014) [18], we can calculate the transfer entropy from one city’s housing prices to those of another. A detailed introduction of the TE method can be found in Supplementary Materials. Based on the TE method, the incidence matrix of housing prices can be calculated as follows:

$$TE\_W = \begin{bmatrix} 0 & TE(1,2) & \cdots & TE(1,n-1) & TE(1,n) \\ TE(2,1) & 0 & \cdots & TE(2,n-1) & TE(2,n) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ TE(n-1,1) & TE(n-1,2) & \cdots & 0 & TE(n-1,n-1) \\ TE(n,1) & TE(n,2) & \cdots & TE(n,n-1) & 0 \end{bmatrix} \quad (1)$$

where the TE  $TE(i, j)$  represents the information transfer from region  $I$  to region  $J$ . The diagonal element is set to 0 because there is no information transmission between housing prices in the same city.

#### 2.1.2. Calculation of Generalized Variance Decomposition

Based on Diebold and Yilmaz (2014) [19], we can construct the contagion matrix by the method of generalized variance decomposition; further details can be found in Supplementary Materials. The elements of the GVD matrix for the H-step forecast are calculated as follows:

$$d_{ij}^H = \frac{\sigma_{ii}^{-1} \sum_{h=0}^{H-1} (e_i' A_h \Sigma e_j)^2}{\sum_{h=0}^{H-1} (e_i' A_h \Sigma A' e_i)} \quad (2)$$

where  $e_j$  is the unit vector, the element of which at  $j$  is 1 and the rest are 0;  $\Sigma$  denotes the covariance matrix of the random disturbance vector ( $\varepsilon_t$ );  $\sigma_{ii}$  is the  $\varepsilon_t$  standard deviation;  $H$  denotes the forecast period; and  $h$  is the disturbance in the moving average formula term lag order.

Because the row sum of the GVD is not always 1, the elements in the generalized variance matrix are standardized to better analyze the spillover relationship between housing prices and can be calculated as follows:

$$\tilde{d}_{ij} = \frac{d_{ij}}{\sum_{j=1}^N d_{ij}} \quad (3)$$

Subsequently, the variance decomposition matrix can be obtained as:

$$D_{ij}(h) = \begin{pmatrix} \tilde{d}_{11} & \tilde{d}_{12} & \cdots & \tilde{d}_{1N} \\ \tilde{d}_{21} & \tilde{d}_{22} & \cdots & \tilde{d}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{d}_{N1} & \tilde{d}_{N2} & \cdots & \tilde{d}_{NN} \end{pmatrix}_{N \times N} \quad (4)$$

The diagonal of the obtained GVD matrix is set to 0; only the elements related to housing price overflow are retained. The processed variance decomposition matrix is:

$$\mathbf{VD\_W} = \begin{bmatrix} 0 & \tilde{d}_{1,2} & \cdots & \tilde{d}_{1,N-1} & \tilde{d}_{1,N} \\ \tilde{d}_{2,1} & 0 & \cdots & \tilde{d}_{2,N-1} & \tilde{d}_{2,N} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \tilde{d}_{N-1,1} & \tilde{d}_{N-1,2} & \cdots & 0 & \tilde{d}_{N-1,N} \\ \tilde{d}_{N,1} & \tilde{d}_{N,2} & \cdots & \tilde{d}_{N,N-1} & 0 \end{bmatrix} \quad (5)$$

### 2.1.3. Steps of the DMST Method

The housing price correlation network built using the two aforementioned methods is fully connected and has a total of  $N(N - 1)$  edges. However, the number of samples is too large to use a fully connected network to identify and analyze important nodes. Therefore, an appropriate method for simplifying the graph is required. Because the two correlation matrices constructed herein are both directed matrices, we simplify network graphs using a DMST (Kwon et al., 2008; Qiu et al., 2020) [20,21] and a DPMFG (Ye et al., 2019) [22]. The DMST procedure is as follows.

1. Randomly select a node as the root node;
2. Travel all edges and find the smallest entry edges of all points except for the root node. Then, sum up the weighted values of edges to form the new graph. Determine the final minimum arborescence if no cycles exist in the new graph;
3. If a ring exists in the new graph, shrink the ring to a point and change the edge weight. The procedure to change edge weights is as follows:
  - (1) Choose a node ( $u$ ) in the ring and set the incoming edge of this node as  $in[u]$  and outgoing edge as  $(u, i, w)$ .  $i$  and  $w$  refer to the source node and weight, respectively;
  - (2) Set the new edge weight of node  $u$  as  $(u, i, w - in[u])$ ;
  - (1) Return to Step 2 if the new weight graph contains rings;
4. Expand the new graph if rings do not exist by the loop-breaking method (Hemminger, 1966; Gabow et al., 1986) [23,24]. The steps of the loop-breaking method are as follows:
  - (1) Find a loop in the graph;
  - (2) Remove the edge with the highest weight among the loops, but keep the graph connected;
  - (3) Repeat this process until there are no loops in the graph (but they are still connected) and obtain the minimum spanning tree.

### 2.1.4. Steps of the DPMFG Method

The network generated using the DMST method is extremely clear and easy to process, but too few edges are retained, and some critical information is frequently overlooked. To retain more edges, it is necessary to make the network as simple as possible while also ensuring that the retained information is sufficient for analysis of the characteristics of the network structure and to identify critical nodes in the network. Thus, we adopt the DPMFG method, which is similar to the DMST method. The calculation steps of the DPMFG method are as follows.

- (1) By summing symmetric elements, convert the directed network into an undirected network;
- (2) Use the undirected PMFG to simplify a fully connected network to a network with only  $3(N - 2)$  edges remaining;
- (3) Restore each edge in the simplified undirected PMFG network to two directed edges, and keep the direction and weight of the side with the highest weight value as the edge and weight of the directed PMFG. Thus, the network is simplified to a directed PMFG network.

## 2.2. Topology Calculation and Community Partitioning Method for Multilayer Networks

### 2.2.1. Centrality in Multilayer Networks

Unlike the aggregate algorithm, which simply aggregates the topology parameters from each single-layer network, the multilayer algorithm used in our study also considers connections between points from different layers. Domenico et al. (2013) [25] and Domenico et al. (2015) [26] used the tensorial formulation to calculate various multilayer network structure indicators.

When considering a multilayer network, a multilayer adjacency tensor can be used as an object for the complex relationship of the multilayer network. According to Domenico et al. (2013) [25], the multilayer network can be expressed as follows:

$$M_{\beta\delta}^{\alpha\tilde{\gamma}} = \sum_{i,j=1}^N \sum_{h,k=1}^L w_{ij}(hk) e^{\alpha}(i) e_{\beta}(j) e^{\tilde{\gamma}}(h) e_{\delta}(k) \tag{6}$$

where  $N$  represents the number of nodes,  $L$  is the number of layers in the network, and  $w_{ij}(hk)$  denotes the connection between the node  $i$  of the layer  $h$  and node  $j$  of layer  $k$ . If it is a weighted network,  $w$  represents the corresponding weight value; otherwise, it is 1. If the network is a directed network,  $w_{ij}(hk)$  and  $w_{ji}(kh)$  are not necessarily equal.  $e^{\alpha}(i)$  denotes the  $\alpha$ th component of the  $i$ th contravariant canonical vector ( $\mathbf{e}_i$ ) in  $\mathbb{R}^N$ , and  $e_{\beta}(j)$  represents the  $\beta$ th component of the  $j$ th covariant canonical vector in  $\mathbb{R}^N$ .  $e^{\tilde{\gamma}}(h)$  is the vector of the canonical basis in space  $\mathbb{R}^L$ , where the Greek index indicates the components of the vector, and the Latin index indicates the  $h$ th canonical vector. The multilayer adjacency tensor ( $M_{\beta\delta}^{\alpha\tilde{\gamma}}$ ) is a general object that can be used to represent the complicated relationships among nodes can be used to conveniently calculate the node centrality in a multilayer network.

#### (1) Degree Centrality

The degrees of nodes are determined by the sum of each layer in the aggregated multilayer network calculation. In reality, the relationship between two nodes may differ between layers. Thus, when the degrees of each layer are directly summed, errors may occur. The same issue arises in the subsequent calculations of various centrality metrics. For multilayer networks, the association between different layers and whether the nodes between the layers can be connected need to be considered comprehensively. This method directly calculates various centrality measures as a whole, which can better reflect the actual relationship. Using the previously introduced multilayer adjacency tensor, we can directly calculate the multilayer degree centrality vector using the following formula:

$$K^{\alpha} = M_{\beta\delta}^{\alpha\tilde{\gamma}} U_{\tilde{\gamma}}^{\delta} u^{\beta} \tag{7}$$

where  $U_{\tilde{\gamma}}^{\delta}$  denotes a second-order tensor, the elements of which are all equal to 1; and  $u^{\beta}$  denotes the 1-vector, the components of which are all equal to 1. After some algebra, the multi-degree centrality vector can be written as:

$$K^{\alpha} = \sum_{h,k=1}^L k^{\alpha}(\tilde{hk}) \tag{8}$$

where  $k^\alpha(\tilde{h}\tilde{k})$  denotes the degree centrality vector that reflects the connections between layers  $\tilde{h}$  and  $\tilde{k}$ . Clearly, the degree centrality vector calculated using this method differs from that calculated by aggregating the network layers into a single-layer network.

(2) Eigenvector Centrality

Eigenvector centrality is an indicator commonly used to describe the importance of a node that often has a high eigenvector centrality when its neighbors also have a high eigenvector centrality. The eigenvector centrality vector in a multilayer network can be used as the solution to the tensorial equation  $(W_j^i v_i = \lambda_1 v_j)$ , where  $\lambda_1$  is the largest eigenvalue of  $W_j^i$  (the adjacency tensor of the monoplex), and  $v_i$  denotes the eigenvector centrality of node  $i$ . In the case of multilayer networks, we can obtain the eigenvector centrality by solving the tensorial equation:

$$M_{j\beta}^{i\alpha} \Theta_{i\alpha} = \lambda_1 \Theta_{j\beta} \tag{9}$$

where  $\lambda_1$  denotes the largest eigenvalue, and  $\Theta_{i\alpha}$  is the corresponding eigentensor centrality of each node in each layer when accounting for the whole interconnected relationship. Thus, the multilayer Bonacich's eigenvector centrality is given by  $\Theta_{j\beta} = \lambda_1^{-1} M_{j\beta}^{i\alpha} \Theta_{i\alpha}$ .

(3) PageRank Centrality

PageRank centrality quantifies how well a node is connected to other nodes. In general, if a node is connected to many other nodes, its PageRank value will be higher. The PageRank calculation simultaneously considers the PageRank value of the node connected to a node; if the PageRank value of the associated node is higher, the PageRank value of the considered node will also be higher. For the multilayer PageRank centrality calculation, we can start from the steady-state solution of the equation  $p_{j\beta}(t + 1) = R_{j\beta}^{i\alpha} p_{i\alpha}(t)$  in the case of a multilayer network, where  $R_{j\beta}^{i\alpha}$  denotes the corresponding transition tensor that the walker jumps to a neighbor with rate  $r$  and teleports to any other node in the network with rate  $1 - r$ , and  $p_{i\alpha}(t)$  represents the time-dependent tensor that described the probability of finding a walker at a particular node in a particular layer. This rank-4 tensor can be written as:

$$R_{j\beta}^{i\alpha} = r T_{j\beta}^{i\alpha} + \frac{(1 - r)}{NL} u_{j\beta}^{i\alpha} \tag{10}$$

where  $T_{j\beta}^{i\alpha}$  denotes the tensor of transition probabilities for jumping between pairs of nodes and switching between pairs of layers (the calculation process is detailed in Domenico et al. (2015) [27]), and  $u_{j\beta}^{i\alpha}$  denotes the rank-4 tensor that with all components equal to 1. We can obtain the eigentensor ( $\Omega_{i\alpha}$ ) of the transition tensor ( $R_{j\beta}^{i\alpha}$ ); the multilayer PageRank is then calculated by simply contracting the layer index of the eigentensor with the 1-vector  $\omega_i = \Omega_{i\alpha} u^\alpha$ , i.e., by summing up across layers.

2.2.2. Multiplex Infomap

Infomap is a stochastic and fast algorithm used to identify the best modular description of network flows according to the map equation, which measures the length required to communicate dynamics in a network and takes advantage of the information theory duality between finding regularities in data and compressing the data. Flow dynamics in multilayer networks are more complex than in single-layer networks because a random walker moves both intralayer and interlayer, and empirical interlayer weights are usually lacking. According to Daniel Edler et al. (2017) [28] and Domenico et al. (2015) [27], the multiplex map equation can be used to reveal the community structure of multilayer networks.

We use  $W_{ij}^\beta$  to describe the intralayer adjacency of layer  $\beta$  and  $D_i^{\alpha\beta}$  to describe the interlayer adjacency of physical node  $i$ . Then, we take the random walker moves according to intralayer links with probability  $1 - \tau$ , and the random walker can move in different layers along any link of the physical node with probability  $\tau$ . Therefore, the random walker switches from layer  $\alpha$  to layer  $\beta$  with probability  $s_i^\beta / S_i^\alpha$ , where  $S_i^\alpha = \sum_\beta D_i^{\alpha\beta}$  represents the



total interlayer outlink weights of node  $i$  in layer  $\alpha$ , and  $s_i^\beta = \sum_j W_{ij}^\beta$  represents the total intralayer outlink weights of node  $i$  in layer  $\beta$ . Transition probabilities are used to describe the random walker dynamics:

$$\mathcal{P}_{ij}^{\alpha\beta}(\tau) = (1 - \tau)\delta_{\alpha\beta} \frac{W_{ij}^\beta}{s_i^\beta} + \tau \frac{W_{ij}^\beta}{S_i} \tag{11}$$

where  $\mathcal{P}_{ij}^{\alpha\beta}$  is the transition probability that node  $i$  in layer  $\alpha$  moves to node  $j$  in layer  $\beta$ ;  $\tau$  is the probability that the random walker moves between layers;  $s_i^\beta = \sum_j W_{ij}^\beta$  represents the total intralayer outlink weights in the same layer;  $S_i = \sum_\beta s_i^\beta$  represents the total interlayer outlink weights across all layers; and  $\delta_{\alpha\beta}$  is the Kronecker delta, which is 1 if  $\alpha = \beta$  and 0 otherwise.  $p_i^\alpha$  denotes the stationary distribution of state node  $i$ ,  $\alpha$ , i.e., node  $i$  in layer  $\alpha$ , and can be derived from the recursive system of equations:

$$p_i^\alpha = \sum_{j,\beta} p_j^\beta \mathcal{P}_{ji}^{\beta\alpha} \tag{12}$$

To ensure a unique ergodic solution in directed networks, we use teleportation at a rate  $\kappa$  to state nodes proportional to their intralayer in-link weights. To make the results more robust to the teleportation rate  $\kappa$ , we use unrecorded teleportation steps and recorded steps along links. First, we can obtain the recorded visit rates by calculating the stationary distribution with teleportation to state nodes proportional to their outlink weights.

$$\tilde{p}_i^\alpha = (1 - \kappa) \sum_{j,\beta} p_j^\beta \mathcal{P}_{ji}^{\beta\alpha} + \kappa \sum_{i,\alpha} \frac{S_i^\alpha}{S_i^\alpha} \tag{13}$$

We derive the recorded steps along links  $q_{ji}^{\alpha\beta}$  and nodes  $p_i^\alpha$  in next step:

$$q_{ji}^{\alpha\beta} = \tilde{p}_j^\beta \mathcal{P}_{ji}^{\beta\alpha} \tag{14}$$

$$p_i^\alpha = \sum_{j,\beta} q_{ji}^{\alpha\beta} \tag{15}$$

For directed networks, the results are robust to the variation of teleportation rate  $\kappa$  in a wide range, and the results are independent of  $\kappa$  for undirected networks (Domenico et al., 2015) [27]. In the interest of simplicity, we use a teleportation rate of  $\kappa = 0.15$  throughout our directed networks.

The map equation can express the description length based on the rates at which a random walker enters and exits modules and visits nodes within modules, which can be used to calculate the rates for a multiplex network.  $q_{\iota\curvearrowright}$  and  $q_{\curvearrowleft\iota}$  denote the transition rates at which the random walker enters and exits each module, respectively, where modules  $\iota = 1, 2, \dots, m$  are assigned from a given partition (M) of state nodes  $i, \alpha$ .

$$q_{\iota\curvearrowright} = \sum_{\{i,\alpha\} \in \iota, \{j,\beta\} \in \iota} q_{ij}^{\alpha\beta} \tag{16}$$

$$q_{\curvearrowleft\iota} = \sum_{\{i,\alpha\} \in \iota, \{j,\beta\} \in \iota} q_{ij}^{\alpha\beta} \tag{17}$$

For module codebook  $\iota$ , the physical node visit rates describe the random walker visiting each of the physical nodes in the module and can be written as:

$$p_{i \in \iota} = \sum_{\{i,\alpha\} \in \iota} p_i^\alpha \tag{18}$$

Now, we have expressed the description length of a random walker in terms of the three rates at which it enters and exits modules and visits state nodes of physical nodes within modules. Subsequently, we can obtain the two-level map equation in terms of multilayer networks with the per-step average description length ( $L(M)$ ) of the trajectory of an ergodic random walker:

$$L(M) = q_{\infty} H\left(\frac{q_{i \in \infty}}{q_{\infty}}\right) + \sum_{i=1}^m p_{i \in \infty} H\left(\frac{p_{i \in i}}{p_{i \in \infty}}\right) \tag{19}$$

$$p_{i \in \infty} = q_{i \in \infty} + p_{i \in i} \tag{20}$$

$$q_{\infty} = \sum_{i=1}^m q_{i \in \infty} \tag{21}$$

where  $H(\bullet)$  denotes the Shannon entropy. For an identically distributed random variable ( $Z$ ),  $P(z_i)$  represents the probability distribution of events ( $z_i$ ), and Shannon entropy can be described as follows:

$$H(Z) = -\sum_i P(z_i) \log_2 P(z_i) \tag{22}$$

Finally, we can use the Infomap search algorithm to acquire the optimal solution to minimize the previously introduced multiplex map equation.

### 2.3. Flowchart of the Methodologies

Sections 2.1–2.3 outline the main methods and steps used in the present study. These methods are mainly divided into two parts. In Section 2.1, we described the construction of multilayer networks, including the construction of fully connected networks (VD and TE methods) and two network-simplification methods (DMST and DPMFG methods).

In Section 2.2, we introduced the topology analysis method for multilayer networks, such as the topology calculation (PageRank, eigenvector, etc.) and community partitioning (Multiplex Infomap) methods.

The specific methodological steps can be represented by the following flowchart in Figure 1:

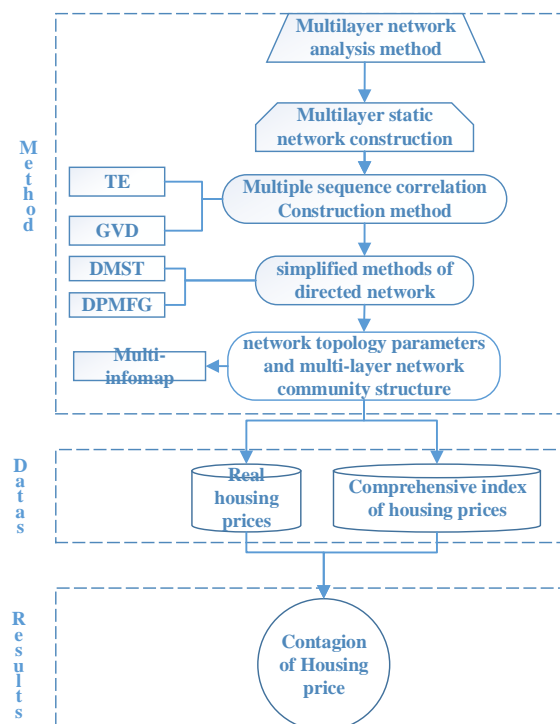


Figure 1. Empirical analysis flowchart.

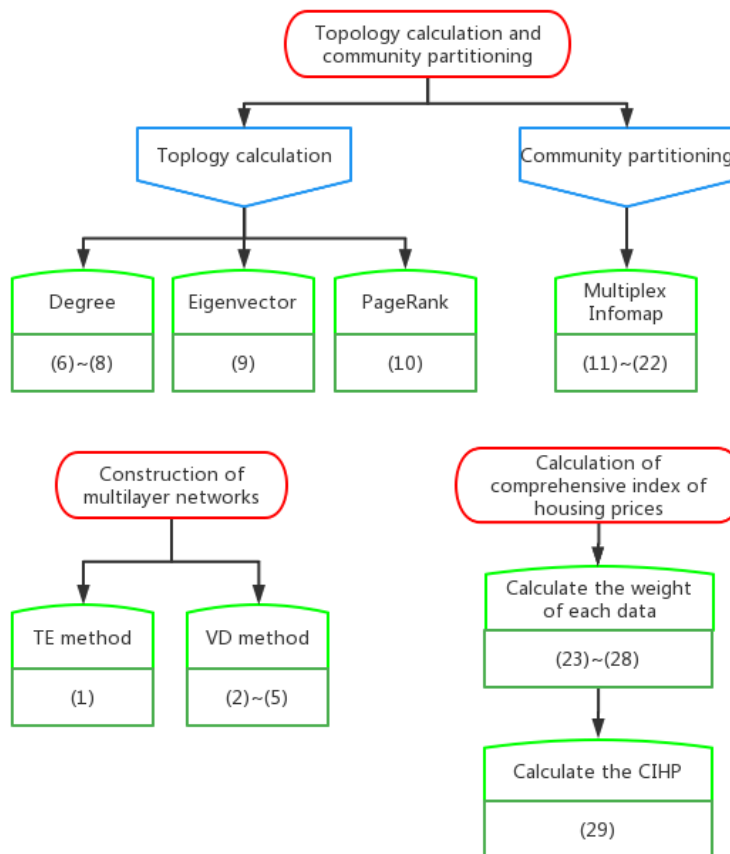


In the interest of simplicity, the formulae discussed in this article are summarized in Table 1 to clarify the calculation logic between the formulae.

**Table 1.** Compilation of the formulae discussed in this article.

| Description  | Formula Number | References  |
|--|----------------|---|
| 2.1 Construction of multilayer networks  |                |   |
| Network constructed by TE method   | (1)            | Schreiber (2000) [17], Chen et al. (2014) [18]                |
| Network constructed by VD method   | (2)~(5)        | Diebold and Yilmaz (2014) [19]                                |
| 2.2 Topology calculation and community partitioning method for multilayer networks |                |   |
| Degree centrality  | (6)~(8)        | Domenico et al. (2013) [25],<br>Domenico et al. (2015) [26]   |
| Eigenvector centrality   | (9)            |   |
| PageRank centrality  | (10)           |   |
| Multiplex Infomap  | (11)~(22)      | Domenico et al. (2015) [27], Daniel Edler et al. (2017) [28], |
| 3.1 Calculation of housing price linkage index                                     |                |   |
| Calculate the weight of each data point  | (23)~(28)      | Chen Minghua et al. (2020) [16],<br>Zhang Xiaoyan (2021) [29] |
| Calculate the comprehensive index of housing price                                 | (29)           |   |

Although the relationships between the formulae used in this article are listed in the table, these relationships are not intuitive, so we also present a diagram detailing the relationships between the formulae in Figure 2.



**Figure 2.** Diagram of the relationships between formulae.

### 3. Data and Indicator Processing

#### 3.1. Calculation of Comprehensive Index of Housing Prices

The sales price index of new commercial houses and secondhand houses is included in the housing price index data. The two indicators must be combined to create a comprehensive index. If the weight of the two indicators is arbitrarily chosen as a proportion, it is highly subjective and speculative. Thus, we employ the entropy method to objectively assign the two indicators. This method's weight value is determined by the structural characteristics of the data. The greater the dispersion of an indicator, the more information it contains and the more weight it is given.

For a city, there is an  $m \times n$  order data matrix  $X = (x_{ij})_{m \times n}$ , where  $m$  represents the length of time, and  $n$  is the number of indicators. The greater the difference between the index values of a column of data in a numerical matrix, the more information the index contains and the greater the role of the index is in the overall evaluation. For  $m$  months and  $n$  indicators of city  $r$ ,  $X_{\theta ij}$  denotes the  $j$ th indicator of city  $i$  in the  $\theta$ th year. According to Chen Minghua et al. (2020) [16] and Zhang Xiaoyan (2021) [29], the improved entropy weight method is used to calculate the weight of each indicator as follows.

- (1) Because the order of magnitude of the indicators may differ, the data must be standardized before calculating the entropy value of each indicator. The indicators selected herein are all positive; therefore, the standardization formula for each indicator is:

$$X'_{\theta ij} = \frac{X_{\theta ij} - \min(X_{\theta j})}{\max(X_{\theta j}) - \min(X_{\theta j})} \tag{23}$$

where  $\max(X_{\theta j})$  is the maximum value of indicator  $j$ , and  $\min(X_{\theta j})$  is the minimum value of indicator  $j$ ;

- (2) To avoid 0 and negative values when calculating entropy, it is necessary to add 0.1 to all values.

$$X''_{\theta ij} = X'_{\theta ij} + 0.1 \tag{24}$$

- (3) Determine the proportion of the  $j$ th index  $P_{\theta ij}$  of each city in each year:

$$P_{\theta ij} = \frac{X''_{\theta ij}}{\sum_{\theta=1}^m \sum_{i=1}^r X''_{\theta ij}} \tag{25}$$

- (4) Calculate the information entropy of the  $j$ th index. The lower the entropy value, the greater the difference between the indices. The information entropy is expressed as follows:

$$e_j = -K \sum_{\theta=1}^m \sum_{i=1}^r P_{\theta ij} \ln P_{\theta ij} \tag{26}$$

where  $K = \frac{1}{\ln(mr)}$ .

- (5) Calculate the difference coefficient of the  $j$ th index:

$$a_j = 1 - e_j \tag{27}$$

- (6) Calculate the weight of  $j$ th index:

$$g_j = \frac{a_j}{\sum_{j=1}^n a_j} \tag{28}$$

- (7) Calculate the comprehensive index of housing price (*ind*). Because there is a minimal difference in the dimension of the sales price index between new commercial houses and the secondhand houses, the original data can be used to calculate the CIHP:

$$ind_{\theta i} = \sum_{j=1}^n (g_j X_{\theta ij}) \quad (29)$$

where  $ind_{\theta i}$  denotes the CIHP of city  $i$  in the  $\theta$ th year.

### 3.2. Analysis of the Comprehensive Index of Housing Prices

We select the capital cities of China's provinces (Urumqi, Lhasa, and Xi'ning are excluded owing to a lack of data) and five well-developed cities (Shenzhen, Xiamen, Dalian, Ningbo, and Qingdao) as samples for our study. Therefore, a total of 33 cities are included for analysis; the time span runs from January 2012 to August 2021. The Wind database is used to calculate the housing price index. For subsequent research and analysis, we employ two methods to divide the selected cities. One method involves dividing them into eastern, central, and western regions according to their geographical location (eastern regions: Beijing, Tianjin, Shijiazhuang, Shenyang, Dalian, Shanghai, Nanjing, Hangzhou, Ningbo, Fuzhou, Xiamen, Jinan, Qingdao, Guangzhou, Shenzhen, Nanning, and Haikou; central regions: Taiyuan, Huhehaote, Changchun, Haerbin, Hefei, Nanchang, Zhengzhou, Wuhan, and Changsha; western regions: Chongqing, Chengdu, Guiyang, Kunming, Xi'an, Lanzhou, and Yinchuan); the method other entails dividing them into first-tier, new first-tier, second-tier, and third-tier cities according to the "Ranking of Cities' Business Attractiveness in China 2021" (<https://www.datayicai.com/readReport/267>) (accessed on 27 January 2022) reported by the Yicai website (first-tier cities: Shanghai, Beijing, Shenzhen, and Guangzhou; new first-tier cities: Chengdu, Hangzhou, Chongqing, Xi'an, Wuhan, Nanjing, Tianjin, Zhengzhou, Changsha, Ningbo, Qingdao, and Shenyang; second-tier cities: Hefei, Kunming, Xiamen, Jinan, Fuzhou, Dalian, Haerbin, Changchun, Shijiazhuang, Nanning, Guiyang, Nanchang, Taiyuan, and Lanzhou; third-tier cities: Haikou, Hohhot, and Yinchuan). Because the selected regions are all in developed urban areas, the number of third-tier cities is relatively small.

To construct the CIHP for the 33 investigated cities, we employ the entropy method and use the average value of the CIHP during the study period for sorting and analysis. Table 2 displays the results.

Table 2 shows that, except for Hefei, Xiamen, and Nanning, the top ten cities in the CIHP are all first-tier or new first-tier cities. The top four cities, in particular, are all first-tier cities, namely Shenzhen, Guangzhou, Beijing, and Shanghai. However, there is no special urban-grade difference among the lower-ranked cities. Except for the first-tier cities, the other three types of cities are represented among the 10 cities with the lowest rankings.

From the perspective of regional distribution, except for Hefei, Wuhan, and Changsha in the central part of the CIHP, the other top 10 cities are all located in the eastern part of the country. The ten lowest-ranking cities have no discernible regional distribution and are located in the eastern, central, and western regions. This may be because the cities in our sample have high levels of economic development, and there is no obvious difference between the low-ranked cities. Consequently, there is little variation in the growth rate of housing prices, and the CIHP is also considerably similar.

The city with the lowest CIHP is Haikou. Nonetheless, its CHIP score is 102.3%, demonstrating that despite minor fluctuations in the growth of housing prices in the major cities in China, they all show an upward trend in general.

**Table 2.** Cities ranked according to the comprehensive index of housing prices.

| Ranking | City         | Abbreviation | CIHP     |
|---------|--------------|--------------|----------|
| 1       | Shenzhen     | SZ           | 110.6588 |
| 2       | Guangzhou    | GZ           | 107.6885 |
| 3       | Beijing      | BJ           | 107.54   |
| 4       | Shanghai     | SH           | 107.1897 |
| 5       | Heifei       | HF           | 106.8774 |
| 6       | Xiamen       | XM           | 106.8587 |
| 7       | Nanjing      | NJ           | 106.3807 |
| 8       | Wuhan        | WH           | 105.8394 |
| 9       | Nanning      | NN           | 104.798  |
| 10      | Changsha     | CS           | 104.6763 |
| 11      | Xi'an        | XA           | 104.6693 |
| 12      | Kunming      | KM           | 104.5756 |
| 13      | Zhengzhou    | ZZ           | 104.5389 |
| 14      | Fuzhou       | FZ           | 104.5032 |
| 15      | Shenyang     | SY           | 104.4324 |
| 16      | Hangzhou     | HZ           | 104.3276 |
| 17      | Huhehaote    | HHHT         | 104.1595 |
| 18      | Shijiazhuang | SJZ          | 103.9711 |
| 19      | Yinchuan     | YC           | 103.8835 |
| 20      | Chongqing    | CQ           | 103.8156 |
| 21      | Guiyang      | GY           | 103.7865 |
| 22      | Tianjin      | TJ           | 103.7282 |
| 23      | Nanchang     | NC           | 103.7266 |
| 24      | Haerbin      | HRB          | 103.6739 |
| 25      | Dalian       | DL           | 103.6114 |
| 26      | Ningbo       | NB           | 103.4872 |
| 27      | Taiyuan      | TY           | 103.4435 |
| 28      | Chengdu      | CD           | 103.4259 |
| 29      | Jinan        | JN           | 103.4219 |
| 30      | Changchun    | CC           | 103.2341 |
| 31      | Lanzhou      | LZ           | 102.9835 |
| 32      | Qingdao      | QD           | 102.8244 |
| 33      | Haikou       | HK           | 102.3365 |

### 3.3. Selection and Source of Real Housing Price Data

The real housing price data used in our statistical analysis were collected from the Anjuke website (<https://www.anjuke.com/fangjia>) (accessed on 2 December 2021). The selected cities and time intervals are consistent with the CIHP.

## 4. Results and Analysis

Based on the aforementioned methods and data, we determine the TE calculation bins, lags, and predictive horizon of GVD to be two, one (one month), and three periods (one quarter), respectively.

Subsequently, we use the DMST and DPMFG methods to translate the TE matrix and GVD matrix, which are generated by the CIHP and real housing prices, into two different four-layer networks. The characteristics of a multilayer network topology are used to describe the influence and contagion of housing prices in various locations.

We depict the multiplex networks of CIHP and the real housing price. To visualize the node cities with more edges in the complex network more intuitively, we use the node degree (the sum of outgoing and incoming degrees) as the node size. The larger the node, the more cities with which the city's housing price is linked in the network. Generally, cities with more connecting nodes play an important role in the housing price association network. When a city's housing prices are affected, the central city is not only more vulnerable to the impact of changes in housing prices in other cities, but changes in its housing prices can also spread to others cities.

#### 4.1. Topology Analysis of Each Layer in Multilayer Complex Networks

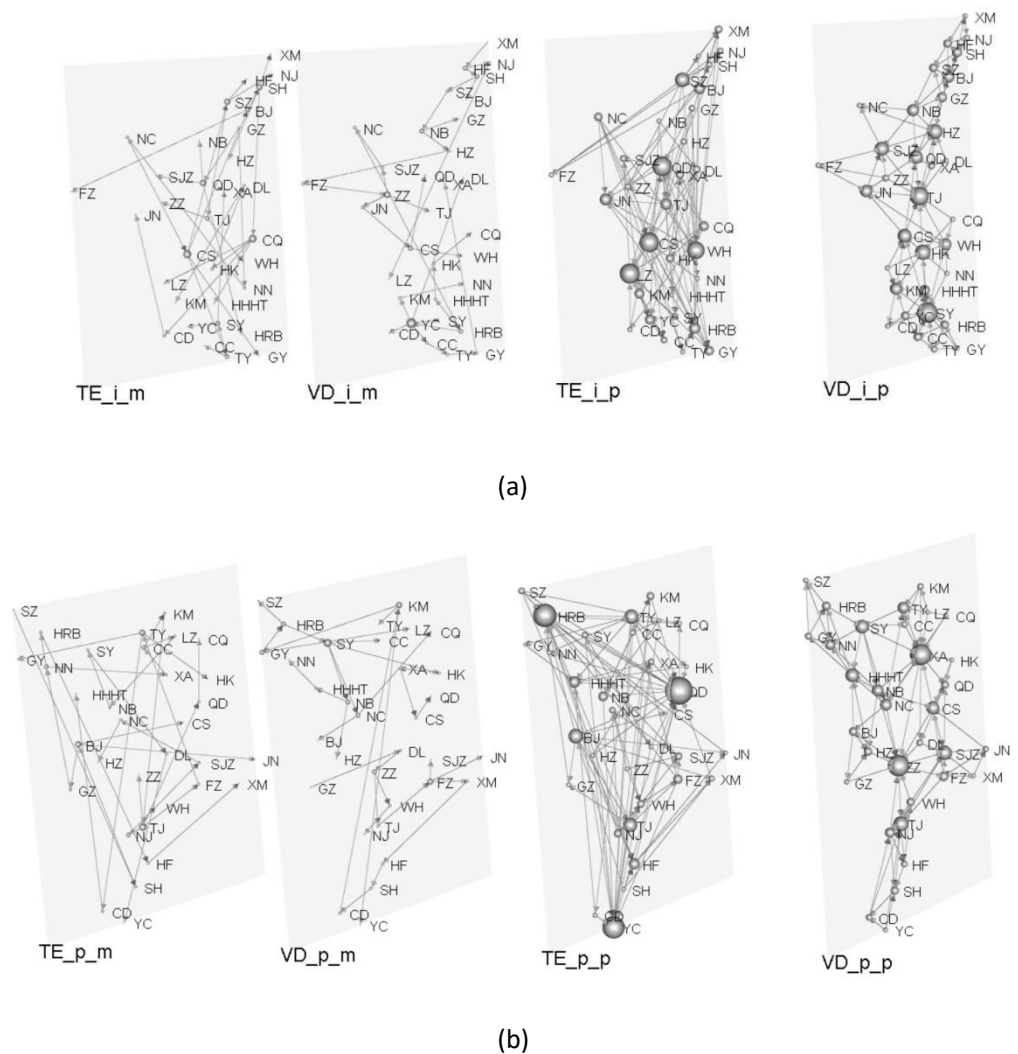
The amount of information required to analyze central cities is insufficient if the number of edges in the simplified network obtained from the DMST is small and the degree difference of each node is not large. Hence, for the DMST-simplified network, we only examine the source nodes, that is, nodes with only outgoing degrees and no incoming degrees. To a considerable extent, the source nodes can represent the source of housing price correlation, which is of significance for analyzing the contagion source of housing price changes and preventing and controlling abnormal housing price changes. In other words, we analyze the source node of the DMST-generated simplified network and the center node of the DPMFG-generated simplified network.

The multilayer network structure of the CIHP and real housing prices can be shown in Figure 3.

The source cities of the CIHP DMST network are Guangzhou and Xiamen, whereas the source cities of the DMST network of real housing prices are Shenzhen and Guangzhou. Therefore, there are three source cities: Guangzhou, Xiamen, and Shenzhen. The CIHP in the source cities is high, particularly in Shenzhen and Guangzhou, which are the top two among all investigated cities. Xiamen is ranked slightly lower, in sixth place. These results show that the source of the contagion of housing price changes is often a city that has a rapidly rising housing prices, which plays a leading role for all cities. Thus, when housing prices in such cities rise or fall, the effect is distributed to other cities, causing housing prices in other cities to change.

The four-layer price complex networks simplified by the DPMFG show that the core cities differ depending on the complex networks. In the network constructed using the CIHP and TE methods, the top cities with the highest degree of nodes are Lanzhou, Changsha, Qingdao, and Wuhan. Only Shenzhen and Shanghai have slightly larger node sizes than the first top central cities among the four traditional first-tier cities. The top central cities are replaced by Shenyang, Tianjin, Hangzhou, and Haikou in the network built using the CIHP and GVD methods. The degree of the first-tier cities is higher than that determined using the TE method, although only a medium degree, with a large gap compared to the top cities.

From the perspective of the network built using real housing prices and the TE method, Qingdao, Harbin, and Yinchuan are the most prominent cities, ranking far higher than Beijing and Taiyuan. Only Beijing ranks in the middle of the first-tier cities, with the other three cities ranking extremely low. Zhengzhou and Xi'an rank far higher than other cities in the network constructed using real housing prices and the GVD method, whereas the first-tier cities rank lower.



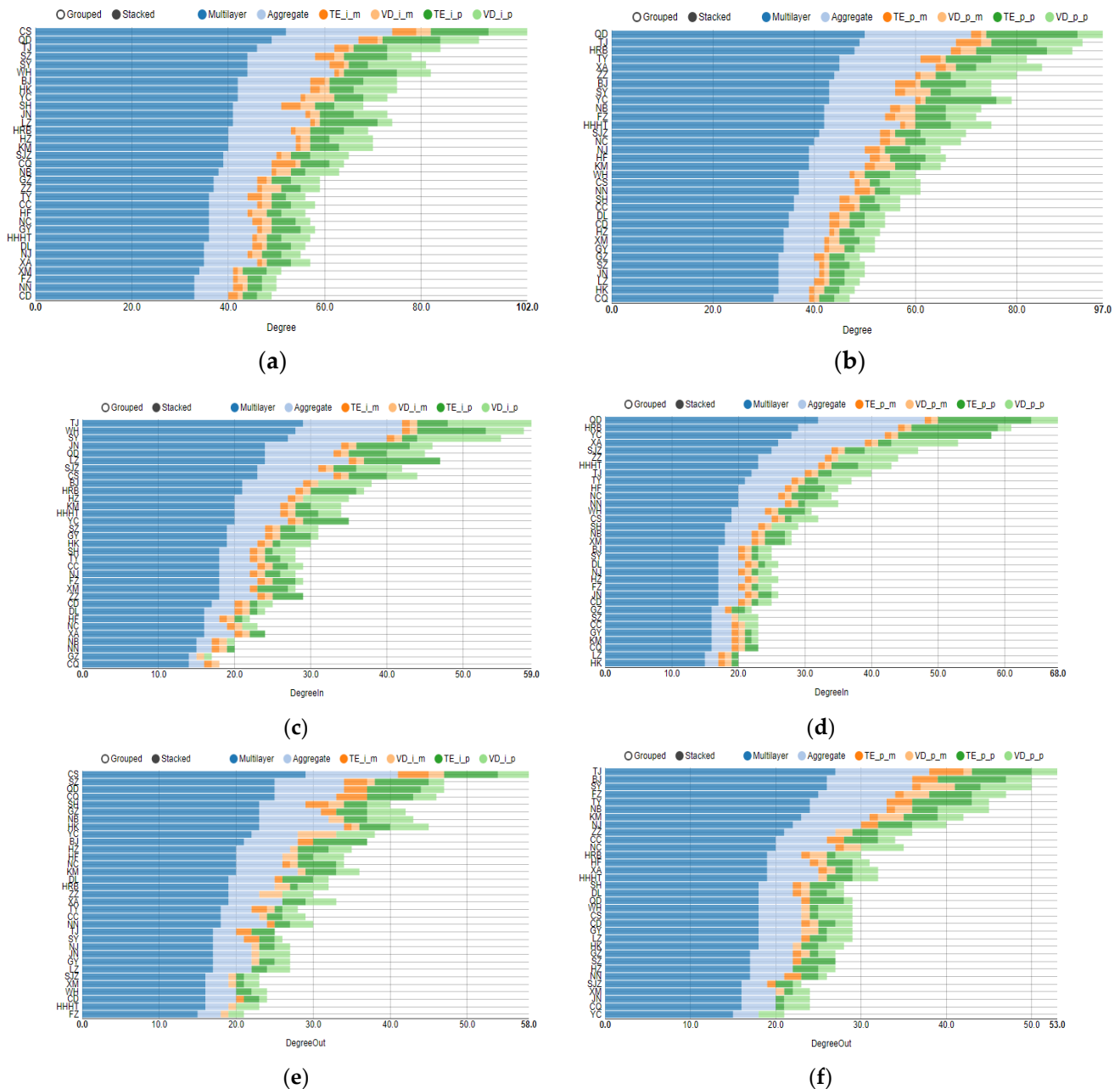
**Figure 3.** Multilayer complex network diagram of the comprehensive index of housing prices and real housing prices. **(a)** The network is built on the comprehensive index of housing prices. From left to right, the first figure depicts a network constructed using TE and DMST methods, the second depicts a network constructed using VD and DMST methods, the third depicts a network built using TE and DPMFG methods, and the fourth depicts a network constructed using VD and DPMFG methods. **(b)** Network built using real-world housing price data. The first figure depicts a network constructed using TE and DMST methods, the second depicts a network built using VD and DMST methods, the third depicts a network constructed using TE and DPMFG methods, and the fourth depicts a network built using VD and DPMFG methods.

Excluding some second-tier cities (Lanzhou and Harbin) and third-tier cities (Haikou and Yinchuan), the rest of the top-ranked cities are new first-tier cities. Qingdao, in particular, has a high degree in the two DPMFG-simplified networks built using the TE method, which is at the heart of the comparison. Conversely, Qingdao has a low degree for networks constructed using the GVD method. The traditional notion that first-tier cities should have a substantial impact on the housing prices of other cities is not reflected in the four DPMFG-simplified networks presented above. Except for Beijing and Shenzhen, which are in the middle of the network built using the TE method, the ranking in the other cases is extremely low.

The light- and dark-green stripes in Figure 4a,b represent the degree of DPMFG-simplified networks (CHPI networks and real housing price networks, respectively). Because the size of the selected node is calculated according to the degree, including both the



part entering and leaving the node, it is impossible to tell whether the third-tier city is in the center because of its high housing price overflow or because it is considerably influenced by other cities. To further analyze the spillover effect of the central city and the extent of spillover from other cities, we calculate some structural indicators of the multilayer complex network to judge whether the central city has a large spillover effect or is affected by spillover from other cities.



**Figure 4.** The multi-degree, multi-degree-in and multi-degree-out centrality of multilayer networks. (a) The multi-degree centrality of CIHP networks; (b) the multi-degree centrality of real housing price networks; (c) the multi-degree-in centrality of CIHP networks; (d) the multi-degree-in centrality of real housing price networks; (e) the multi-degree-out centrality of CIHP networks; (f) the multi-degree-out centrality of real housing price networks.

4.2. Multilayer Network Centrality Index Analysis

We calculate the multi-PageRank centrality with reference to Domenico et al. (2015) [26] and the multi-degree and multi-eigenvector according to the work of Domenico et al.



calculated using the CIHP and direct centrality calculated using real housing price data. Therefore, it can be argued that Qingdao has a prominent position with respect to the contagion of housing price changes.

Regarding the comparison of Haikou and Yinchuan presented in the previous section, the multi-centrality indicators presented in Figures 4 and 5 show that the centrality of the two cities ranks lower, possibly because they are highly ranked in some networks. Hence, to correctly evaluate whether a city is located in the center of the housing price correlation network, multiplex networks must be combined. It is inadvisable to rely solely on one or two complex networks to determine the status of a city in the correlation network.

Beijing ranked highest among the top 10 cities in terms of centrality of the four first-tier cities of Beijing, Shanghai, Guangzhou, and Shenzhen a total of six times, with the highest ranking of second in the multi-degree-out centrality of real housing prices (Figure 4f). Guangzhou only appeared in the top 10 once for the multi-degree-out centrality of the CIHP network (Figure 4e). Shanghai and Shenzhen were only in the top 10 three or four times, with a top ranking of third place. These findings also show that first-tier cities are not always at the center of the housing price correlation network. Tianjin, Qingdao, Shenyang, Beijing, and Taiyuan are the top 10 cities, with more than six occurrences; excluding Beijing and the second-tier city Taiyuan, the other three cities are new first-tier cities. These results also demonstrate that as the economy gradually develops, the dominance of the traditional first-tier cities with respect to housing price changes is gradually replaced by newly developed next-tier cities. Therefore, when controlling for the severe fluctuations in housing prices, we can focus on these new first-tier cities and conduct necessary housing price supervision to prevent rapid fluctuations in housing prices in a short period of time, which is not only conducive to the stability of housing prices in local cities but can also stabilize housing prices in other cities through the housing price association network.

Table 3 depicts the multi-centrality of the multilayer network of the CIHP and real housing prices. Here, “CIHP” represents the multi-centrality of the multilayer network constructed using the CIHP, and “price” represents the multi-centrality of the multilayer network constructed using the real housing prices. In addition, we count the number of times each city has been ranked in the top 10 in various multi-centrality categories, as shown in the last column of Table 3.

**Table 3.** Multi-centrality ranking of cities.

| City | PageRank |       | Eigenvector |       | Degree |       | Degree-In |       | Degree-Out |       | Rank in Top 10 |
|------|----------|-------|-------------|-------|--------|-------|-----------|-------|------------|-------|----------------|
|      | CIHP     | Price | CIHP        | Price | CIHP   | Price | CIHP      | Price | CIHP       | Price |                |
| TJ   | 1        | 7     | 2           | 7     | 3      | 2     | 1         | 8     | 22         | 1     | 9              |
| QD   | 9        | 1     | 8           | 1     | 2      | 1     | 4         | 1     | 2          | 16    | 9              |
| SY   | 2        | 20    | 1           | 9     | 4      | 7     | 3         | 18    | 22         | 2     | 7              |
| BJ   | 3        | 13    | 23          | 19    | 7      | 7     | 9         | 18    | 10         | 2     | 6              |
| TY   | 24       | 6     | 10          | 8     | 21     | 4     | 18        | 9     | 19         | 5     | 6              |
| SJZ  | 6        | 3     | 14          | 2     | 16     | 13    | 7         | 5     | 28         | 29    | 5              |
| WH   | 4        | 31    | 3           | 31    | 4      | 18    | 2         | 13    | 28         | 16    | 4              |
| XA   | 10       | 11    | 11          | 6     | 27     | 4     | 26        | 4     | 15         | 12    | 4              |
| CS   | 12       | 26    | 7           | 13    | 1      | 18    | 7         | 13    | 1          | 16    | 4              |
| SH   | 16       | 5     | 30          | 3     | 10     | 21    | 18        | 15    | 5          | 16    | 4              |
| ZZ   | 23       | 9     | 24          | 16    | 19     | 6     | 18        | 6     | 15         | 9     | 4              |
| YC   | 30       | 14    | 21          | 28    | 7      | 7     | 11        | 3     | 9          | 33    | 4              |
| NC   | 5        | 30    | 15          | 25    | 21     | 14    | 26        | 10    | 11         | 10    | 3              |
| CQ   | 8        | 18    | 5           | 29    | 16     | 33    | 32        | 26    | 2          | 29    | 3              |
| HK   | 14       | 33    | 4           | 26    | 7      | 28    | 15        | 32    | 5          | 16    | 3              |
| FZ   | 15       | 16    | 25          | 10    | 31     | 10    | 18        | 18    | 33         | 4     | 3              |

Table 3. Cont.

| City | PageRank |       | Eigenvector |       | Degree |       | Degree-In |       | Degree-Out |       | Rank in Top 10 |
|------|----------|-------|-------------|-------|--------|-------|-----------|-------|------------|-------|----------------|
|      | CIHP     | Price | CIHP        | Price | CIHP   | Price | CIHP      | Price | CIHP       | Price |                |
| SZ   | 18       | 4     | 29          | 15    | 4      | 28    | 15        | 26    | 2          | 22    | 3              |
| NN   | 19       | 2     | 19          | 4     | 31     | 18    | 30        | 10    | 19         | 22    | 3              |
| HRB  | 20       | 27    | 16          | 22    | 13     | 3     | 9         | 2     | 15         | 12    | 3              |
| JN   | 22       | 10    | 20          | 12    | 10     | 28    | 4         | 18    | 22         | 29    | 3              |
| HHHT | 25       | 12    | 13          | 5     | 21     | 10    | 11        | 6     | 28         | 12    | 3              |
| NB   | 27       | 32    | 27          | 20    | 18     | 10    | 30        | 15    | 5          | 5     | 3              |
| CD   | 11       | 8     | 6           | 18    | 31     | 23    | 25        | 18    | 28         | 16    | 2              |
| LZ   | 13       | 21    | 26          | 27    | 10     | 28    | 4         | 32    | 22         | 16    | 2              |
| HZ   | 7        | 19    | 17          | 21    | 13     | 25    | 11        | 18    | 11         | 22    | 1              |
| GY   | 17       | 28    | 9           | 23    | 21     | 25    | 15        | 26    | 22         | 16    | 1              |
| GZ   | 21       | 29    | 28          | 33    | 19     | 28    | 32        | 26    | 5          | 22    | 1              |
| KM   | 26       | 23    | 12          | 24    | 13     | 15    | 11        | 26    | 11         | 7     | 1              |
| NJ   | 29       | 15    | 31          | 30    | 27     | 15    | 18        | 18    | 22         | 8     | 1              |
| HF   | 31       | 17    | 32          | 14    | 21     | 15    | 26        | 10    | 11         | 12    | 1              |
| DL   | 28       | 24    | 22          | 32    | 27     | 23    | 26        | 18    | 15         | 16    | 0              |
| CC   | 32       | 25    | 18          | 17    | 21     | 21    | 18        | 26    | 19         | 10    | 0              |
| XM   | 33       | 22    | 33          | 11    | 30     | 25    | 18        | 15    | 28         | 29    | 0              |

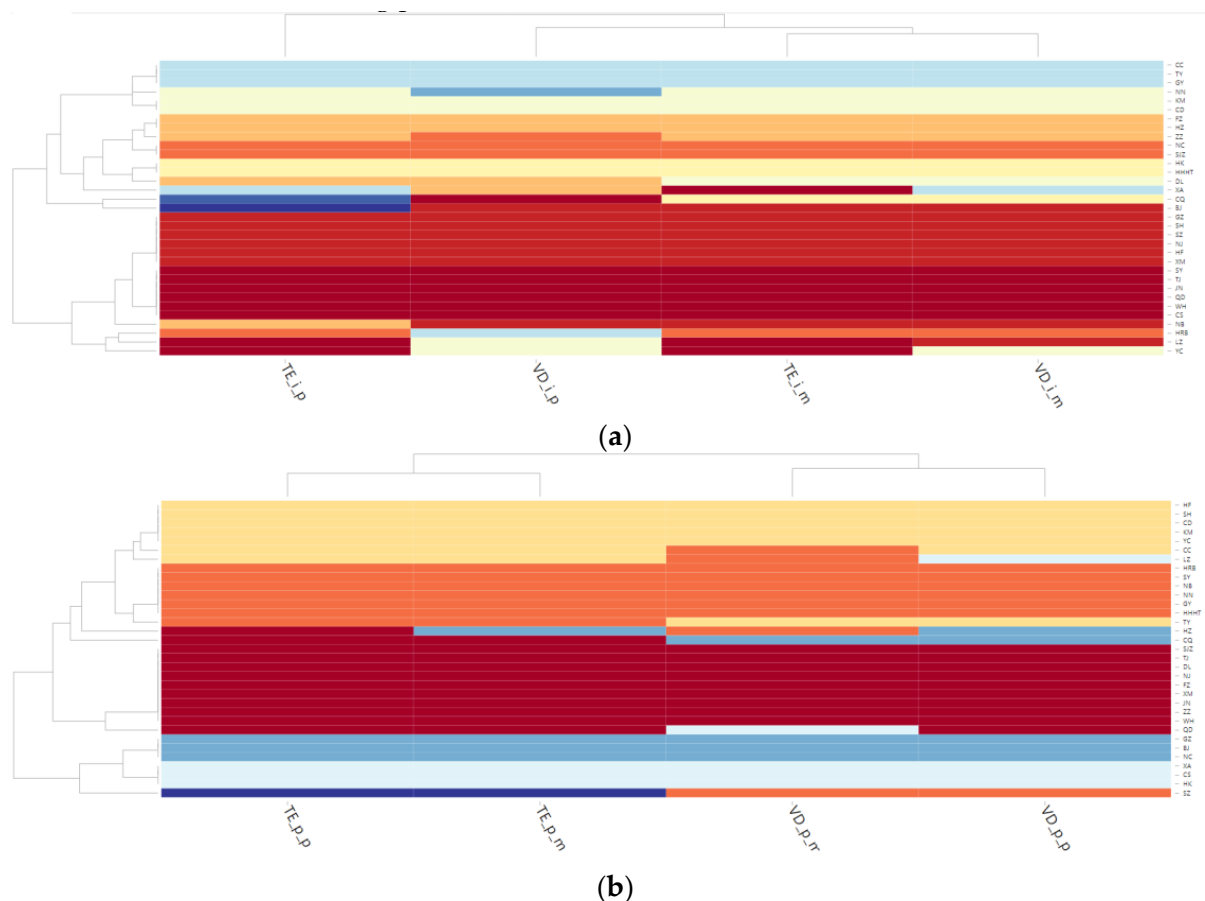
#### 4.3. Community Analysis of Multilayer Networks

By employing the Multiplex Infomap method, we reveal the community structure of multilayer networks constructed using the CIHP and real housing prices. The Multiplex Infomap can integrate the characteristics of multiple networks and group individuals with similar characteristics into the same community, resulting in a large gap between different communities and a small gap between the same community. In this manner, we can investigate which cities have similar characteristics and gain an improved understanding of the linkage of housing prices, which is critical for subsequent housing price control. As shown in Figure 6, we use the Multiplex Infomap method to cluster the communities of the two four-layer networks and obtain a community clustering diagram of the CIHP and real housing prices.

According to the community clustering diagram of the CIHP, the cities can be divided into six communities. The first-tier cities of Beijing, Shanghai, Guangzhou, and Shenzhen are all part of the same community, as are Nanjing, Hefei, and Xiamen. These three cities are either adjacent to Shanghai (Nanjing and Hefei) or in the same coastal special economic zone as Shenzhen (Xiamen) and thus do not have a large gap relative to the first-tier cities. The first-tier cities are not all part of the same community according to the community clustering diagram of real housing prices. Except for Beijing and Guangzhou, which are in the same community, Shenzhen and Shanghai are in separate communities with other cities and are separated by a considerable distance. For example, Shanghai is in the same community as Hefei, Chengdu, Kunming, and Yinchuan, whereas Shenzhen is not clearly in the same community as the other cities. In general, the community clustering result obtained by using the CIHP is more effective for first-tier cities and is highly correlated with their actual economic level and geographical location.

The three cities of Tianjin, Qingdao, and Shenyang, which are among the top 10 cities in various categories of multi-centrality, are in the same community clustering of the CIHP, as are Jinan, Wuhan, and Changsha. Tianjin and Qingdao remain in the same community in terms of real housing price clustering, whereas Shenyang is a member of a separate community. Jinan and Wuhan are also in the same community as Tianjin and Qingdao, whereas Changsha is not in the same community as Tianjin or Shenyang. The clustering

results of the two communities show that the central cities with a higher importance are more likely to share the same characteristics and attributes, making it easier for them to be grouped together.



**Figure 6.** Community clustering graph of multilayer networks. (a) Community clustering of multilayer networks using the comprehensive index of housing prices. (b) Community clustering of multilayer networks using real housing prices.

It can be concluded that the multilayer complex network community clustering results obtained from the CIHP are superior; these results not only accurately divide the first-tier cities into the same community but also cluster the cities with higher rankings as per the multilayer network structure characteristics into the same community. There is still a substantial gap between the outcomes of real housing prices and the actual situation, and the outcomes of community division are less effective than those from the CIHP. Therefore, the multilayer complex network constructed using the CIHP is more consistent with the actual situation and has a high degree of recognition with respect to identifying important nodes.

## 5. Conclusions and Discussion

Focusing on the risk contagion of housing prices, we used the TE and GVD methods to construct a fully connected network with respect to housing price correlation. Subsequently, we simplified this network using the DMST and DPMFG, yielding two different four-layer multiplex network graphs of the CIHP and real housing prices, respectively. By analyzing the source cities from the DMST network, we found that the source cities are Guangzhou, Xiamen, and Shenzhen, which are all among the top cities in the CIHP. In addition, according to the node size of the DPMFG network, the central city is not a traditional first-tier city; rather, new first-tier cities appears more frequently in the central location.

Calculation the multi-centrality of multilayer networks shows that traditional first-tier cities do not rank extremely high in the various categories of multi-centrality. However, Tianjin, Qingdao, and Shenyang, which are new first-tier cities, occupy the central position in the multilayer networks. Furthermore, new first-tier cities replace first-tier cities as the core of the price fluctuation correlation network. Subsequently, we used the Multiplex Infomap method to cluster multilayer networks and found that the community cluster based on the CIHP is most effective. The first-tier cities and some cities near Shanghai are members of the same community, as are Tianjin, Qingdao, and Shenyang, with high network structure indicators.

The methods of building multilayer networks and dynamic topology extraction can also be applied to financial networks. For example, when building stock/bank multilayer networks, we can use the closing prices of multiple stocks as the research object, build a multilayer financial network through VD and TE methods, and build dynamic financial networks for crisis early warning research. The method used in the present study has good applicability in the financial field.

In the future, we will consider using machine learning to identify early warning signs of rising housing prices. We intend to convert housing data into a symbolic sequence (0 and 1) corresponding to price rises and falls. Through the sliding window method and multilayer network structured indicators (multi-PageRank, multi-eigenvector, etc.) constructed above as the characteristics (X1, X2, X3, etc.) of a machine learning model, various machine learning models can be used to forecast the rise or fall in housing prices.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/e24091305/s1>, Appendix: Construction of multilayer networks.

**Author Contributions:** Data curation, R.S.; Supervision, Z.W.; Writing—original draft, R.S.; Writing—review & editing, L.Q. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the General project of NSFC: Research on multi-dimensional and multiple contagions of Banking Systemic Financial Risk in structural change (grant number 71973098) and Youth project of NSFC: Financial risk early-warning based on high dimensional supervised network (12105178).

**Data Availability Statement:** Data available in a publicly accessible repository. (<https://www.anjike.com/fangjia>) (accessed on 27 June 2022).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Li, Z.; Wang, X.J.; Liu, Q. Housing Price Volatility Spillovers in China: From the Perspective of Risk Network. *J. Cent. Univ. Financ. Econ.* **2021**, *4*, 114–128. [[CrossRef](#)]
- Ding, R.X.; Ni, P.F. Regional Spatial Linkage and Spillover Effect of House Prices of China's Cities—Based on the Panel Data of 285 Cities from 2005 to 2012. *Financ. Trade Econ.* **2015**, *36*, 136–150. [[CrossRef](#)]
- Chen, L.N.; Wang, H. Empirical Investigation on the Regional Interactions of Real Estate Prices in China. *Stat. Res.* **2012**, *29*, 37–43. [[CrossRef](#)]
- Gong, Y.L.; Haan, J.; Boelhouwer, P. Cross-city spillovers in Chinese housing markets: From a city network perspective. *Pap. Reg. Sci.* **2020**, *99*, 1065–1085. [[CrossRef](#)]
- MacDonald, R.; Taylor, M.P. Regional House Prices in Britain: Long-Run Relationships and Short-Run Dynamics. *Scott. J. Political Econ.* **1993**, *40*, 43–55. [[CrossRef](#)]
- Alexander, C. Seasonality and Cointegration of Regional House Prices in the UK. *Urban Stud.* **1994**, *31*, 1667–1689. [[CrossRef](#)]
- Wang, J.Y.; Liu, X.L. Residential Fundamental Value, Bubble Component and Regional Spillover Effect. *China Econ. Q.* **2014**, *13*, 1283–1302. [[CrossRef](#)]
- Yang, J.; Yu, Z.L.; Deng, Y.H. Housing price spillovers in China: A high-dimensional generalized VAR approach. *Reg. Sci. Urban Econ.* **2017**, *68*, 98–114. [[CrossRef](#)]
- Lv, L.; Liu, H.Y. Research on the Measurement, Network Structure and Influencing Factors of the Spillover Effect of Urban Housing Prices. *Econ. Rev.* **2019**, *2*, 125–139. [[CrossRef](#)]
- Li, Z.; Liu, Q.; Liang, Q. A Study on Forestalling China's Systemic Risk Based on Financial Industry and Real Economy Interacted Network. *Stat. Res.* **2019**, *36*, 23–37. [[CrossRef](#)]



11. Liu, X.L.; Gu, S.T. Research on risk spillovers from the real estate department to financial system based on AR-GARCH-CoVaR. *Syst. Eng. Theory Pract.* **2014**, *34*, 106–111.
12. Liu, H.H.; Chen, S.H. Nonlinear relationships and volatility spillovers among house prices, interest rates and stock market prices. *Int. J. Strateg. Prop. Manag.* **2016**, *20*, 371–383. [[CrossRef](#)]
13. Yamaka, W.; Liu, J.X.; Li, M.Y.; Maneejuk, P.; Dinh, H.Q. Analyzing the Causality and Dependence between Exchange Rate and Real Estate Prices in Boom-and-Bust Markets: Quantile Causality and DCC Copula GARCH Approaches. *Axioms* **2022**, *11*, 113. [[CrossRef](#)]
14. Zeng, X.W.; Liu, Z.D.; Liu, W.Y. Study on Transmission and Volatility Spillover of Housing Prices in China's Urban Agglomerations. *Manag. Rev.* **2015**, *27*, 3–13. [[CrossRef](#)]
15. Chen, M.H.; Liu, H.J.; Sun, Y.N.; He, L.W. Empirical Study on the Characteristics of Network Structure and Its Influence Factors of Urban Housing Price Linkage: Based on Monthly Data of 69 Large and Medium Cities in China. *South China J. Econ.* **2016**, *1*, 71–88. [[CrossRef](#)]
16. Chen, M.H.; Wang, S.; Liu, W.F.; Liu, Y.X. Measurement and Analysis of Correlation Effect of Urban Housing Price under the Nonlinear Perspective. *China Soft Sci.* **2020**, *10*, 96–106.
17. Schreiber, T. Measuring information transfer. *Phys. Rev. Lett.* **2000**, *85*, 461–464. [[CrossRef](#)]
18. Chen, Y.C. *Market Efficiency and Information Flow between Stocks with Different Systematic Risk Based on the Transfer Entropy*; Beijing Jiaotong University: Beijing, China, 2014.
19. Diebold, F.X.; Yilmaz, K. On the network topology of variance decompositions: Measuring the connectedness of financial firms. *J. Econom.* **2014**, *182*, 119–134. [[CrossRef](#)]
20. Kwon, O.; Yang, J.S. Information flow between stock indices. *EPL Europhys. Lett.* **2008**, *82*, 68003. [[CrossRef](#)]
21. Qiu, L.; Nan, W.Y. Brain Network Constancy and Participant Recognition: An Integrated Approach to Big Data and Complex Network Analysis. *Front. Psychol.* **2020**, *11*, 1003. [[CrossRef](#)]
22. Ye, L.; Wang, Y.Z.; Chen, Y.Y. Study on the Risk Spillover Effect between Chinese Financial Institution. *J. Stat. Inf.* **2019**, *34*, 54–63.
23. Hemminger, R.L. On the group of a directed graph. *Can. J. Math.* **1966**, *18*, 211–220. [[CrossRef](#)]
24. Gabow, H.N.; Galil, Z.; Spencer, T.H.; Tarjan, R.E. Efficient algorithms for finding minimum spanning trees in undirected and directed graphs. *Combinatorica* **1986**, *6*, 109–122. [[CrossRef](#)]
25. De Domenico, M.; Solé-Ribalta, A.; Cozzo, E.; Kivelä, M.; Moreno, Y.; Porter, M.A.; Gómez, S. Mathematical Formulation of Multilayer Networks. *Phys. Rev. X* **2013**, *3*, 041022. [[CrossRef](#)]
26. De Domenico, M.; Solé-Ribalta, A.; Omodei, E.; Gómez, S.; Arenas, A. Ranking in interconnected multilayer networks reveals versatile nodes. *Nat. Commun.* **2015**, *6*, 6868. [[CrossRef](#)]
27. De Domenico, M.; Lancichinetti, A.; Arenas, A.; Rosvall, M. Identifying Modular Flows on Multilayer Networks Reveals Highly Overlapping Organization in Interconnected Systems. *Phys. Rev. X* **2015**, *5*, 011027. [[CrossRef](#)]
28. Daniel, E.; Ludvig, B.; Martin, R. Mapping Higher-Order Network Flows in Memory and Multilayer Networks with Infomap. *Algorithms* **2017**, *10*, 112.
29. Zhang, X.Y. Chinese Commercial Banks: Balance Development and Security. *J. Quant. Technol. Econ.* **2021**, *38*, 66–87. [[CrossRef](#)]
30. Boccaletti, S.; Bianconi, G.; Criado, R.; Del Genio, C.I.; Gómez Gardeñes, J.; Romance, M.; Sendiña Nadal, I.; Wang, Z.; Zanin, M. The structure and dynamics of multilayer networks. *Phys. Rep.* **2014**, *544*, 121–122. [[CrossRef](#)]