Full paper

International Journal of Real Estate Studies

INTREST

Dynamic Relationships among Composite Property Prices of Major Chinese Cities: Contemporaneous Causality through Vector Error Corrections and Directed Acyclic Graphs

Xiaojie Xu[1]*, Yun Zhang[2]

[1]Department of Economics, North Carolina State University, Raleigh, NC, United States
[2]Department of Materials Science and Engineering, North Carolina State University, Raleigh, NC, United States

*Corresponding author's email: xxu6@alumni.ncsu.edu

Article history: Received: *1 March 2023* Received in revised form: *13 April 2023*
Accepted: *7 May 2023* Published online: *30 June 2023*

Abstract

The present study is the first one that investigates dynamic relations among composite real estate price indices of ten different cities in China during the years from 2005 to 2021. Utilizing the data recorded on a monthly basis, we apply VECM (vector error-correction modeling) and DAGs (directed acyclic graphs) in order to characterize contemporaneous causal relations among the ten real estate price indices. We use the PC algorithm to identify a pattern with non-directed edges and the LiNGAM algorithm to determine the causal ordering, based on which we calculate the results of innovation accounting. The LiNGAM algorithm adopted here effectively utilizes non-normality for facilitating the arrival of complete causal orderings. Our results show that price dynamics revealed through processes of price adjustments due to shocks to prices are rather sophisticated and such dynamics are, in general, dominated by price indices of Shanghai and Shenzhen, which are two top-tier cities among the four top-tier cities in China. This indicates that policy design on composite property prices should be focusing on price indices of Shanghai and Shenzhen.

Keywords: Composite real estate price indices, vector error-correction modeling, directed acyclic graphs, PC algorithm, LiNGAM algorithm

1.0 INTRODUCTION

Real estate markets have dramatically grown during the past decade in China (Peng et al., 2020). Various property types and their prices have attracted the people's close attention. Regional and spatial price trends of properties, as well as their interrelationships, are of great importance (Yang et al., 2018) as they largely influence one's decisions on where to invest, work, and reside. Previous studies in this direction include Yang et al. (2013), Gong et al. (2016), and Yang et al. (2021) for the Chinese market. In particular, Yang et al. (2013) have explored dynamics of housing prices of four Chinese cities, namely Beijing, Shanghai, Shenzhen, and Guangzhou, during the period from December 2000 to May 2010. Gong et al. (2016) have investigated relationships among housing prices of ten Chinese cities, namely Guiyang, Xiamen, Shenzhen, Fuzhou, Guangzhou, Chengdu, Nanchang, Kunming, Changsha, and Nanning, during the period from June 2005 to May 2015. Yang et al. (2021) have examined transmissions of information among housing prices of eight Chinese cities, namely Wenzhou, Shanghai, Hangzhou, Nanjing, Xuzhou, Suzhou, Yangzhou, and Wuxi, during the period from 2009 to 2018 by focusing on the transaction volume. The focus of these three studies is the residential property market.

In the present study, we focus on examining the problem of contemporaneous causal orderings among composite real estate price indices of ten different Chinese cities, namely Chengdu, Beijing, Wuhan, Shanghai, Nanjing, Tianjing, Hangzhou, Chongqing, Guangzhou, and Shenzhen, during the period from 2005 to 2021. To fulfill this purpose of analysis, we utilize the technique of directed acyclic graphs (DAGs), through which one could arrive at data-driven results of innovation accounting rather than relying on traditional approaches that are based upon subjective judgment (Awokuse, 2007). As compared to previous work by Xu and Zhang (2022b), where contemporaneous causality is investigated based upon regional residential housing markets in China and Beijing, Shenzhen, and Guangzhou are found to dominate price adjustment processes of the residential sub-sector, the examination of composite property price indices in the present work aims at the issue for overall real estate markets as sub-sectors are different from each other and the composite index is designed to reveal the status of the comprehensive real estate market. Different from the finding in Xu and Zhang (2022b) for the residential sub-sector, this work determines that for regional composite property price indices, only Shanghai and Shenzhen dominate regional price adjustment processes.

The empirical framework that facilitates our analysis of issues of contemporaneous causal orderings and corresponding impulse responses of prices following shocks here includes vector error-correction modeling (VECM) and DAGs. In particularly, two algorithms are considered for inferring for DAGs, namely the PC algorithm and the Linear Non-Gaussian Acyclic Model (LiNGAM) algorithm. To the best of our knowledge, the present study represents the first research on applying VECM and DAGs for analyzing contemporaneous causality among the composite real estate price indices of ten different major Chinese cities. Particularly for our analysis here, we use the PC algorithm to identify a pattern with non-directed edges and the LiNGAM algorithm to determine causal orderings. The LiNGAM algorithm adopted here effectively utilizes non-normality for facilitating the arrival of complete causal orderings. Having causal orderings inferred, we calculate impulse responses for examining the problem of dynamic relationships among all of the composite real estate price indices under consideration. Our results show that price dynamics revealed through processes of price adjustments due to shocks to prices are rather sophisticated and such dynamics are, in general, dominated by price indices of Shanghai and Shenzhen, which are two top-tier cities among the four top-tier cities in China. This indicates that policy design on composite property prices should be focusing on price indices of Shanghai and Shenzhen.

For the remainder of this work, Section 2 provides a review of the literature, Section 3 presents the price data used for analysis, Section 4 describes the methods adopted, Section 5 discusses our results, and Section 6 concludes the study.

2.0 LITERATURE REVIEW

For the purpose of analyzing dynamics among regional real estate prices, the lead-lag relationship is one of research directions that attracts attention of economists. For example, Grigoryeva and Ley (2019) concentrate on real estate markets of the Vancouver metropolitan area for studying price ripple effects and determine that it takes three months to have the price information transmitted from an epicenter of origination to other metropolitan areas and longer time periods to have the price information transmitted from an originating epicenter to municipalities that are more distant. Chiang and Tsai (2016) concentrate on real estate markets in the United States for examining regional housing price dynamics through vector autoregression modeling, Granger lead-lag relation testing, and innovation accounting analysis. Their results determine three cities that serve as regional centers of price information, including Miami for the southern US, New York for the eastern US, and Los Angeles for the western US. Fan et al. (2019) explore the real estate market in China by analyzing lead and lag dynamics among five different cities' housing prices, including Beijing, Guangzhou, Shenzhen, Shanghai, and Tianjin, during 2008–2014, and suggest that understandings of price dynamics could help investors optimize portfolios and manage risks. Zhang and Liu (2009) also investigate real estate markets in China with the focus of ripple effects of eight different cities' housing prices, including Shenzhen, Shanghai, Chengdu, Hangzhou, Xi'an, Nanjing, Haerbin, and Beijing, during 1998–2007 through analysis based upon vector error correction modeling and Granger lead-lag relation testing. Their results show that Shenzhen is the most important source of price discovery.

In addition to the lead-lag relationship, contemporaneous causality represents another important research direction of analyzing dynamics among regional real estate prices. For example, Zohrabyan et al. (2007) focus on real estate markets in the United States and analyze regional price dynamics of eight regions during 1975–2006 through VECM and DAGs. Their results suggest that New England and West North Central tend to serve as the most essential information sources of real estate prices due to their relative economic advantages over other regions. Teye and Ahelegbey (2017) concentrate on real estate markets in the Netherlands and examine spatio-temporal relationships among housing prices of twelve provinces by employing Bayesian vector autoregressive modeling and directed acyclic graphs. They find that Noord-Holland dominates the price adjustment process during earlier periods and Drenthe dominates during later periods of their sample. Hu et al. (2020) utilize directed acyclic graphs for assessing price diffusion processes among intra-urban real estate markets in Shanghai, China and determine that there exist spiraling effects of local real estate prices due to bidirectional spillover effects between high-priced and low-priced markets. Wang et al. (2018) propose dividing cities in China into five different tiers and adopt directed acyclic graphs for investigating price spillovers among the corresponding regional real estate markets during 2011–2016. They find that the second tier real estate markets have the most important price spillover effects.

3.0 DATA

We assembled data of composite property price indices from the China Real Estate Index System (CREIS)[1]. The composite real estate price indices have the coverage of ten different major Chinese cities, namely Chengdu, Beijing, Wuhan, Shanghai, Nanjing, Tianjing, Hangzhou, Chongqing, Guangzhou, and Shenzhen. According to CREIS, their price data collection processes involve multiple channels, such as surveys through field visits, telephone calls, and the internet. For calculations of composite real estate price indices, CREIS uses samples of three types of listed-for-sale properties, namely retail properties, office properties, and residential properties, from each of the ten cities in a given month. New developments are added for calculations of the indices on a monthly basis across the ten cities. If a particular new construction project is multi-phase in nature and its p−th phase has been sold out while its $(p + 1)$−th phase has not been ready yet for sales purposes, calculations of the price indices would utilize the results associated with the p−th phase. In addition, if available-for-sale units are below 5% for a multi-phase project's total units, the project would no longer be included as part of the calculations of the price indices. Collected samples by CREIS include corresponding information as follows: average listed prices, clubhouses, ratios of parking spots to property units, acreage of the land and construction, categories of buildings, types of properties, districts of properties, ratios of floors to areas, average rates of occupancy, fees associated with property management, standards of furnishments, square footage of main types of units, sales incentives, landscaping ratios, numbers of units, and outdoor playground

[1] More institutional background about the system might be found from Xu and Zhang (2021, 2022a).

facilities. CREIS' data collected also went through multiple crosscheck processes for quality assurance purposes. It should be noted here that only the time-series data of composite real estate price indices are obtained from CREIS for analysis in this work.

Given city j, its residential, office, or retail property price index is represented as:

$$P_j^t = \frac{\sum P_{ij}^t \times Q_{ij}}{\sum Q_{ij}}. \tag{1}$$

In Equation (1), P_j^t is utilized to denote the average property price in city j at time t, P_{ij}^t is utilized to denote the property price of the ith construction project in city j at time t, and Q_{ij} is utilized to denote the ith construction project's construction area in city j. According to CREIS, the composite property price index is calculated as the result of the weighted average of price indices of three types of properties, namely retail properties, office properties, and residential properties. Specifically, the residential, office, and retail sectors receive weights of 75%, 15%, and 10%, respectively.

The period of the data analyzed ranges from July of 2005 to April of 2021. Composite real estate price indices of the ten cities are depicted in Figure 1 (top panel), as well as the first differences of the ten price indices (bottom panel). It should be noted that Beijing's composite property price in December of 2000 is utilized as the price index for the base period and the index value is set to 1,000. Composite property price indices of the ten cities in different months are normalized based upon this base period price index.

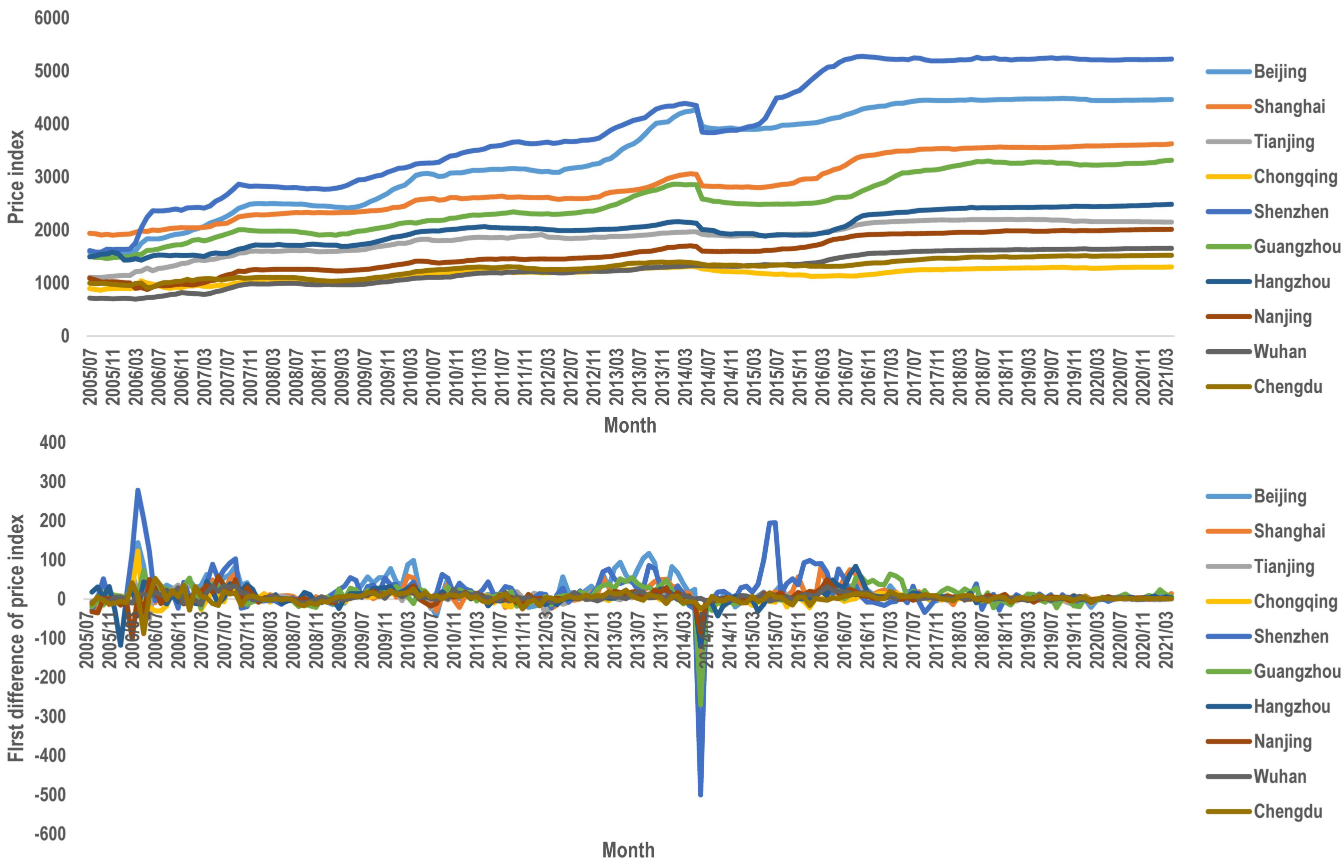

Figure 1 Composite property price indices (top panel) and their first differences (bottom panel) of Chengdu, Beijing, Wuhan, Shanghai, Nanjing, Tianjing, Hangzhou, Chongqing, Guangzhou, and Shenzhen. The horizontal axis plots the month and year and the vertical axis plots the composite property price index (top panel) or the first differences of the composite property price index (bottom panel).

4.0 Method

Although time series models are often seen in the literature when researchers investigate lead-lag causal relations, one generally should confront the problem of contemporaneous causal orderings as part of gaining understandings of contemporaneous consequences due to interventions and shocks (Xu, 2017, 2019). Specifically, a standard approach that one could take for answering the research question of whether a particular economic variable's lags might benefit forecasting another economic variable's value associated with the current period is to examine the Granger causality (Campbell et al., 2001; Guo & Savickas, 2008). However, if one infers causal relationships based upon Granger causality's standard meanings, there exist obvious reasons why it could lead to misleading results (Wang, 2010b). In particular, if one finds that the economic variable B is Granger caused by the economic variable A, the only implication is that the economic variable A contains forward-looking information. One could not determine whether the economic variable A will cause the economic variable B to go up or go down (Wang, 2010b). In addition, even if it is true that the economic variable B is caused by the economic variable A, the data aggregation problem could lead to the result of one not being able to identify the causal relationship in the manner of how one implements Granger causal testing (Wang, 2010b). Performing inference of contemporaneous causal relations is generally difficult in a non-structural model. For this situation, the technique of the DAG can be applied, which is helpful in constructing contemporaneous innovations' covariance orthogonalization. Actually, the DAG technique represents a powerful data-driven method for fulfilling the purpose of obtaining inferences of innovation accounting (Swanson & Granger, 1997). In economic research, the DAG technique has helped solve problems of causal orderings across many different economic variables (Bessler & Akleman, 1998; Bessler, Yang and Wongcharupan, 2003; Chopra & Bessler, 2005; Ciarli et al., 2019; Coad & Grassano, 2019; Haigh & Bessler, 2004). For arriving at the DAG, different algorithms could be employed. One of such algorithms that has been widely sought for economic research (e.g. Awokuse & Bessler, 2003; Bessler & Akleman, 1998; Bessler & Yang, 2003; Bessler et al., 2003; Haigh & Bessler, 2004; Lai & Bessler, 2015; Wang, 2010a; Wang et al., 2007; Xu, 2019; Yang, 2003; Yang & Bessler, 2004) is the PC algorithm (Spirtes et al., 2000). We also consider this algorithm in our work. One potential issue of employing this algorithm is that one might encounter observational equivalence. If this is the case, it is possible that one could not obtain complete causal orderings as some edges might not be able to be directed. Therefore, other than the PC algorithm, we also employ the LiNGAM algorithm by following previous studies (Lai & Bessler, 2015; Moneta et al., 2013; Shimizu et al., 2006; Xu, 2019) on time-series data. The LiNGAM algorithm is recently developed powerful technique that makes effective explorations of non-normal features in the data for enabling one to arrive at complete causal orderings.

4.1 VECM

Let X_t represent a vector of $p \times 1$ and consider the following VECM in Equation (2):

$$\Delta X_t = \mu + \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + e_t, \text{ where } t = 1, \ldots, T. \tag{2}$$

Here in Equation (2), $\Delta \mathrm{X_t} = \mathrm{X_t} - \mathrm{X_{t-1}}$, μ is utilized to denote a deterministic term, and Π and Γ_i are utilized to denote coefficient matrices[2]. In order to determine cointegrating relationships among the ten price indices, we employ the trace testing approach (Johansen, 1988, 1991). In addition, for the purpose of evaluating stabilities of any identified cointegrating relationships, we employ the recursive cointegration testing approach (Hansen & Johansen, 1999) by following Bessler et al. (2003), Yang (2003), Yang et al. (2012), and Xu (2017, 2019).

We conduct two hypothesis testing. The first one tests for if a specific price index needs to be included as part of the determined cointegration relationships. The second one tests for if a specific price index would be responsive to the disturbance to the long-run equilibrium relationships. To perform corresponding analysis, we need to impose appropriate restrictive conditions on β and α, where β and α are utilized to denote $p{\times}r$ matrices (r is utilized to denote the rank of cointegration relationships) and they satisfy $\Pi = \alpha\beta'$[3]. The null hypothesis of the first test that we would pursue is represented in Equation (3):

$$R'_{1\times p}\beta_{p\times r} = 0_{1\times r}, \tag{3}$$

which means that a price index is not to be included as part of the cointegrating relationships. Accordingly, column i of Π would be 0 for i = 1,…, p. $R'_{1\times p}$ would be expressed as (0,…, 0, 1, 0,…, 0), where 1 is utilized to denote the ith element for i = 1,…, p, if one would like to carry out the testing based upon the ith price index. Under the null, the ith price index is not part of the cointegrating relationships, and the associated test statistic, asymptotically, would follow a χ^2 distribution. Similarly, the null of the second test that we would pursue is represented in Equation (4):

$$\mathrm{B}'_{1\times \mathrm{p}}\alpha_{\mathrm{p}\times \mathrm{r}} = 0_{1\times \mathrm{r}}, \tag{4}$$

which means that a price index is not going to be responsive to the disturbance. Accordingly, row i of Π would be 0 for i = 1,…, p. $\mathrm{B}'_{1\times \mathrm{p}}$ would be expressed as (0,…0, 1, 0,…, 0), where 1 is utilized to denote the ith element for i = 1,…, p, if one would like to carry out the testing based upon the ith price index. The associated test statistic, asymptotically, would follow a χ^2 distribution as well. Previous work that includes Xu (2017, 2019), Bessler et al. (2003), and Yang et al. (2012) has carried out similar hypothesis tests for the purpose of arriving at appropriate specifications of VECM.

4.2 DAG and Analysis of Innovation Accounting

Let e_t in Equation (2) be expressed as $Ae_t = v_t$ that satisfies $E(Ae_te_t'A') = E(v_tv_t')$, where A is utilized to denote a $p{\times}p$ matrix and vt is utilized to denote a $p{\times}1$ vector that contains orthogonal structural shocks satisfying $E(v_{i,t}, v_{j,t}) = 0$ for $i \neq j$. During the process of inferring for DAGs, many algorithms would generally carry out searching and placing zeros on the matrix A. The identification condition for the matrix A (Doan, 1996) is that when $i \neq j$ $(i, j = 1, 2, \ldots, p)$, no element of the matrix A would simultaneously meet A_{ij} and $A_{ji} \neq 0$. For the purpose of carrying out analysis of innovation accounting, we would convert the VECM into its equivalent representation based upon the vector auto-regressive (VAR) model in levels for calculations of impulse responses. Considering that such a VAR has already embedded cointegrating relationships of the VECM, its resultant analysis of innovation accounting would be consistent (Phillips, 1998). Let the VAR be represented as follows in Equation (5):

$$\mathrm{X_t} = \mathrm{A}^{-1}\mathrm{A_1}\mathrm{X_{t-1}} + \mathrm{A}^{-1}\mathrm{A_2}\mathrm{X_{t-2}} + \cdots + \mathrm{A}^{-1}\mathrm{A_k}\mathrm{X_{t-k}} + \mathrm{A}^{-1}\mathrm{v_t}\ for\ \mathrm{t} = 1, \ldots, \mathrm{T}. \tag{5}$$

Here in Equation (5), A_i's are $p \times p$ matrices. The corresponding structural VAR is represented in Equation (6):

$$\mathrm{AX_t} = \mathrm{A_1}\mathrm{X_{t-1}} + \mathrm{A_2}\mathrm{X_{t-2}} + \cdots + \mathrm{A_k}\mathrm{X_{t-k}} + \mathrm{v_t}\ for\ \mathrm{t} = 1, \ldots, \mathrm{T}. \tag{6}$$

In order to understand relative effects of different price indices, X_t's responses to vt will need to be sorted out.

As one of the earliest developed algorithms for the purpose of inferring for DAGs, the PC algorithm has been widely sought in economic research. We thus employ this algorithm as well in our present work. Detailed introductions of this algorithms could be located from previous economic studies (e.g. Awokuse & Bessler, 2003; Bessler & Akleman, 1998; Bessler & Yang, 2003; Bessler et al., 2003;

---

[2] For our case, p = 10 and $X_t = (X_{1,t}X_{2,t}X_{3,t}X_{4,t}X_{5,t}X_{6,t}X_{7,t}X_{8,t}X_{9,t}X_{10,t})'$, where $X_{1,t}$, $X_{2,t}$, $X_{3,t}$, $X_{4,t}$, $X_{5,t}$, $X_{6,t}$, $X_{7,t}$, $X_{8,t}$, $X_{9,t}$, $X_{10,t}$ correspond to composite real estate price indices of the ten cities under consideration. Specifically, X_1 corresponds to Beijing, X_2 corresponds to Shanghai, X_3 corresponds to Tianjing, X_4 corresponds to Chongqing, X_5 corresponds to Shenzhen, X_6 corresponds to Guangzhou, X_7 corresponds to Hangzhou, X_8 corresponds to Nanjing, X_9 corresponds to Wuhan, and X_{10} corresponds to Chengdu.

[3] α is utilized to denote the loading matrix and β is utilized to denote the cointegration matrix.

Haigh & Bessler, 2004; Lai & Bessler, 2015; Wang, 2010a; Wang et al., 2007; Xu, 2019; Yang, 2003; Yang & Bessler, 2004). The correlation matrix V of contemporaneous innovations based upon the estimated VECM would be used for inferring DAGs through the PC algorithm.

Because using the PC algorithm could potentially lead to observational equivalence and one might not be able to arrive at complete causal orderings, we also explore the use of the LiNGAM algorithm for inferring DAGs. In particular, the LiNGAM algorithm could make effective use of non-normal data features that help address non-directed edges from the use of the PC algorithm. In this way, one could construct complete causal orderings. Detailed introductions of this algorithms could be located from previous economic studies (e.g. Lai & Bessler, 2015; Moneta et al., 2013; Shimizu et al., 2006; Xu, 2019).

5.0 Results

5.1 Preliminary Analysis

Table 1 reports the summary statistics of our price index data, from which one should be able to observe that based upon the significance level of 5%, none of the ten price indices is following a normal distribution. According to the results of the augmented Dickey-Fuller (ADF) tests (Dickey & Fuller, 1981), each price index is integrated of order one. Particularly, the penultimate column of Table 1 shows p-values associated with the ADF tests based upon the ten composite property price indices, which suggest that the levels of the price indices are not stationary. The last column of Table 1 shows $p-$values associated with the ADF tests based upon the first differences of the ten composite property price indices, which suggest that the first differences are stationary. The unit root analysis results thus suggest that it is necessary to test for cointegration, which we turn to next.

Table 1 Summary statistics of composite real estate price indices of ten cities

City	Minimum	Mean	Median	Standard Deviation	Maximum	1st Per	5th Per	95th Per	99th Per	Jarque-Bera (p-value)	ADF (raw series)	ADF (1st diff)
Beijing	1567	3426	3605	934	4483	1575	1785	4471	4478	0.005	0.999	0.001
Shanghai	1903	2824	2745	549	3627	1911	1977	3596	3612	0.009	0.999	0.001
Tianjing	1101	1849	1888	294	2201	1107	1228	2195	2199	0.003	0.999	0.001
Chongqing	866	1160	1196	123	1310	887	916	1297	1304	0.002	0.986	0.001
Shenzhen	1583	3922	3863	1114	5275	1608	2129	5245	5267	0.011	0.999	0.001
Guangzhou	1463	2486	2487	549	3315	1478	1573	3281	3304	0.017	0.999	0.001
Hangzhou	1432	2014	2007	309	2484	1435	1514	2449	2476	0.020	0.999	0.001
Nanjing	877	1554	1560	336	2012	912	974	1998	2010	0.014	0.999	0.001
Wuhan	696	1256	1243	297	1654	704	719	1645	1652	0.014	0.999	0.001
Chengdu	912	1284	1316	172	1523	947	991	1514	1521	0.009	0.999	0.001

5.2 Cointegration

Based upon different criteria, namely the Hannan Quinn's criterion, AIC criterion, or BIC criterion, 4 is the optimal number of lags for Equation (2) via the representation of the level VAR. The results of trace testing (Johansen, 1988, 1991) are presented in Table 2. We can see that 7 is the rank of cointegrating relationships, i.e. r = 7. Hansen and Johansen (1999)'s recursive testing method is employed for assessing any possible structural breaks among the ten composite real estate price indices' long-run relationships reflected through cointegration. Accordingly, results of trace test statistics, which have been normalized, are presented in Figure 2. These statistics are calculated based upon each observation from November 2013 (i.e. the 101th observation) to April 2021 (i.e. the 190th observation). The first 100 observations from July 2005 to October 2013 are employed as the base period. We also note that critical values corresponding to the level of 5% are utilized for scaling the test statistics shown in Figure 2. Therefore, when the corresponding entry is greater than one in Figure 2, we would be able to reject the null. Results in Figure 2 indicate that there are at least 7 and never more than 7 cointegration vectors. The establishment of these cointegration relations could be through the practice of leadership of the price indices (Bessler et al., 2003) of particular cities and/or similar influences (Awokuse, 2007) of external factors, such as national-level policies, on the real estate market.

Table 2 Cointegration test results for composite real estate price indices of ten cities

r	Test statistics	Critical values (5% significance level)	p-values
0	415.121	239.121	0.000
1	298.475	197.22	0.000
2	234.086	159.319	0.000
3	172.994	125.417	0.000
4	124.509	95.514	0.000
5	78.535	69.611	0.007
6	48.481	47.707	0.042
7	26.76	29.804	0.110
8	10.659	15.408	0.237

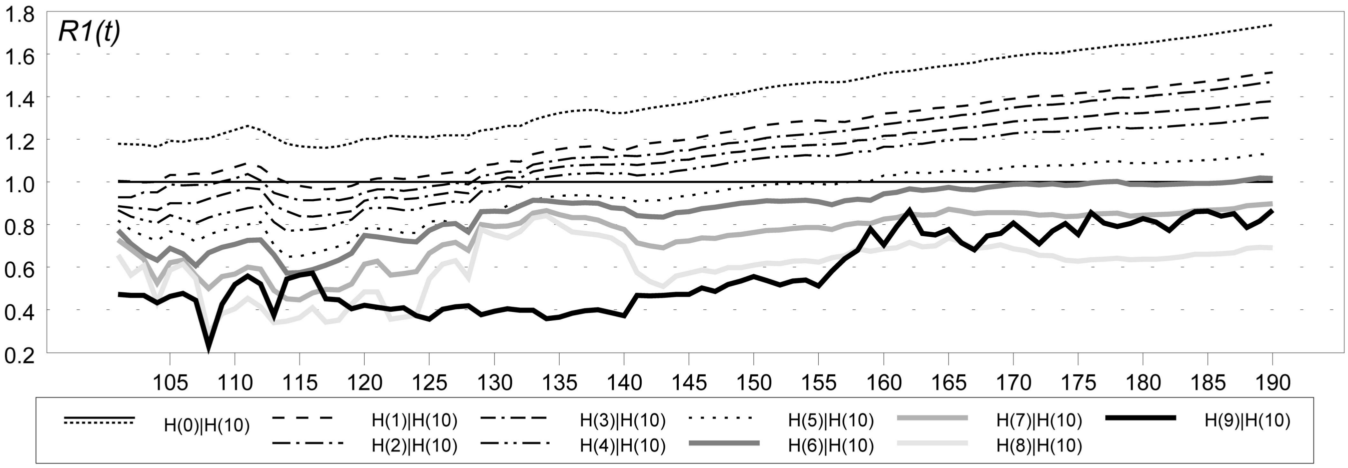

Figure 2 Analysis of recursive cointegration based upon composite real estate price indices of ten cities: the plot of statistics of trace tests. These statistics are calculated based upon each observation from November 2013 (i.e. the 101th observation) to April 2021 (i.e. the 190th observation). The first 100 observations from July 2005 to October 2013 are employed as the base period. Note that critical values corresponding to the level of 5% are utilized for scaling the test statistics. The results show that there are at least 7 and never more than 7 cointegration vectors.

5.3 Analysis of Hypothesis Testing

Having $r = 7$, we now proceed to analysis of hypothesis testing discussed in Section 4.1. The corresponding results are shown in Table 3. We can see that none of the ten composite real estate price indices is excluded from the cointegrating relationships (Panel A) and only the price index of Shenzhen is determined to be weakly exogenous (Panel B).

Table 3 Analysis of hypothesis testing based upon composite real estate price indices of ten cities

Hypothesis	Degrees of freedom	χ^2 test statistics	p-values
Panel A: Testing being excluded from the cointegrating vector			
Beijing	7	51.781	0.000
Shanghai	7	45.849	0.000
Tianjing	7	44.723	0.000
Chongqing	7	53.379	0.000
Shenzhen	7	24.593	0.001
Guangzhou	7	25.977	0.001
Hangzhou	7	62.595	0.000
Nanjing	7	38.442	0.000
Wuhan	7	32.294	0.000
Chengdu	7	30.253	0.000
Panel B: Testing weak exogeneity			
Beijing	7	16.892	0.018
Shanghai	7	14.831	0.038
Tianjing	7	27.517	0.000
Chongqing	7	62.311	0.000
Shenzhen	7	10.759	0.149
Guangzhou	7	12.765	0.078
Hangzhou	7	30.398	0.000
Nanjing	7	29.371	0.000
Wuhan	7	26.456	0.000
Chengdu	7	20.218	0.005
Shenzhen & Guangzhou	14	30.271	0.007

Long-run relations among the ten composite real estate price indices can be represented as follows in Equation (7):

$$\alpha\beta' X_{t-1} = \begin{bmatrix} -0.018 & 0.060 & -0.014 & 0.006 & -0.002 & -0.076 & 0.068 \\ -0.017 & -0.003 & -0.009 & -0.035 & -0.005 & -0.119 & 0.016 \\ -0.082 & -0.022 & -0.006 & -0.111 & 0.002 & -0.011 & 0.047 \\ -0.162 & 0.019 & 0.007 & -0.070 & 0.001 & -0.027 & -0.006 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.019 & -0.065 & -0.003 & -0.015 & -0.002 & -0.065 & 0.016 \\ 0.159 & 0.007 & 0.012 & -0.175 & -0.002 & -0.047 & 0.046 \\ 0.126 & -0.005 & -0.001 & 0.032 & 0.005 & -0.072 & -0.004 \\ 0.011 & -0.004 & -0.006 & -0.114 & 0.001 & 0.002 & -0.024 \\ -0.078 & -0.025 & 0.006 & 0.042 & -0.001 & 0.010 & 0.015 \end{bmatrix} \times$$
$$\begin{bmatrix} -0.227 & 0.606 & 0.349 & 1.000 & 0.036 & -0.070 & -0.925 & -0.010 & -0.336 & 0.344 \\ -0.365 & -0.726 & 1.000 & -0.019 & -0.125 & 0.773 & -1.163 & 2.782 & -2.249 & 1.456 \\ -0.966 & 1.000 & 4.905 & -3.653 & -1.395 & -0.351 & -1.912 & 0.925 & 3.545 & 2.708 \\ 0.157 & -0.431 & -0.324 & -0.159 & -0.012 & 0.162 & 0.468 & -0.617 & 1.000 & -0.603 \\ 0.722 & 1.000 & 7.811 & -20.088 & -0.578 & 1.593 & -2.705 & -13.965 & 0.453 & 20.559 \\ 0.050 & 0.978 & 0.023 & 1.000 & -0.257 & -0.322 & -0.461 & -0.480 & 0.189 & -0.047 \\ -0.194 & -0.015 & -0.296 & 0.640 & 0.019 & 0.202 & -0.354 & -0.112 & 1.000 & -0.643 \end{bmatrix} X_{t-1} \quad (7)$$

The matrix V based upon the estimated VECM is given in Equation (8):

$$V = \begin{bmatrix} & \text{Beijing} & \text{Shanghai} & \text{Tianjing} & \text{Chongqing} & \text{Shenzhen} & \text{Guangzhou} & \text{Hangzhou} & \text{Nanjing} & \text{Wuhan} & \text{Chengdu} \\ \text{Beijing} & 1.000 & & & & & & & & & \\ \text{Shanghai} & 0.804 & 1.000 & & & & & & & & \\ \text{Tianjing} & 0.417 & 0.436 & 1.000 & & & & & & & \\ \text{Chongqing} & 0.467 & 0.306 & 0.189 & 1.000 & & & & & & \\ \text{Shenzhen} & 0.815 & 0.779 & 0.393 & 0.379 & 1.000 & & & & & \\ \text{Guangzhou} & 0.801 & 0.750 & 0.378 & 0.436 & 0.806 & 1.000 & & & & \\ \text{Hangzhou} & 0.598 & 0.601 & 0.355 & 0.336 & 0.523 & 0.540 & 1.000 & & & \\ \text{Nanjing} & 0.568 & 0.531 & 0.230 & 0.309 & 0.494 & 0.470 & 0.411 & 1.000 & & \\ \text{Wuhan} & 0.136 & 0.135 & 0.173 & 0.164 & 0.149 & 0.058 & 0.019 & 0.307 & 1.000 & \\ \text{Chengdu} & 0.330 & 0.209 & 0.135 & 0.283 & 0.261 & 0.201 & 0.273 & 0.239 & 0.053 & 1.000 \end{bmatrix} \quad (8)$$

The matrix V contains information of contemporaneous correlations among innovations of the price indices of the ten cities. Thus, V has incorporated information of contemporaneous causality. However, one would need additional structural assumptions or atheoretical analysis like DAGs for revealing the contemporaneous causal relationships.

5.4 DAG

Figure 3 reports a "pattern" based upon the PC algorithm at the significance levels of 15% and 20% (Awokuse & Bessler, 2003; Bessler & Akleman, 1998; Haigh & Bessler, 2004; Spirtes et al., 2000; Yang, 2003) as two edges, namely Shanghai–Guangzhou and Chongqing–Guangzhou, cannot be directed. We then turn to the LiNGAM algorithm based upon prune factors (Lai & Bessler, 2015; Shimizu et al., 2006) of 0.5 and 0.6 (Bizimana et al., 2015) and we are able to reach the DAG as shown in Figure 4. The corresponding matrix, A, can be expressed as follows in Equation (9):

$$\begin{pmatrix} a_{1,1} & a_{1,2} & 0 & 0 & 0 & a_{1,6} & 0 & 0 & 0 & 0 \\ 0 & a_{2,2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_{3,3} & 0 & 0 & 0 & 0 & 0 & a_{3,9} & 0 \\ a_{4,1} & 0 & 0 & a_{4,4} & 0 & a_{4,6} & 0 & 0 & 0 & 0 \\ a_{5,1} & a_{5,2} & 0 & 0 & a_{5,5} & a_{5,6} & 0 & 0 & 0 & 0 \\ 0 & a_{6,2} & 0 & 0 & 0 & a_{6,6} & 0 & 0 & 0 & 0 \\ 0 & a_{7,2} & 0 & 0 & 0 & 0 & a_{7,7} & 0 & 0 & 0 \\ a_{8,1} & a_{8,2} & 0 & 0 & 0 & 0 & 0 & a_{8,8} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{9,8} & a_{9,9} & 0 \\ a_{10,1} & 0 & 0 & a_{10,4} & 0 & 0 & 0 & 0 & 0 & a_{10,10} \end{pmatrix}. \quad (9)$$

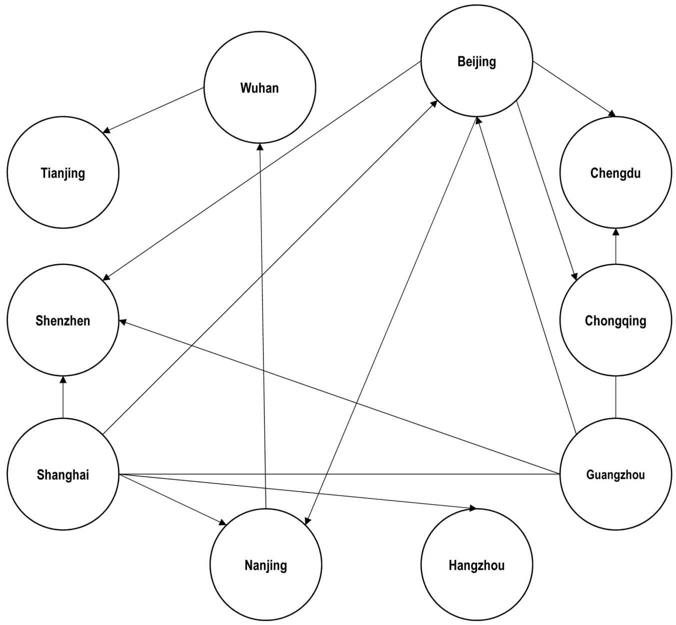

Figure 3 A causal pattern based upon innovations on composite real estate price indices of ten cities according to the PC algorithm. It is called a pattern because the edge between Shanghai and Guangzhou and the edge between Chongqing and Guangzhou cannot be directed.

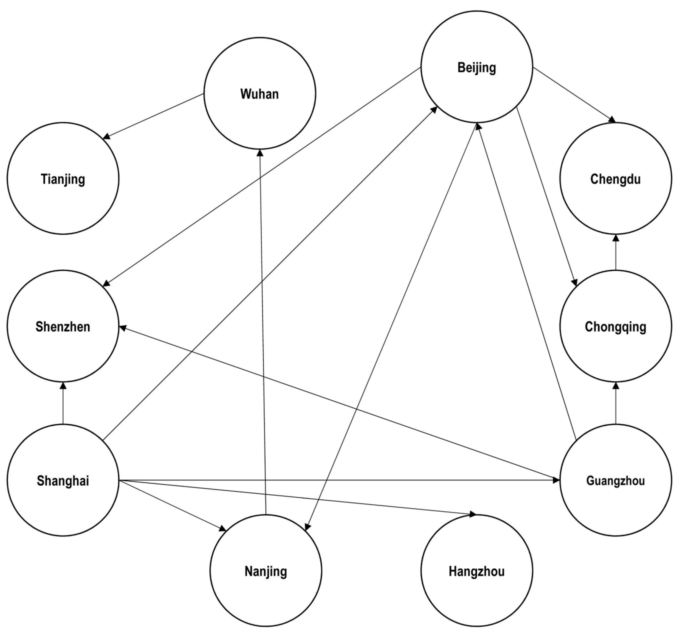

Figure 4 The DAG based upon innovations on composite real estate price indices of ten cities according to the LiNGAM algorithm. As compared to Figure 3, all edges are directed in this figure.

5.5 Innovation Accounting

With A in Equation (9), we proceed to innovation accounting analysis through calculating impulse responses for the period of sixty months and report the associated results in Figure 5. One should be able to make comparisons of the results shown in different sub-figures of Figure 5 as they all have been normalized based upon the standard deviation of the historical innovations for a certain price index. Results in Figure 5 show that price dynamics revealed through processes of price adjustments due to shocks to prices are rather sophisticated and such dynamics are, in general, dominated by price indices of Shanghai and Shenzhen, which are two top-tier cities among the four top-tier cities in China. A price shock to the composite real estate price index of Shanghai or Shenzhen would generate permanent increases in indices of all other cities. A price shock to the composite real estate price index of a city other than Shanghai and Shenzhen generally do not have profound influences on indices of other cities. This indicates that policy design on composite property prices should be focusing on price indices of Shanghai and Shenzhen.

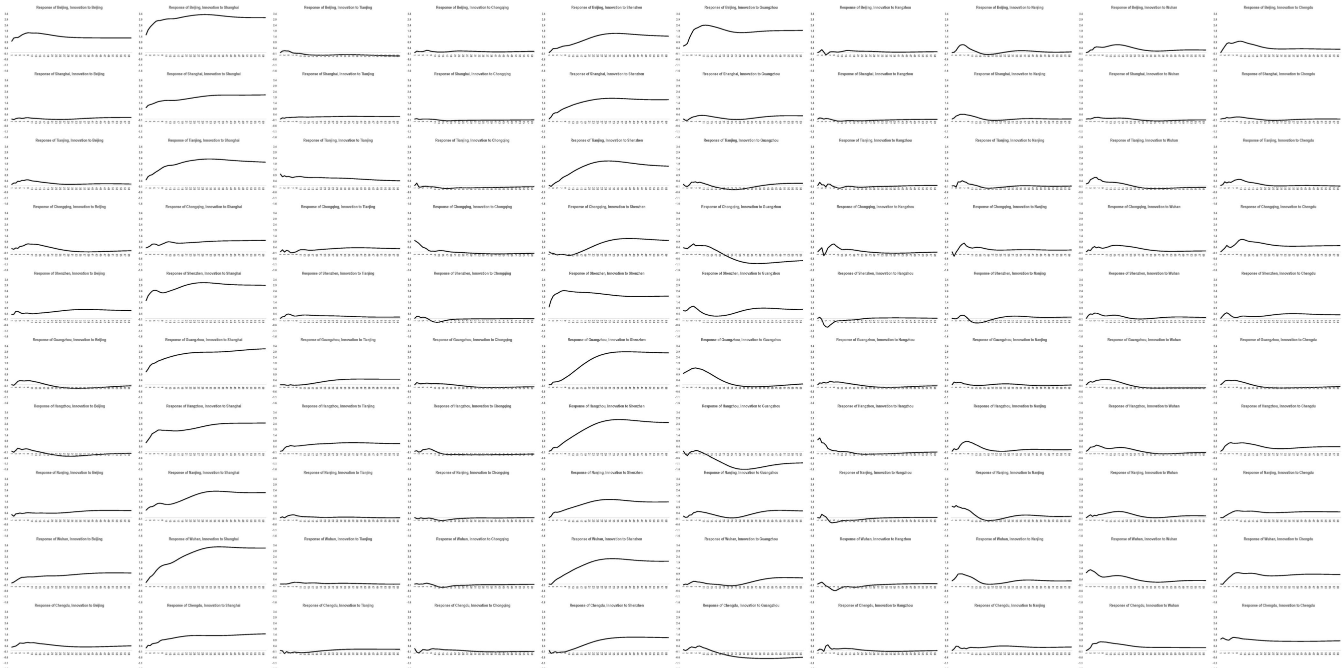

Figure 5 Impulse responses of one city's composite real estate price index to a one-time-only shock in the innovations in another city's composite real estate price index for ten cities. For each subfigure, the horizontal axis plots the month into the future and the vertical axis plots the impulse response. The title of each subfigure contains information about where the innovation occurs and where the impulse response follows.

6.0 CONCLUSION

The composite property price index is designed to reflect the overall status of real estate markets in China. In the present work, we examine the issue of contemporaneous causal orderings among composite property price indices of ten different major cities in China during the period from July 2005 to April 2021, the coverage of which not only addresses potential dimensionality constraints (Zhang & Fan, 2019) but also represents an economic natural way to assess the issue. With the real estate market in China going through policy governance back and forth over the past decade, which aims at preventing market over-heating while maintaining healthy growth (Zhao et al., 2021), analysis of dynamics among the composite property price indices of different major cities is particularly valuable to investors and policymakers as results here could help understandings of how sophisticated real estate price information reverberates through the complex system of different regional markets. Such knowledge could be useful to policymakers in deigning, implementing, and monitoring different policies targeting organic growth of the real estate business sector and useful to investors in planning, adjusting, and optimizing investment portfolios and managing market risks. We apply the framework of VECM (vector error-correction modeling) and the DAG (directed acyclic graph) to fulfill our analysis goal and find that price adjustment processes are generally dominated by two top-tier cities, namely Shanghai and Shenzhen. This empirical finding is somewhat different from the potential general perception for residential real estate markets that price adjustment processes should be dominated by all of the four top-tier cities, namely Shenzhen, Beijing, Guangzhou, and Shanghai. Our results thus shed light on the two targeted cities for price discovery processes of the overall real estate markets in China and narrow down the list of four top-tier cities to two for investors and policymakers.

Acknowledgement

We thank the three anonymous referees and Maimunah Sapri (the Editor-in-Chief of INTREST) for their helpful and insightful comments.

References

Awokuse, T. O. (2007). Market reforms, spatial price dynamics, and China's rice market integration: A causal analysis with directed acyclic graphs. *Journal of Agricultural and Resource Economics, 32*(1), 58-76.

Awokuse, T. O., & Bessler, D. A. (2003). Vector autoregressions, policy analysis, and directed acyclic graphs: An application to the U.S. economy. *Journal of Applied Economics, 6*(1), 1-24.

Bessler, D. A., & Akleman, D. G. (1998). Farm prices, retail prices, and directed graphs: Results for pork and beef. *American Journal of Agricultural Economics, 80*(5), 1144–1149.

Bessler, D. A., & Yang, J. (2003). The structure of interdependence in international stock markets. *Journal of International Money and Finance, 22*(2), 261-287.

Bessler, D. A., Yang, J., & Wongcharupan, M. (2003). Price dynamics in the international wheat market: Modeling with error correction and directed acyclic graphs. *Journal of Regional Science, 43*(1), 1-33.

Bizimana, J.-C., Angerer, J. P., Bessler, D. A., & Keita, F. (2015). Cattle markets integration and price discovery: The case of Mali. *The Journal of Development Studies, 51*(3), 319-334.

Campbell, J. Y., Lettau, M., Malkiel, B. G., & Xu, Y. (2001). Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *The Journal of Finance, 56*(1), 1-43.

Chiang, M.-C., & Tsai, I.-C., (2016). Ripple effect and contagious effect in the US regional housing markets. *The Annals of Regional Science, 56*(1), 55-82.

Chopra, A., & Bessler, D. A. (2005). Price discovery in the black pepper market in Kerala, India. *Indian Economic Review, 40*(1), 1-21.

Ciarli, T., Coad, A., & Moneta, A. (2019, December). *Exporting and productivity as part of the growth process: Causal evidence from a data-driven structural VAR* (LEM Working Paper No. 2019/39). Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy.

Coad, A., & Grassano, N. (2019). Firm growth and R&D investment: SVAR evidence from the world's top R&D investors. *Industry and Innovation, 26*(5), 508-533.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica, 49*(4), 1057-1072.

Doan, T. (1996). *User's manual: RATS 4.0*. Evanston, IL: Estima.

Fan, Y., Yang, Z., & Yavas, A. (2019). Understanding real estate price dynamics: The case of housing prices in five major cities of China. *Journal of Housing Economics, 43*, 37-55.

Gong, Y., Hu, J., & Boelhouwer, P. J. (2016). Spatial interrelations of Chinese housing markets: Spatial causality, convergence and diffusion. *Regional Science and Urban Economics, 59*, 103-117.

Grigoryeva, I., & Ley, D. (2019). The price ripple effect in the Vancouver housing market. *Urban Geography, 40*(8), 1168-1190.

Guo, H., & Savickas, R. (2008). Average idiosyncratic volatility in G7 countries. *The Review of Financial Studies, 21*(3), 1259-1296.

Haigh, M. S., & Bessler, D. A. (2004). Causality and price discovery: An application of directed acyclic graphs. *Journal of Business, 77*(4), 1099-1121.

Hansen, H., & Johansen, S. (1999). Some tests for parameter constancy in cointegrated VAR-models. *Econometrics Journal, 2*(2), 306-333.

Hu, J., Xiong, X., Cai, Y., & Yuan, F. (2020). The ripple effect and spatiotemporal dynamics of intra-urban housing prices at the submarket level in Shanghai, China. *Sustainability, 12*(12), Article 5073.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control, 12*(2-3), 231-254.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica, 59*(6), 1551-1580.

Lai, P.-C., & Bessler, D. A. (2015). Price discovery between carbonated soft drink manufacturers and retailers: A disaggregate analysis with PC and LiNGAM algorithms. *Journal of Applied Economics, 18*(1), 173-197.

Moneta, A., Entner, D., Hoyer, P. O., & Coad, A. (2013). Causal inference by independent component analysis: Theory and applications. *Oxford Bulletin of Economics and Statistics,* 75(5), 705-730.

Peng, Z., Yu, Z., & Nong, H. (2020). Inter-type investment connectedness: A new perspective on China's booming real estate market. *Global Economic Review, 49*(2), 186-204.

Phillips, P. C. B. (1998). Impulse response and forecast error variance asymptotics in nonstationary VARs. *Journal of Econometrics, 83*(1-2), 21-56.

Shimizu, S., Hoyer, P. O., Hyvärinen, A., & Kerminen, A. (2006). A linear non-Gaussian acyclic model for causal discovery. *Journal of Machine Learning Research, 7*, 2003-2030.

Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction, and search* (2nd ed.). Cambridge, MA: MIT Press.

Swanson, N. R., & Granger, C. W. J. (1997). Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions. *Journal of the American Statistical Association, 92*(437), 357-367.

Teye, A. L., & Ahelegbey, D. F. (2017). Detecting spatial and temporal house price diffusion in the Netherlands: A Bayesian network approach. *Regional Science and Urban Economics,* 65, 56-64.

Wang, X., Han, Y.-h., & Wang, C. (2018). Measurement of connectedness and spillovers in China's housing market-based on the DAG and spillover index model. *Journal of Applied Statistics and Management, 37*(4), 713-727.

Wang, Z. (2010a). Directed graphs, information structure and forecast combinations: An empirical examination of US unemployment rates. *Journal of Forecasting, 29*(4), 353-366.

Wang, Z. (2010b). Dynamics and causality in industry-specific volatility. *Journal of Banking & Finance, 34*(7), 1688-1699.

Wang, Z., Yang, J., & Li, Q. (2007). Interest rate linkages in the Eurocurrency market: Contemporaneous and out-of-sample Granger causality tests. *Journal of International Money and Finance, 26*(1), 86-103.

Xu, X. (2017). Contemporaneous causal orderings of US corn cash prices through directed acyclic graphs. *Empirical Economics, 52*(2), 731-758.

Xu, X. (2019). Contemporaneous and Granger causality among US corn cash and futures prices. *European Review of Agricultural Economics, 46*(4), 663-695.

Xu, X., & Zhang, Y. (2021). House price forecasting with neural networks. *Intelligent Systems with Applications, 12*, Article 200052.

Xu, X., & Zhang, Y. (2022a). Contemporaneous causality among one hundred Chinese cities. *Empirical Economics, 63*(4), 2315-2329.

Xu, X., & Zhang, Y. (2022b). Contemporaneous causality among residential housing prices of ten major Chinese cities. *International Journal of Housing Markets and Analysis.* Advance online publication. doi:10.1108/IJHMA-03-2022-0039

Yang, J., (2003). Market segmentation and information asymmetry in Chinese stock markets: A VAR analysis. *The Financial Review, 38*(4), 591-609.

Yang, J., & Bessler, D. A. (2004). The international price transmission in stock index futures markets. *Economic Inquiry, 42*(3), 370-386.

Yang, J., Liu, H., & Leatham, D. J. (2013). The multi-market analysis of a housing price transmission model. *Applied Economics, 45*(27), 3810-3819.

Yang, J., Tong, M., & Yu, Z. (2021). Housing market spillovers through the lens of transaction volume: A new spillover index approach. *Journal of Empirical Finance, 64*, 351-378.

Yang, J., Yang, Z., & Zhou, Y. (2012). Intraday price discovery and volatility transmission in stock index and stock index futures markets: Evidence from China. *Journal of Futures Markets, 32*(2), 99-121.

Yang, J., Yu, Z., & Deng, Y. (2018). Housing price spillovers in China: A high-dimensional generalized VAR approach. *Regional Science and Urban Economics, 68*, 98-114.

Zhang, D., & Fan, G.-Z. (2019). Regional spillover and rising connectedness in China's urban housing prices. *Regional Studies, 53*(6), 861-873.

Zhang, H., & Liu, L. (2009, July 11-12). Empirical research on housing price ripple effect between China's cities. In *SSME 2009. Proceedings of the 2009 IITA International Conference on Services Science, Management and Engineering* (pp. 43-46). Los Alamitos, CA: IEEE Computer Society.

Zhao, S., Zhao, K., & Zhang, P. (2021). Spatial inequality in China's housing market and the driving mechanism. *Land, 10*(8), Article 841.

Zohrabyan, T., Leatham, D., & Bessler, D. (2007, October 4-5). *Cointegration analysis of regional house prices in U.S*. Paper presented at the Regional Research Committee, NC-1014 Annual Meeting, St. Louis, MO.