Empirical Economics (2023) 64:2661–2682
https://doi.org/10.1007/s00181-023-02414-w

Assessing the impacts of pandemic and the increase in minimum down payment rate on Shanghai housing prices

Hongjun Li[1] · Zheng Li[2] · Cheng Hsiao[3,4]

Received: 6 May 2022 / Accepted: 12 March 2023 / Published online: 6 April 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract
In 2016, the city of Shanghai increased the minimum down payment rate requirement for purchasing various types of properties. We study the treatment effect of this major policy change on Shanghai's housing market by employing panel data from March 2009 to December 2021. Since the observed data are either in the form of no treatment or under the treatment but before and after the outbreak of COVID-19, we use the panel data approach suggested by Hsiao et al. (J Appl Econ, 27(5):705–740, 2012) to estimate the treatment effects and a time-series approach to disentangle the treatment effects and the effects of the pandemic. The results suggest that the average treatment effect on the housing price index of Shanghai over 36 months after the treatment is −8.17%. For time periods after the outbreak of the pandemic, we find no significant impact of the pandemic on the real estate price indices between 2020 and 2021.

Keywords Housing prices · Down payment rate · Pandemic · Panel data

1 Introduction

Chinese cities are categorized into four tiers, according to their size and economic influence. First-tier cities include Beijing, Shanghai, Guangzhou and Shenzhen, while second-tier cities mostly consist of provincial capitals such as Nanjing and Hangzhou.

We are grateful to Qi Li's thoughtful inputs at various stages of this work and two referees for helpful comments and suggestions. Cheng Hsiao also wishes to acknowledge partial research support by China NSF Grant No. #72033008.

✉ Zheng Li
zli42@ncsu.edu

1 Institute of Economics, School of Social Sciences, Tsinghua University, Beijing, China

2 Department of Agricultural and Resource Economics, North Carolina State University, Raleigh, NC, USA

3 Department of Economics, University of Southern California, Los Angeles, CA, USA

4 WISE, Xiamen University, Xiamen, China

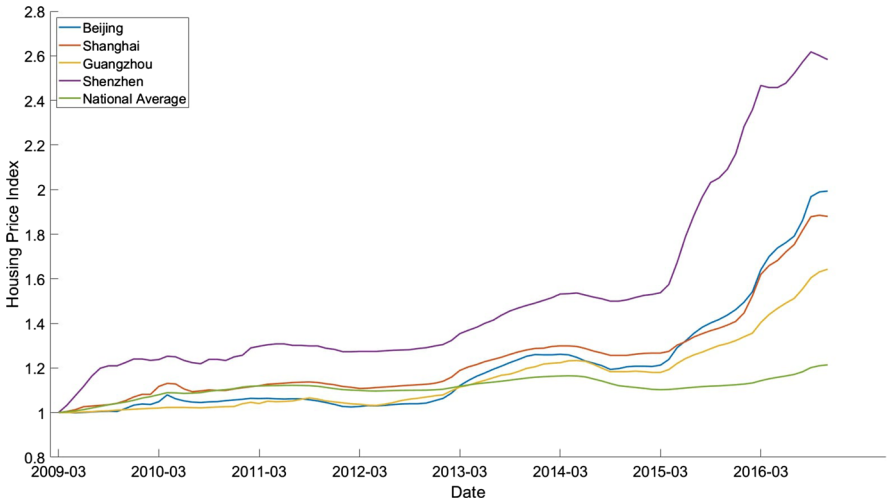

Fig. 1 Housing price index comparison among first-tier cities and national average

Third- and fourth-tier cities are smaller, often newly developed urban areas. The top-tier cities have experienced overheating in housing demand, and as a result, inflated house prices, while the lower tiers have suffered from a supply glut and a subsequent spate of unsold houses. Figure 1 shows the housing price index (normalized as one at the first time period) increases over time in first-tier cities versus the national average. As a result, the average house in Shanghai would take the average worker 41 years to pay off versus 21 years in London. Home-buyers in Shanghai, especially those whose incomes fall in the bottom 10 percent of homeowners, take on extreme financial burdens to buy homes. Their mortgages are typically more than 10 times their annual disposable income. In an effort to slow down the speculative purchasing of real estate, various policies have been implemented. For instance, Shanghai introduced a property tax in 2011 and increased the minimum down payment requirement in November 2016. COVID-19 also broke out in December 2019. This paper uses monthly data from March 2009 to December 2021 to assess the impacts of increased minimum down payment requirements and the pandemic on Shanghai's housing price index. This paper is related to studies on the mortgage and housing market (e.g.,Duan et al. 2021; Ye et al. 2014; Ricks 2021).

The traditional approach to measuring the treatment effects is to assume the observed data are either subject to treatment or not (e.g.,Rosenbaum and Rubin 1983; Abadie et al. 2010; Hsiao et al. 2012). However, our data cover the period of no treatment (before November 2016), under the treatment (November 2016 to December 2019), and under the treatment and pandemic (January 2020 to December 2021). It provides a unique possibility to investigate the impact of the pandemic and the impact of increased minimum down payment requirements on the Shanghai housing price index separately. In this paper, we are interested in measuring the treatment effects, as well as considering methods to disentangle the impact of different treatments.

The analysis relies on the Hsiao et al. (2012) (HCW) panel nonparametric regression method[1] and the time-series ARIMA model.

The HCW method has several appealing features that have been documented in the literature. Bai et al. (2014) suggested that the HCW method can be extended to non-stationary time-series data. Li and Bell (2017) showed the HCW method's robustness to regression functional form misspecification. Because of these advantages, the HCW method has been popular in analyzing treatment effects. Du and Zhang (2015) used their method to study the home-purchase restriction that was first started in Beijing in 2010, as well as the pilot property tax program implemented in Shanghai and Chongqing in 2011. They concluded that the home-purchase restriction reduced the annual growth rate of housing prices in Beijing by 7.69%. The trial property tax program reduced the annual growth rate of housing prices in Chongqing by 2.52%, while this program had no significant effects on the housing prices in Shanghai. Ke and Hsiao (2021) examined the 76-day COVID-19 lockdown policy placed on Hubei Province, the epicenter of COVID-19, in the first quarter of 2020. They concluded that this lockdown policy reduced the GDP of Hubei Province by 37% in the first quarter of 2020, and the majority of the economy quickly recovered after the government lifted the lockdown policy in April 2020. Other applications include Ouyang and Peng (2015), Ke et al. (2017), Du et al. (2021), Gao et al. (2019), and Li and Long (2018).

We present the data description in Sect. 2. Section 3 discusses the econometric modeling methodology. Section 4 provides the empirical analysis. Concluding remarks are in Sect. 5.

2 Data description

We use monthly housing price indices of 68 major Chinese cities from March 2009 to December 2021 downloaded from the Chinese National Bureau of Statistics. The housing price index for each of the cities considered is normalized to 100 for the year 2015. The treatment unit is Shanghai. We exclude the first-tier cities, Beijing, Guangzhou and Shenzhen, from the control group because they have experienced different policy changes during our sample periods. After the exclusion, the control group contains 64 cities. Table 1 presents the list of all cities in our dataset.

Shanghai in November 2016 raised the minimum down payment rate requirement from 30% to 35% for primary residence and from 60% to 70% for investment property. We let T denote the total number of time periods in the dataset. Let T_0 denote November 2016. Let T_1 denote December 2019, the date COVID-19 broke out. Thus, for $t = 1, 2, \cdots, T_0$, observed data contain neither treatment nor pandemic. For $t = T_0 + 1, T_0 + 2, \cdots, T_1$, Shanghai was under treatment, but in the absence of the pandemic, while other units received no treatment, nor under the pandemic. For $t = T_1 + 1, T_1 +$

[1] The term "nonparametric" used here is different from the conventional usages of "nonparametric" methods where a causal interpretation between y and $\boldsymbol{x}$ are assumed (e.g., Li and Racine 2007). Here, the "nonparametric" is to mean that the relation between y and $\boldsymbol{x}$ need not be "causal."

Table 1 City list

Beijing (excluded)	Xiamen	Mudanjiang	Changde
Shanghai (treatment unit)	Nanchang	Wuxi	Shaoguan
Guangzhou (excluded)	Jinan	Xuzhou	Zhanjiang
Shenzhen (excluded)	Qingdao	Yangzhou	Huizhou
Tianjin	Zhengzhou	Wenzhou	Nanning
Shijiazhuang	Wuhan	Jinhua	Guilin
Taiyuan	Changsha	Bengbu	Beihai
Hohhot	Chongqing	Quanzhou	Haikou
Shenyang	Chengdu	Jiujiang	Sanya
Dalian	Xi'an	Ganzhou	Luzhou
Changchun	Lanzhou	Yantai	Nanchong
Harbin	Tangshan	Jining	Guiyang
Nanjing	Qinhuangdao	Luoyang	Zunyi
Hangzhou	Baotou	Pingdingshan	Kunming
Ningbo	Dandong	Yichang	Xining
Hefei	Jinzhou	Xiangyang	Yinchuan
Fuzhou	Jilin	Yueyang	Urumqi

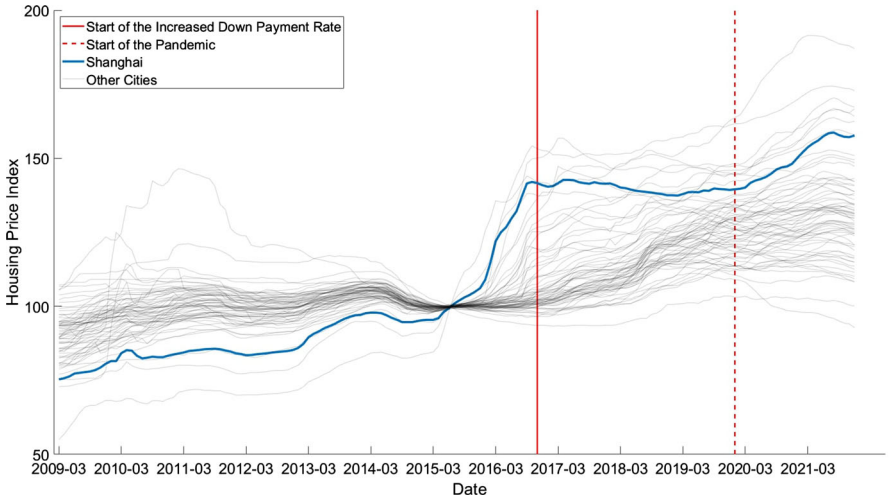

Fig. 2 Housing price indices of 68 cities

$2, \cdots, T$, all units are under pandemic, while Shanghai remained under increased minimum down payment requirement policy, but not for other units. Figure 2 shows the housing price indices of all 68 cities.

3 Methodology

To measure the treatment effects, we need to know the outcomes with or without receiving the treatment. However, rarely one has information about both. In other words, the counterfactuals for the missing outcomes need to be constructed in order to estimate the treatment effects. There are numerous methods being suggested (e.g., Chapter12, Hsiao 2022). In this paper, we adopt the panel nonparametric regression method suggested by Hsiao et al. (2012) (HCW).

Let $(y_t, x_{1t}, x_{2t}, \cdots, x_{Nt})' = (y_t, \boldsymbol{x}_t')'$ be the $(N+1)$ random vectors, $t = 1, 2, \cdots, T$. If y_t and $\boldsymbol{x}_t$ are cross-correlated, we can decompose y_t as

$$y_t = E(y_t|\boldsymbol{x}_t) + \eta_t, \tag{3.1}$$

where by construction $\eta_t \perp \boldsymbol{x}_t$. Hence one may predict y_t by $E(y_t|\boldsymbol{x}_t)$. HCW suggest to approximate $E(y_t|\boldsymbol{x}_t)$ by a linear function of $\boldsymbol{x}_t$, $\mu + \boldsymbol{a}'\boldsymbol{x}_t$. However, $(\mu, \boldsymbol{a}')$ is unknown. It needs to be estimated. To balance the closeness of within-sample fit and the accuracy of post-sample prediction, various methods are suggested. For instance, HCW suggest using model selection criteria such as AIC (Akaike 1973, 1974) or AICC (Hurvich and Tsai 1989), Li and Bell (2017) suggest using Lasso (Tibshirani 1996), and Shi and Huang (2021) suggest the forward step-wise procedure.

To predict y_t in the absence of higher down payment requirement and pandemic, we use data before T_0. We select the relevant subset of $\boldsymbol{x}_t$, $\boldsymbol{x}_t^*$, following Shi and Huang (2021) forward step-wise procedure. We start from a null model which includes no unit in the control group. At each step, we select one unit from the candidate control units and add it to the model. The model with the selected unit has the best in-sample fit compared with models with other candidate control units. The in-sample goodness of fit is measured using mean squared errors (MSE),

$$\frac{1}{T_0}\sum_{t=1}^{T_0}(y_t - \hat{y}_t)^2, \tag{3.2}$$

where $\hat{y}_t$ is the estimate of $E(y_t|\boldsymbol{x}_t)$.

The forward step-wise selection generates a path of models from a model with only one control unit to a model with all of the candidate control units. To evaluate the performance of each model on the path, we use out-of-sample forecasting accuracy as the criterion. Suppose we use the first k months as a training set and forecast the next h months ($k+1$-th to $k+h$-th month). Denote $\hat{y}_{t,k}^m$ as the counterfactual from models trained with the first k time periods of the dataset using m out of N control units,

$$\hat{y}_{t,k}^m = \hat{\beta}_0 + \hat{\beta}_1 x_{1t} + \cdots + \hat{\beta}_m x_{mt}. \tag{3.3}$$

We calculate the forecasting MSE as

$$\frac{1}{h}\sum_{t=k+1}^{k+h}(y_t - \hat{y}_{t,k}^m)^2 \tag{3.4}$$

The model with the smallest forecasting MSE is selected as the best model out of all models on the path. In the following procedures, we denote $\hat{y}_t$ as the counterfactual generated by the best model estimated with the pre-treatment data.

Under the assumption that treatment or the absence of treatment in the treatment units has no effects on the control units, the so-identified model can be used to predict the outcomes of the treatment unit in the absence of treatment for the post-treatment periods. Although COVID-19 broke out in December 2019, as long as the impact of COVID-19 does not affect the fundamental relations between the treatment unit and control units in the absence of treatment, Equation (3.3) can still be used to predict the outcomes in the absence of treatment under the pandemic because all units are under the pandemic. Since the observed y_t are the outcomes under the joint impact of the treatment and the pandemic, the treatment effects for the post-treatment periods can still be estimated as

$$\Delta_t^1 = y_t - \hat{y}_t, \text{ for } t = T_0 + 1, T_0 + 2, \cdots, T, \tag{3.5}$$

independent of whether the pandemic broke out or not in December 2019.

The observed housing price indices in the post-pandemic period represent the outcomes under both the pandemic and the increased minimum down payment requirement. They provide the possibility to separately identify the impact of the pandemic, provided that the counterfactuals in the absence of the pandemic and treatment can be constructed. One way to predict such outcomes is to use pre-treatment and pre-pandemic data $(1, 2, \cdots, T_0)$ to identify and estimate a univariate time-series model, say the Box and Jenkins (1976) type autoregressive integrated moving average model (ARIMA) for the units considered, then use the identified model to predict the outcomes for periods $T_1+1, T_1+2, \cdots, T$. However, the treatment started in November 2016 and the pandemic broke out in December 2019. The constructed pre-treatment and pre-pandemic time-series models would not make use of the most recent 37 time-series observations before the breakout of COIVD-19. Moreover, it is well known that a univariate (ARIMA) model in general predicts short-run outcomes well, but not necessarily so for long-run predictions. In our case, the prediction involves predicting 38, 39, $\cdots$, periods ahead.

To increase the degrees of freedom and reduce the prediction horizon, we note that Eq. (3.3) generates predictions in the absence of treatment for post-pandemic periods because $\boldsymbol{x}_t$ are the outcomes under pandemic but in the absence of treatment. If we replace $\boldsymbol{x}_t$ by the outcomes in the absence of treatment and pandemic, denoted as $\tilde{\boldsymbol{x}}_t$, then

$$\tilde{y}_t = \hat{\beta}_0 + \hat{\beta}_1 \tilde{x}_{1t} + \cdots + \hat{\beta}_m \tilde{x}_{mt}, \quad t = T_1 + 1, T_1 + 2, \cdots, T, \tag{3.6}$$

can give us the counterfactual paths in the absence of treatment and pandemic. We note that $\hat{y}_t$ for $t = T_1 + 1, T_1 + 2, \cdots, T$, are the predicted outcomes in the absence of treatment but under the pandemic, then the pandemic impact can be estimated as

$$\Delta_t^2 = \hat{y}_t - \tilde{y}_t, \quad t = T_1 + 1, T_1 + 2, \cdots, T. \tag{3.7}$$

Since COVID-19 started in December 2019, all control units before this date were not subject to treatment or pandemic, we suggest using Box and Jenkins (1976) identifi-

Table 2 The parameter estimates of two HCW forecasting models

	Model 1	Model 2
Changchun	−1.069***	−0.830***
	(0.226)	(0.223)
Nanjing	0.466***	0.474***
	(0.0686)	(0.0721)
Hefei	0.364***	0.392***
	(0.0438)	(0.0450)
Xiamen	−0.0476	−0.0702
	(0.0722)	(0.0755)
Baotou	−0.118	−0.00996
	(0.212)	(0.219)
Jinzhou	−0.741***	−0.699***
	(0.0874)	(0.0908)
Yichang	0.353***	0.463***
	(0.115)	(0.115)
Shaoguan	0.712***	0.743***
	(0.107)	(0.113)
Beihai	0.393***	
	(0.128)	
Luzhou	0.434**	0.780***
	(0.168)	(0.131)
Nanchong	−0.164	−0.550**
	(0.287)	(0.271)
Kunming	0.304***	0.358***
	(0.0960)	(0.0992)
Constant	11.56	−4.726
	(9.230)	(7.957)

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

cation, estimation and prediction method to predict $\tilde{x}_{it}$ for $t = T_1 + 1, T_1 + 2, \cdots, T$, then use Eq. (3.6) to construct $\tilde{y}_t$ for $t = T_1 + 1, T_1 + 2, \cdots, T$. The training dataset contains all pre-pandemic time periods, including the 37 time periods between the treatment and the pandemic. Therefore, the prediction involves predicting 1, 2, $\cdots$, periods ahead.

4 Empirical results

4.1 The treatment effect before the pandemic

We first consider the time periods before the pandemic of 2019. Let y_t be the treatment city's housing price index at time period t, and x_{it} be the i-th control city's housing

Table 3 The comparison between actual and HCW predicted values with forecasting window of 12

Date	Actual	HCW predicted	Difference	Prediction standard error	Difference in percentage
2016–11	141.60	144.07	−2.47	0.98	−1.71
2016–12	140.90	145.60	−4.70	1.08	−3.23
2017–01	140.40	146.05	−5.65	1.10	−3.87
2017–02	140.60	145.90	−5.30	1.08	−3.63
2017–03	141.60	146.42	−4.82	1.28	−3.29
2017–04	142.70	147.83	−5.13	1.29	−3.47
2017–05	142.70	149.17	−6.47	1.30	−4.34
2017–06	142.60	150.28	−7.68	1.35	−5.11
2017–07	141.90	151.15	−9.25	1.37	−6.12
2017–08	141.60	151.26	−9.66	1.39	−6.39
2017–09	141.40	151.08	−9.68	1.38	−6.41
2017–10	141.90	150.71	−8.81	1.44	−5.85
2017–11	141.50	150.39	−8.89	1.39	−5.91
2017–12	141.40	150.65	−9.25	1.34	−6.14
2018–01	141.50	149.85	−8.35	1.36	−5.57
2018–02	140.90	150.25	−9.35	1.38	−6.22
2018–03	140.10	150.12	−10.02	1.45	−6.67
2018–04	139.90	150.38	−10.48	1.55	−6.97
2018–05	139.40	150.93	−11.53	1.54	−7.64
2018–06	139.00	150.01	−11.01	1.70	−7.34
2018–07	138.80	152.28	−13.48	1.70	−8.85
2018–08	138.60	155.45	−16.85	2.00	−10.84
2018–09	138.40	158.43	−20.03	2.19	−12.64
2018–10	138.10	158.60	−20.50	2.27	−12.92
2018–11	137.90	158.14	−20.24	2.25	−12.80
2018–12	137.50	157.68	−20.18	2.26	−12.80
2019–01	137.50	157.75	−20.25	2.26	−12.83
2019–02	137.40	158.20	−20.80	2.20	−13.15
2019–03	137.90	158.17	−20.27	2.09	−12.81
2019–04	138.50	157.21	−18.71	2.06	−11.90
2019–05	138.60	157.24	−18.64	2.10	−11.85
2019–06	138.50	156.51	−18.01	2.22	−11.51
2019–07	139.10	157.00	−17.90	2.20	−11.40
2019–08	139.00	157.05	−18.05	2.31	−11.49
2019–09	139.80	155.81	−16.01	2.55	−10.27
2019–10	139.60	155.28	−15.68	2.67	−10.10
Average			−12.61		−8.17

Table 4 The comparison between actual and HCW predicted values with forecasting window of 24

Date	Actual	HCW predicted	Difference	Prediction standard error	Difference in percentage
2016–11	141.60	144.27	−2.67	1.03	−1.85
2016–12	140.90	145.68	−4.78	1.13	−3.28
2017–01	140.40	146.32	−5.92	1.16	−4.05
2017–02	140.60	146.00	−5.40	1.13	−3.70
2017–03	141.60	146.58	−4.98	1.34	−3.40
2017–04	142.70	147.68	−4.98	1.36	−3.37
2017–05	142.70	148.30	−5.60	1.33	−3.78
2017–06	142.60	148.97	−6.37	1.34	−4.27
2017–07	141.90	149.51	−7.61	1.33	−5.09
2017–08	141.60	149.33	−7.73	1.30	−5.18
2017–09	141.40	149.14	−7.74	1.29	−5.19
2017–10	141.90	148.53	−6.63	1.32	−4.47
2017–11	141.50	148.60	−7.10	1.33	−4.78
2017–12	141.40	149.29	−7.89	1.33	−5.29
2018–01	141.50	148.40	−6.90	1.34	−4.65
2018–02	140.90	149.16	−8.26	1.40	−5.54
2018–03	140.10	149.08	−8.98	1.48	−6.02
2018–04	139.90	149.27	−9.37	1.59	−6.28
2018–05	139.40	150.45	−11.05	1.61	−7.35
2018–06	139.00	149.76	−10.76	1.79	−7.18
2018–07	138.80	152.57	−13.77	1.79	−9.02
2018–08	138.60	155.31	−16.71	2.10	−10.76
2018–09	138.40	158.22	−19.82	2.30	−12.53
2018–10	138.10	158.87	−20.77	2.39	−13.07
2018–11	137.90	158.91	−21.01	2.36	−13.22
2018–12	137.50	158.86	−21.36	2.34	−13.45
2019–01	137.50	159.10	−21.60	2.33	−13.58
2019–02	137.40	159.60	−22.20	2.26	−13.91
2019–03	137.90	159.78	−21.88	2.13	−13.69
2019–04	138.50	158.78	−20.28	2.10	−12.77
2019–05	138.60	158.72	−20.12	2.15	−12.68
2019–06	138.50	157.67	−19.17	2.30	−12.16
2019–07	139.10	158.24	−19.14	2.27	−12.09
2019–08	139.00	158.29	−19.29	2.40	−12.19
2019–09	139.80	157.93	−18.13	2.58	−11.48
2019–10	139.60	157.57	−17.97	2.70	−11.41
Average			−12.61		−8.13

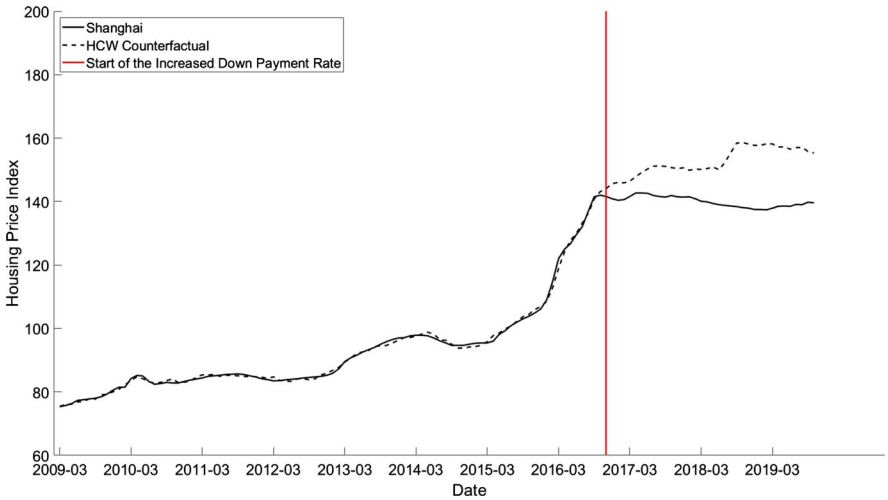

Fig. 3 Counterfactual estimated by the HCW method with forecasting window of 12

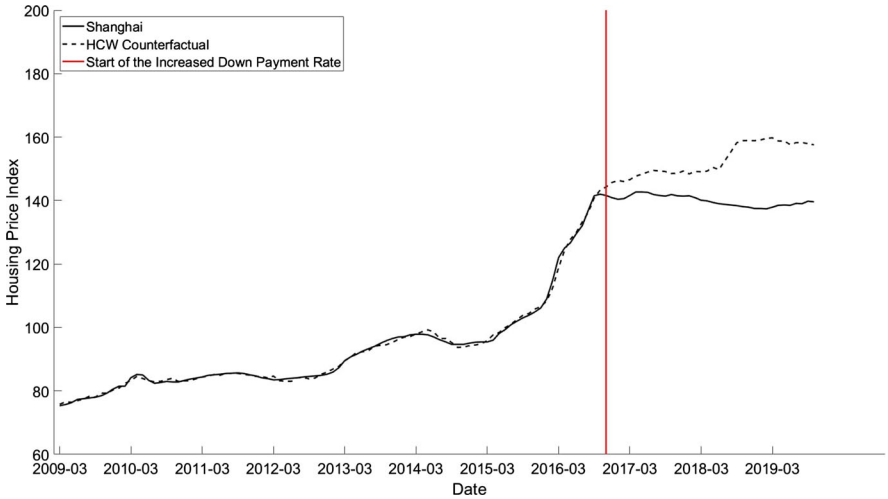

Fig. 4 Counterfactual estimated by the HCW method with forecasting window of 24

price index at time period t. Denote the number of control cities as N. We use the Shi and Huang (2021) forward step-wise selection rule to select the model to generate counterfactual paths for $\hat{y}_t, t = T_0+1, T_0+2, \cdots, T$. We use two sets of time windows to measure the out-of-sample forecasting windows, $h = 12$ or 24. For $h = 12$, we use $k = 40, 50, 60, 70, 80$ months as a training set and forecast the next 12 months ($k+1$-th to $k+12$-th month). Denote $\hat{y}_{t,k}$ as the counterfactual from models trained with the first k time periods of the dataset. We calculate the forecasting MSE as

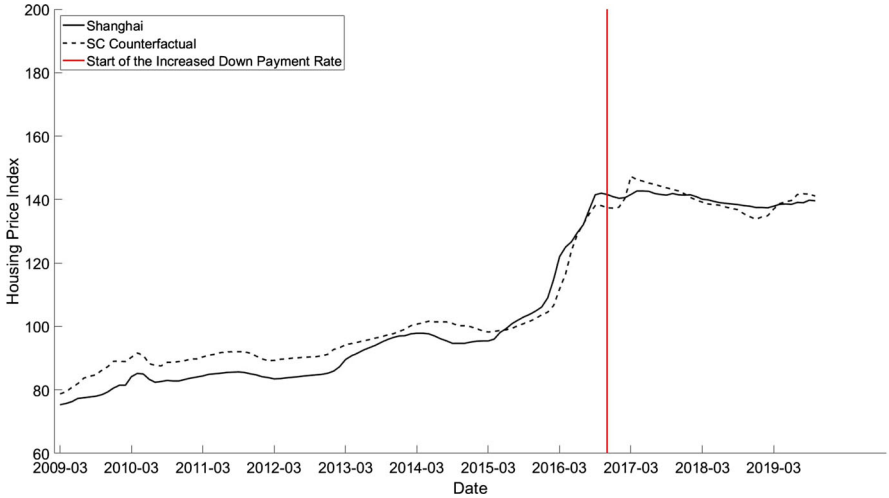

Fig. 5 Counterfactual estimated by the synthetic control method

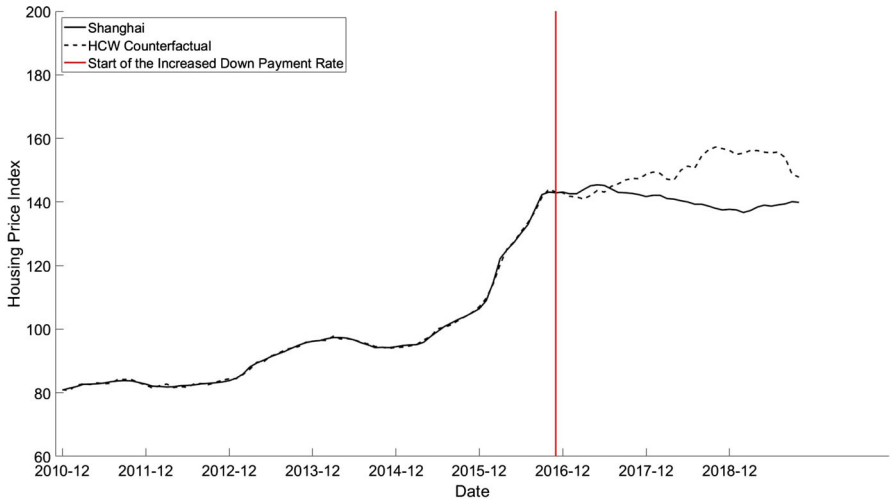

Fig. 6 Counterfactual estimated by the HCW method for small-sized houses

$$\sum_{k\in\{40,50,60,70,80\}} \frac{1}{12} \sum_{t=k+1}^{k+12} (y_t - \hat{y}_{t,k})^2 \tag{4.1}$$

The model with the smallest forecasting MSE is selected as the best model out of all models on the path. The best model contains 12 cities: Changchun, Nanjing, Hefei, Xiamen, Baotou, Jinzhou, Yichang, Shaoguan, Beihai, Luzhou, Nanchong, and Kunming. For $h = 24$, we use the first $k = 40$, 50, 60 months as a training set and forecast the next 24 months ($k + 1$-th to $k + 24$-th month). We calculate the forecasting MSE as

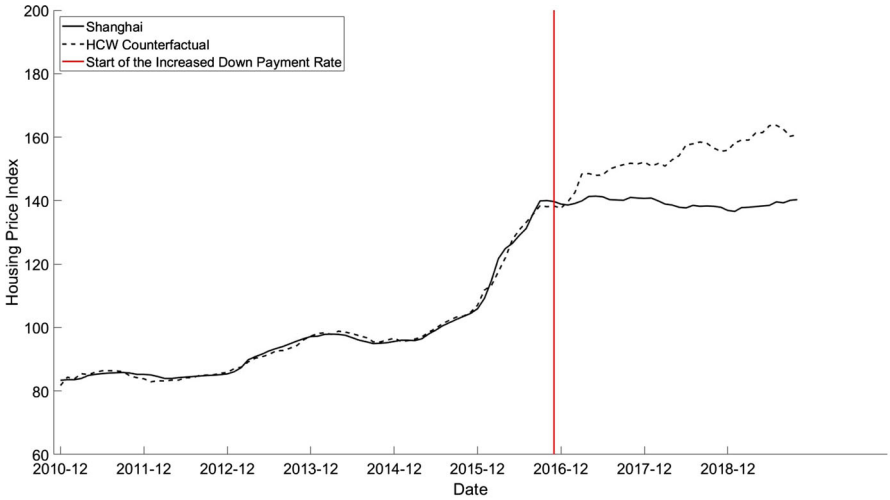

Fig. 7 Counterfactual estimated by the HCW method for medium-sized houses

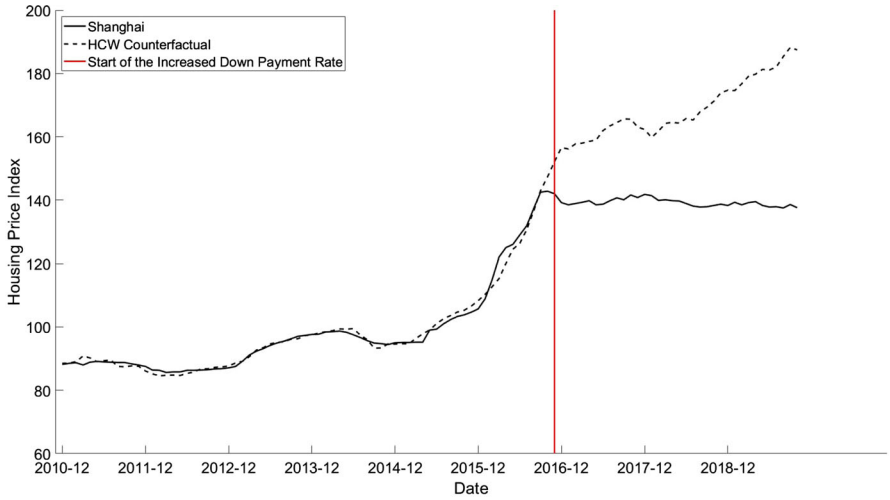

Fig. 8 Counterfactual estimated by the HCW method for large-sized houses

$$\sum_{k\in\{40,50,60\}} \frac{1}{24} \sum_{t=k+1}^{k+24} (y_t - \hat{y}_{t,k})^2 \tag{4.2}$$

With these settings, the best model contains 11 cities: Changchun, Nanjing, Hefei, Xiamen, Baotou, Jinzhou, Yichang, Shaoguan, Luzhou, Nanchong, and Kunming. These 11 cities are identical to the control cities selected based on Eq. (4.1) except that the city "Beihai" is removed. It is worth noting that the selected control units are unlikely to be affected by the policy spillover because they are all far away from Shanghai and not experiencing the overheating as Shanghai.

Table 5 The comparison between actual and HCW predicted values for small-sized houses

Date	Actual	HCW predicted	Difference	Prediction standard error	Difference in percentage
2016–11	142.90	143.17	−0.27	0.99	−0.19
2016–12	143.10	142.92	0.18	1.10	0.12
2017–01	142.60	141.81	0.79	1.26	0.56
2017–02	142.60	141.56	1.04	1.26	0.74
2017–03	143.90	140.95	2.95	1.57	2.09
2017–04	145.10	142.01	3.09	1.71	2.17
2017–05	145.40	143.68	1.72	1.86	1.20
2017–06	145.20	143.11	2.09	2.38	1.46
2017–07	144.10	144.94	−0.84	2.62	−0.58
2017–08	143.00	145.73	−2.73	2.95	−1.87
2017–09	142.90	146.89	−3.99	2.99	−2.72
2017–10	142.70	147.41	−4.71	3.04	−3.19
2017–11	142.30	147.34	−5.04	3.03	−3.42
2017–12	141.70	148.68	−6.98	3.07	−4.70
2018–01	142.10	149.42	−7.32	3.15	−4.90
2018–02	142.10	149.01	−6.91	2.98	−4.64
2018–03	141.10	147.18	−6.08	3.22	−4.13
2018–04	140.90	146.86	−5.96	3.28	−4.06
2018–05	140.40	150.00	−9.60	3.19	−6.40
2018–06	140.00	151.30	−11.30	3.19	−7.47
2018–07	139.30	150.66	−11.36	3.77	−7.54
2018–08	139.30	154.49	−15.19	4.37	−9.84
2018–09	138.70	156.42	−17.72	4.51	−11.33
2018–10	138.00	157.28	−19.28	4.96	−12.26
2018–11	137.50	156.80	−19.30	5.38	−12.31
2018–12	137.70	156.19	−18.49	5.52	−11.84
2019–01	137.50	154.98	−17.48	5.47	−11.28
2019–02	136.70	155.36	−18.66	5.33	−12.01
2019–03	137.30	156.18	−18.88	5.63	−12.09
2019–04	138.40	156.19	−17.79	5.94	−11.39
2019–05	139.00	155.67	−16.67	6.08	−10.71
2019–06	138.70	155.46	−16.76	6.36	−10.78
2019–07	139.10	155.71	−16.61	6.37	−10.67
2019–08	139.40	153.99	−14.59	7.11	−9.47
2019–09	140.10	148.73	−8.63	7.97	−5.80
2019–10	139.90	147.82	−7.92	8.34	−5.36
Average			−8.76		−5.68

Table 6 The comparison between actual and HCW predicted values for medium-sized houses

Date	Actual	HCW predicted	Difference	Prediction standard error	Difference in percentage
2016–11	139.70	138.27	1.43	1.31	1.03
2016–12	138.90	137.62	1.28	1.34	0.93
2017–01	138.60	139.68	−1.08	1.34	−0.77
2017–02	139.10	142.50	−3.40	1.34	−2.38
2017–03	139.90	148.35	−8.45	1.37	−5.70
2017–04	141.30	148.52	−7.22	1.38	−4.86
2017–05	141.40	147.95	−6.55	1.42	−4.43
2017–06	141.20	148.08	−6.88	1.49	−4.65
2017–07	140.30	149.95	−9.65	1.52	−6.44
2017–08	140.20	150.68	−10.48	1.56	−6.96
2017–09	140.10	151.33	−11.23	1.59	−7.42
2017–10	141.00	151.77	−10.77	1.62	−7.09
2017–11	140.80	151.55	−10.75	1.63	−7.10
2017–12	140.70	152.08	−11.38	1.67	−7.48
2018–01	140.80	150.94	−10.14	1.68	−6.72
2018–02	139.90	151.66	−11.76	1.72	−7.75
2018–03	138.90	150.89	−11.99	1.73	−7.95
2018–04	138.60	152.86	−14.26	1.92	−9.33
2018–05	137.90	154.16	−16.26	2.08	−10.55
2018–06	137.70	157.52	−19.82	2.38	−12.58
2018–07	138.50	157.88	−19.38	2.46	−12.28
2018–08	138.20	158.46	−20.26	2.63	−12.79
2018–09	138.30	158.04	−19.74	2.77	−12.49
2018–10	138.20	156.55	−18.35	2.84	−11.72
2018–11	137.90	155.51	−17.61	2.87	−11.32
2018–12	136.90	155.86	−18.96	2.90	−12.16
2019–01	136.60	158.17	−21.57	2.99	−13.64
2019–02	137.80	159.10	−21.30	3.02	−13.39
2019–03	137.90	159.05	−21.15	2.98	−13.30
2019–04	138.10	161.26	−23.16	3.07	−14.36
2019–05	138.30	161.43	−23.13	3.11	−14.33
2019–06	138.50	163.62	−25.12	3.18	−15.35
2019–07	139.60	163.72	−24.12	3.07	−14.73
2019–08	139.30	162.51	−23.21	2.94	−14.28
2019–09	140.10	160.25	−20.15	2.75	−12.57
2019–10	140.30	160.82	−20.52	2.76	−12.76
Average			−14.36		−9.16

Table 7 The Comparison between actual and HCW predicted values for large-sized houses

Date	Actual	HCW predicted	Difference	Prediction standard error	Difference in percentage
2016–11	142.00	152.11	−10.11	1.89	−6.64
2016–12	139.20	156.58	−17.38	1.95	−11.10
2017–01	138.50	156.14	−17.64	1.96	−11.30
2017–02	138.90	157.72	−18.82	1.96	−11.93
2017–03	139.30	157.99	−18.69	1.98	−11.83
2017–04	139.80	158.54	−18.74	1.98	−11.82
2017–05	138.50	158.95	−20.45	1.99	−12.87
2017–06	138.70	161.97	−23.27	2.06	−14.37
2017–07	139.80	163.43	−23.63	2.15	−14.46
2017–08	140.70	164.51	−23.81	2.21	−14.47
2017–09	140.10	165.66	−25.56	2.34	−15.43
2017–10	141.60	165.55	−23.95	2.35	−14.47
2017–11	140.80	163.09	−22.29	2.31	−13.67
2017–12	141.80	162.30	−20.50	2.32	−12.63
2018–01	141.40	159.78	−18.38	2.38	−11.50
2018–02	139.90	161.79	−21.89	2.43	−13.53
2018–03	140.10	164.24	−24.14	2.54	−14.70
2018–04	139.80	164.56	−24.76	2.48	−15.05
2018–05	139.70	164.30	−24.60	2.47	−14.97
2018–06	138.90	165.78	−26.88	2.51	−16.21
2018–07	138.10	165.30	−27.20	2.52	−16.46
2018–08	137.80	167.87	−30.07	2.73	−17.91
2018–09	137.90	169.36	−31.46	2.74	−18.57
2018–10	138.30	171.27	−32.97	2.75	−19.25
2018–11	138.70	173.91	−35.21	2.89	−20.24
2018–12	138.30	174.70	−36.40	3.15	−20.84
2019–01	139.30	174.59	−35.29	3.28	−20.21
2019–02	138.50	176.73	−38.23	3.39	−21.63
2019–03	139.20	179.14	−39.94	3.42	−22.30
2019–04	139.50	179.77	−40.27	3.53	−22.40
2019–05	138.30	181.30	−43.00	3.67	−23.72
2019–06	137.80	181.10	−43.30	3.82	−23.91
2019–07	137.90	182.08	−44.18	3.84	−24.26
2019–08	137.50	185.33	−47.83	4.05	−25.81
2019–09	138.60	188.21	−49.61	4.15	−26.36
2019–10	137.60	187.47	−49.87	4.17	−26.60
Average			−29.18		−17.04

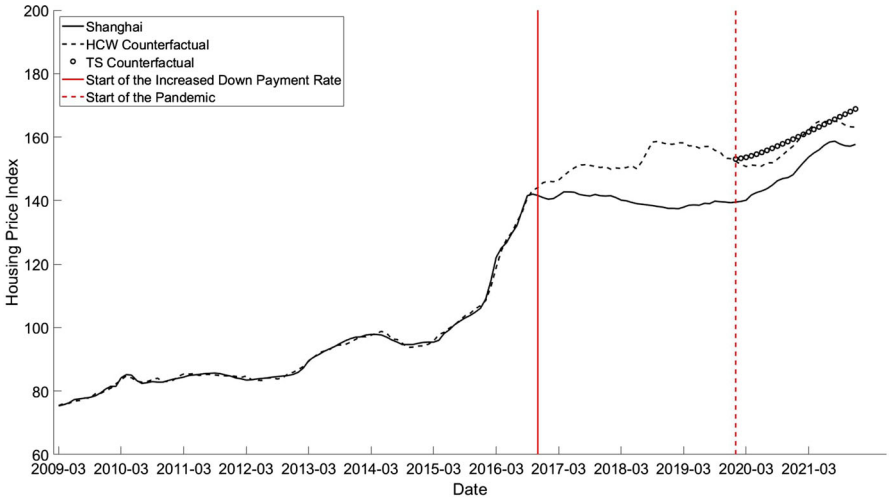

Fig. 9 Counterfactual estimated by the TS method with forecasting window of 12

We will call the method using the 12 selected control cities based on Eq. (4.1) as Model 1, and the method using the 11 selected control cities based on Eq. (4.2) as Model 2. Table 2 provides the estimation results of Model 1 and Model 2. Figures 3 and 4 plot the counterfactual estimated by the models. As one can see there is hardly any difference in the predicted counterfactual between the two models. The estimated difference between the actual and predicted counterfactual paths are reported in Tables 3 and 4, respectively. The average treatment effect is −8.17% and −8.13% on the price index over 36 months, respectively, with hardly any noticeable difference. Tables 3 and 4 also report the prediction standard errors for the HCW predicted outcomes using the formulas for constructing prediction error variance as in Fujiki and Hsiao (2015). The standard errors indicate that the differences between the actual and predicted values are significant at the 5% level.

4.2 Comparison with the synthetic control method

For comparison purposes, in addition to the HCW method, we also estimate the treatment effects before the pandemic using the synthetic control method (SCM) following Abadie and Gardeazabal (2003) and Abadie et al. (2010). Both the HCW and SCM are non-causal, nonparametric methods to construct counterfactuals. However, the HCW method places no restrictions on the coefficients of predictors while SCM restricts the coefficients to be nonnegative and sum to 1. Moreover, the HCW method allows the presence of intercept while the SCM does not. Constructing counterfactuals is essentially an issue of prediction. In principle, whatever variables are not affected by the treatment but help prediction should be included and there should be no restrictions on the sign and value of the coefficients (Hsiao and Wan 2014). Moreover, predictive models including intercepts typically yield more accurate predictions than predictive models not allowing intercepts, not to mention that intercepts in the predictive models

Table 8 The comparison between actual and TS predicted values with forecasting window of 12

Date	Actual	HCW predicted	TS predicted	Actual - HCW	HCW - TS	Actual - TS
2020–01	139.50	152.32	153.06	−12.82	−0.74	−13.56
2020–02	139.70	151.42	153.28	−11.72	−1.87	−13.58
2020–03	140.10	150.70	153.61	−10.60	−2.91	−13.51
2020–04	141.70	151.17	154.06	−9.47	−2.89	−12.36
2020–05	142.50	151.02	154.59	−8.52	−3.56	−12.09
2020–06	143.00	150.84	155.17	−7.84	−4.33	−12.17
2020–07	143.70	151.90	155.80	−8.20	−3.90	−12.10
2020–08	144.80	151.88	156.46	−7.08	−4.58	−11.66
2020–09	146.20	153.04	157.15	−6.84	−4.11	−10.95
2020–10	146.90	154.35	157.86	−7.45	−3.51	−10.96
2020–11	147.20	155.81	158.58	−8.61	−2.78	−11.38
2020–12	148.10	157.10	159.33	−9.00	−2.23	−11.23
2021–01	150.08	158.93	160.08	−8.85	−1.15	−10.00
2021–02	151.95	160.74	160.84	−8.80	−0.10	−8.90
2021–03	153.67	162.12	161.62	−8.45	0.50	−7.95
2021–04	154.96	164.10	162.40	−9.14	1.71	−7.44
2021–05	155.97	164.91	163.19	−8.94	1.72	−7.22
2021–06	157.40	164.88	163.98	−7.47	0.90	−6.58
2021–07	158.41	164.97	164.78	−6.56	0.19	−6.37
2021–08	158.70	165.03	165.58	−6.33	−0.56	−6.89
2021–09	157.83	164.79	166.39	−6.96	−1.60	−8.56
2021–10	157.26	163.77	167.21	−6.51	−3.44	−9.95
2021–11	157.12	163.23	168.03	−6.11	−4.80	−10.91
2021–12	157.69	163.13	168.85	−5.44	−5.72	−11.16
Average				−8.24	−2.07	−10.31

can be considered as capturing the unit-specific effects (Hsiao et al. 2012). Wan et al. (2018) have conducted Monte Carlo studies to compare the accuracy of predictions by the HCW and SCM methods in a number of data generating processes and found that in general HCW method yields more accurate predictions. This is also the case here.

From the counterfactuals constructed by the synthetic control method reported in Fig. 5, we observed that it does not track the treatment unit very well in the pre-treatment periods. The forecasting MSE using a forecasting window of 12 as described in Eq. (4.1) is 2440.5 for the synthetic control method, and 263.4 for the HCW method. Given its unsatisfactory performance in forecasting MSE, the estimation of the treatment effects does not seem reliable. The SCM fits the in-sample data (for $t \leq T_0$) poorly because of two reasons: (i) The SCM fitted curve (for $t \leq T_0$) has a sample mean above the real data. Since the SCM does not allow for an intercept, it cannot move the estimated curve downward to improve the fit. (ii) The SCM fitted curve has

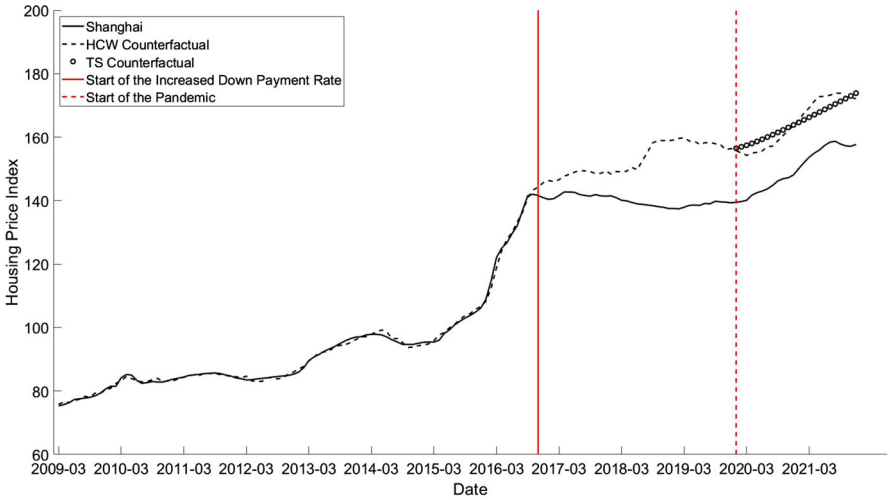

Fig. 10 Counterfactual estimated by the TS method with forecasting window of 24

a flatter slope than the real data. Therefore, the parallel trend assumption (during the pre-treatment periods) is violated. This is because the SCM weights are restricted to sum to one. Since all control cities' price curves have flatter upward trends, no SCM curve can match the upward trend of Shanghai's price index. Due to these two reasons, the SCM counterfactual is biased toward finding zero treatment effects.

4.3 Heterogeneous treatment effects on house size

The increased down payment rate suppresses the overall demand in the housing market. As suggested by a referee, there could be demand shifting among market segments. To explore heterogeneous treatment effects on house size, we obtain the segment-wise housing price indices for small (area $< 90\text{m}^2$), medium ($90\text{m}^2 <$ area $< 144\text{m}^2$) and large (area $> 144\text{m}^2$) houses. These price indices started in December 2010. Therefore, the time series is slightly shorter than the overall housing price index. We redo the estimation for the three housing market segments. The corresponding results are reported in Figs. 6, 7 and 8 and Tables 5, 6 and 7. The average treatment effects for small, medium and large houses are -5.68%, -9.16%, and -17.04%, respectively. For the small house price index, during the first 13 months after the treatment, the estimated treatment effects are mostly insignificant with mixed signs. We observe a similar phenomenon for the medium house price index but it lasts much shorter, for 3 months. After these time periods, the estimated treatment effects are negative and mostly significant at the 5% level for both the small and medium house price indices. For the large house price index, the estimated treatment effects are negative and significant from the first month after the treatment. The magnitude of the treatments is much larger than that of small and medium houses. These empirical results suggest that there is demand shifting from larger houses to smaller houses due to the increased down payment rate.

Table 9 The Comparison between actual and TS predicted values with forecasting window of 24

Date	Actual	HCW predicted	TS predicted	Actual - HCW	HCW-TS	Actual - TS
2020–01	139.50	155.73	156.49	−16.23	−0.77	−16.99
2020–02	139.70	155.04	156.94	−15.34	−1.91	−17.24
2020–03	140.10	154.25	157.43	−14.15	−3.18	−17.33
2020–04	141.70	154.96	158.02	−13.26	−3.06	−16.32
2020–05	142.50	155.21	158.65	−12.71	−3.44	−16.15
2020–06	143.00	155.65	159.32	−12.65	−3.67	−16.32
2020–07	143.70	156.98	160.03	−13.28	−3.05	−16.33
2020–08	144.80	157.19	160.76	−12.39	−3.57	−15.96
2020–09	146.20	158.85	161.51	−12.65	−2.66	−15.31
2020–10	146.90	160.49	162.27	−13.59	−1.78	−15.37
2020–11	147.20	162.25	163.05	−15.05	−0.79	−15.85
2020–12	148.10	163.50	163.84	−15.40	−0.34	−15.74
2021–01	150.08	165.49	164.64	−15.42	0.86	−14.56
2021–02	151.95	167.56	165.44	−15.61	2.11	−13.50
2021–03	153.67	169.21	166.26	−15.54	2.95	−12.59
2021–04	154.96	171.38	167.08	−16.42	4.30	−12.12
2021–05	155.97	172.77	167.91	−16.80	4.86	−11.94
2021–06	157.40	172.85	168.74	−15.45	4.11	−11.34
2021–07	158.41	173.21	169.58	−14.80	3.62	−11.17
2021–08	158.70	173.81	170.43	−15.12	3.39	−11.73
2021–09	157.83	173.87	171.28	−16.03	2.59	−13.44
2021–10	157.26	172.82	172.13	−15.56	0.69	−14.87
2021–11	157.12	172.15	172.99	−15.03	−0.84	−15.87
2021–12	157.69	172.23	173.85	−14.54	−1.62	−16.16
Average				−14.71	−0.05	−14.76

4.4 Separating the treatment effect and the pandemic effect

For the pandemic time periods, the control cities are affected by the pandemic. Using the HCW predicted $\hat{y}_t$ obtained by Eq. (3.3) reflects the counterfactual of Shanghai housing price index *without* the treatment but *under* the pandemic. To predict the counterfactuals without the pandemic and without the increase in the minimum down payment requirement, as well as to reduce the potential distortions due to ignoring the time-series information from November 2016 to December 2019, we opt to construct ARIMA models for each control unit using data from pre-December 2019, and then use the estimated models to generate predictions for the control units for the post-pandemic periods, which reflect the scenarios *without* pandemic and *without* treatment. Denote each of such predicted values by $\tilde{x}_{it}$. The ARIMA models for each $\tilde{x}_{it}$ for the used control units are reported in the appendix. We use $\tilde{x}_{it}$ to construct the counterfactual for Shanghai in the absence of pandemic and treatment using Eq. (3.6). Then the impact

Table 10 Estimation and diagnosis results of ARIMA models

	Constant ($\times 10^3$)	AR(1)	AR(2)	Q(4)	p-value of Q(4)
Changchun	0.650	0.708		1.377	0.848
	(0.341)	(0.044)			
Nanjing	0.773	0.809		6.427	0.169
	(0.589)	(0.042)			
Hefei	0.951	0.806		5.891	0.207
	(1.004)	(0.030)			
Xiamen	1.439	0.688		0.262	0.992
	(1.089)	(0.040)			
Baotou	0.408	0.506		7.735	0.102
	(0.364)	(0.064)			
Jinzhou	−0.106	0.384	0.406	1.390	0.846
	(0.307)	(0.055)	(0.063)		
Yichang	0.798	0.559		7.228	0.124
	(0.507)	(0.040)			
Shaoguan	0.663	0.553		7.164	0.127
	(0.370)	(0.062)			
Beihai	0.650	0.714		2.758	0.599
	(0.486)	(0.055)			
Luzhou	0.601	0.436	0.272	2.952	0.566
	(0.355)	(0.072)	(0.081)		
Nanchong	0.484	0.716		1.599	0.809
	(0.315)	(0.039)			
Kunming	1.597	0.479		1.450	0.835
	(0.654)	(0.060)			

of the pandemic is estimated as

$$\Delta_t^2 = \hat{y}_t - \tilde{y}_t, \text{ for } t = T_1 + 1, T_1 + 2, \cdots, T, \tag{4.3}$$

and the impact of the increased minimum down payment requirement under the pandemic is

$$\Delta_t^1 = y_t - \hat{y}_t, \text{ for } t = T_1 + 1, T_1 + 2, \cdots, T. \tag{4.4}$$

For the control group selected by the forward step-wise selection with a forecasting window of 12, Fig. 9 shows the paths of $\hat{y}_t$ (HCW counterfactual) and $\tilde{y}_t$ (TS counterfactual), and Table 8 shows the corresponding month-by-month values. Figure 10 and Table 9 report the results for the control group selected by the forward step-wise selection with a forecasting window of 24. The difference between the counterfactuals predicted by TS and HCW indicates that the pandemic has a negligible impact on Shanghai's housing price index. On the other hand, by comparing the difference between Actual and HCW after December 2019, we observe a lasting

impact of increased minimum down payment requirement. It reduces the Shanghai housing price index by about 8%.

5 Concluding remarks

Our empirical analysis showed that raising the minimum down payment requirement appeared to be an effective policy to reduce the growth of Shanghai housing prices. The policy reduced its housing price index by about 8% had such a policy not been implemented. Moreover, the effects appeared to last as long as such a policy is in place. On the other hand, the pandemic hardly had any impact on the real estate prices in Shanghai, contrary to the findings in the real sector, where Ke and Hsiao (2022) find the pandemic reduces China's outputs by about 10% because of the break out of pandemic.

Funding This study was partially funded by China NSF (Grant No. 72033008).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

A Time series models for the control units

Based on the Box-Jenkins diagnosis, we adopt ARIMA(1,1,0) or ARIMA(2,1,0) to model the log housing price indices of the control cities, $\log x_{it}$. The predicted value $\tilde{x}_{it}$ is calculated by $\exp(\widehat{\log x_{it}}) \cdot \exp(\frac{1}{2}\sigma_\epsilon^2)$, where $\widehat{\log x_{it}}$ and σ_ϵ^2 are the predicted output and estimated variance of the residual of the ARIMA model, respectively. The estimates are reported in Table 10.

References

Abadie A, Diamond A, Hainmueller J (2010) Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. J Am stat Assoc 105(490):493–505

Abadie A, Gardeazabal J (2003) The economic costs of conflict: a case study of the basque country. Am Econ Rev 93(1):113–132

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *Proc. 2nd Inter. Symposium on Information Theory*, 267–281

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Contr 19(6):716–723

Bai C, Li Q, Ouyang M (2014) Property taxes and home prices: a tale of two cities. J Econ 180(1):1–15

Box GE, Jenkins GM (1976) Time series analysis: forecasting and control. Wiley, Hoboken

Du Z, Li J, Pei P (2021) The impacts of china's exchange rate regime reform in 2005: a counterfactual analysis. Rev Devel Econ 25(1):430–448

Du Z, Zhang L (2015) Home-purchase restriction, property tax and housing price in china: a counterfactual analysis. J Econom 188(2):558–568

Duan J, Tian G, Yang L, Zhou T (2021) Addressing the macroeconomic and hedonic determinants of housing prices in Beijing metropolitan area, China. Habitat Int 113:102374
Fujiki H, Hsiao C (2015) Disentangling the effects of multiple treatments-measuring the net economic impact of the 1995 great hanshin-awaji earthquake. J Econ 186(1):66–73
Gao Y, Gan L, Li Q (2019) Chinese trade price and yuan's valuation. World Econ 42(7):2215–2243
Hsiao C (2022) Analysis of panel data, 4th edn. Cambridge University Press, Cambridge
Hsiao C, Steve Ching H, Ki Wan S (2012) A panel data approach for program evaluation: measuring the benefits of political and economic integration of hong kong with mainland china. J Appl Econ 27(5):705–740
Hsiao C, Wan SK (2014) Is there an optimal forecast combination? J Econ 178:294–309
Hurvich CM, Tsai C-L (1989) Regression and time series model selection in small samples. Biometrika 76(2):297–307
Ke X, Chen H, Hong Y, Hsiao C (2017) Do china's high-speed-rail projects promote local economy?-new evidence from a panel data approach. China Econ Rev 44:203–226
Ke X, Hsiao C (2021) Economic impact of the most drastic lockdown during covid-19 pandemic–the experience of Hubei, China. J Appl Econ 37(1):187–209
Ke X, Hsiao C (2022) Data subject to multiple treatment effects-disentangle the impacts of global pandemic and a specific disease control policy. The Singapore Economic Review, pp 1–21
Li KT, Bell DR (2017) Estimation of average treatment effects with panel data: asymptotic theory and implementation. J Econ 197(1):65–75
Li Q, Long W (2018) Do parole abolition and truth-in-sentencing deter violent crimes in virginia? Emp Econ 55(4):2027–2045
Li Q, Racine JS (2007) Nonparametric econometrics: theory and practice. Princeton University Press, Princeton
Ouyang M, Peng Y (2015) The treatment-effect estimation: a case study of the 2008 economic stimulus package of china. J Econ 188(2):545–557
Ricks JS (2021) Mortgage subsidies, homeownership, and marriage: effects of the va loan program. Reg Sci Urban Econ 87:103650
Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. Biometrika 70(1):41–55
Shi Z, Huang J (2021) Forward-selected panel data approach for program evaluation. J Econ. https://doi.org/10.1016/j.jeconom.2021.04.009
Tibshirani R (1996) Regression shrinkage and selection via the lasso. J Royal Stat Soci: Series B (Methodological) 58(1):267–288
Wan S-K, Xie Y, Hsiao C (2018) Panel data approach vs synthetic control method. Econ Lett 164:121–123
Ye G, Deng G, Li Z (2014) Mortgage rate and choice of mortgage length: a quasi-experimental evidence from chinese transaction-level data. J Hous Econ 25:96–103

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.