

The Wharton/NUS/Tsinghua
Chinese Residential Land Price Indexes (CRLPI)
White Paper

Draft of September 1st, 2014

Yongheng Deng
Institute of Real Estate Studies
NUS Business School and School of Design and Environment
National University of Singapore
ydeng@nus.edu.sg

Joseph Gyourko
The Wharton School, University of Pennsylvania and NBER
gyourko@wharton.upenn.edu

Jing Wu
Hang Lung Center for Real Estate,
Institute of Real Estate Studies, and
Department of Construction Management
Tsinghua University
ireswujing@tsinghua.edu.cn

We appreciate the advice and insight of our colleagues Professor Hongyu Liu and Dr. Siqi Zheng on various parts of this project, although we obviously are responsible for any remaining errors. Excellent research assistance was provided by Bo Zhang, Pu Wang, Wei Guo, Mingyue Li, Baoyi Hu, Jingting Huang, and Chenxi Zhao from Tsinghua University, Ying Chen, Hui Liu and Chen Zheng from Wharton, and Jia He and Mingying Xu from the National University of Singapore. Gyourko thanks the Global Research Initiatives Project of the Wharton School at the University of Pennsylvania for financial support. Deng thanks the Institute for Real Estate Studies at the National University of Singapore for financial support. Wu thanks Hang Lung Center for Real Estate at Tsinghua University and the National Natural Science Foundation of China for financial support (No. 712003060 & No. 71373006).

I. Introduction

The Wharton/NUS/Tsinghua Chinese Residential Land Price Indexes (CRLPI) are constant-quality series that track changes in the real value of residential land parcels purchased via public bidding or auction from local governments in the 35 major cities across China identified in Figure 1. Data on the market share of these 35 markets in China's aggregate land sales are not available. However, newly-built housing transactions in these cities as measured by floor area constitutes at least 30% of the national total for any year during 2004-2013 as depicted in Figure 2, with the market share in value of these transactions being around 50% to 70%.

Index values are reported at quarterly, semi-annual, and annual levels for different levels of geographic aggregation (i.e., national, regional and city) depending upon data availability and quality, as is described more fully below. This white paper provides a technical description of the underlying land parcel data and the statistical methods used to estimate the reported indices.

All are welcome to download the index data. If these data are used in any publication or presentation, a reference to the following publication should be included:

Wu, Jing, Joseph Gyourko and Yongheng Deng (2012). "Evaluating Conditions in Major Chinese Housing Markets," *Regional Science and Urban Economics*, Vol. 42, No. 2: 531-543.

II. Background on the Development of Land Markets in China

Because the Chinese government retains ultimate ownership of all urban land, the transactions prices we observe technically are for a leasehold interest. A 1988 constitutional amendment allowed the purchase of use rights for long periods of time—up to 70 years for residential properties. In a typical private housing project development, local governments first sell the leasehold interest to a developer. The developer then builds housing units on the parcels, and sell those units to households. Households have the right to live in, rent out or sell their housing units during the remaining period of the leasehold estate. Our indexes are based on the prices paid by developers for the right to use the land. Because there is a single up-front payment made by the developer to the local government and because the usage rights are for many decades (with none being close to its terminal period as yet), we treat the initial lump sum payment by the developer to the local government as the transactions price of the land parcel.

Land parcels wholly or partly designated for residential usage typically are included in our samples except when the site is wholly designated for public housing units. The pricing mechanism in those cases is very different because of government regulation of prospective prices or rents and because there can be direct governmental allocation of housing units after construction.¹

We report index values beginning in 2004 because that is the first year for which we have confidence that data quality is consistently high. Prior to that in May

¹ Public housing programs in China include low-rent units (*lian zu fang*), rent-controlled units (*gong gong zu lin fang*), affordable housing units (*jing ji shi yong fang*) and price-controlled units (*xian jia fang*). Typically the parcels designated for low-rent and rent-controlled units are directly allocated by local governments, with parcels for affordable and price-controlled units required to transact via public auction or other bidding process, but subject to government controlled pricing of the completed units.

2002, the Ministry of Land and Resources (MLR) required all residential and commercial land parcel leasehold purchases subsequent to July 2002 to be sold via public auction process (Document No. 2002-11 of MLR).² This requirement was further institutionalized in March of 2004 (Document No. 2004-71 of MLR). Residential land prices since then reasonably can be expected to reflect fair market values. Local land authorities also were required to post information about the land parcels as part of the public auction/bidding process, and these data are reliably reported from 2004-on. Finally, 2004 is the first year for which there were a sufficiently large number of transactions with which to credibly estimate hedonic price indexes.

Table 1 documents the three different levels of geographic aggregation for which we report constant-quality real land price indexes. A national-level index is reported quarterly beginning in 2004(1), and is based on land parcel transactions in each of the 35 markets we track.

Three regional-level indexes are reported at semi-annual frequency. The three regions are East, Middle, and West based on the official definitions of China's National Bureau of Statistics. There are 16 cities included in the East region: Beijing, Tianjin, Shijiazhuang, Shenyang, Dalian, Shanghai, Nanjing, Hangzhou, Ningbo, Fuzhou, Xiamen, Jinan, Qingdao, Guangzhou, Shenzhen, and Haikou. There are 8 cities included in the Middle region: Taiyuan, Changchun, Harbin, Hefei, Nanchang,

² Prior to this ruling called the 11th Provision, most transactions of urban land parcels were done by negotiation between a developer and a local government. This process was criticized for being opaque and open to corruption (Cai, Henderson and Zhang, 2009). For our purposes, the prices that resulted seem likely to be below free market levels, with the degree unknown and possibly changing over time depending upon local circumstances.

Zhengzhou, Wuhan, and Changsha. There are 11 cities included in the West region: Hohhot, Nanning, Chongqing, Chengdu, Guiyang, Kunming, Xian, Lanzhou, Xining, Yinchuan, and Urumqi.

Finally, we report annual indexes for 12 individual cities: Beijing, Tianjin, Dalian, Shanghai (since 2006), Nanjing, Hangzhou, Wuhan, Changsha, Guangzhou (since 2007), Chongqing, Chengdu, and Xian (since 2007). This choice of cities is determined by data availability and reliability. While we observe sales in each city and can estimate hedonic price indexes for them, we report results only if there were at least 15 observations per year. That hurdle was reached only in 2006 for Shanghai, and 2007 for Guangzhou and Xian.

III. Data and Index Creation Procedures

1. Data Sources

The MLR requires local land authorities to publish detailed information on each parcel's address, size, designated usage, land conditions on delivery, and major planning indicators when listing the parcel on the market. Typically, such a "listing bulletin" appears on the official website of the local land authority and/or in major local newspapers. Following the sale, the local land authority is required to publicly report the winning bidder and the transaction price in a "results bulletin" via similar channels. In some cities such as Beijing and Shanghai, the local land authority also provides a database on its website with the history of all public auction/bidding transactions.

The land transaction data in the 35 major cities underlying our indexes were first acquired from Soufun, a leading real estate data vendor in China. We then checked each residential land transaction recorded in Soufun's land database against the "listing bulletins" and "result bulletins" posted by the local land authorities and addressed any inconsistencies. We then mapped each parcel using a GIS system. These mapped data form the micro-level data base for estimation of each of our reported land price indexes.

To be included in our data base, the residential land parcel generally must be in one of a city's urban districts (*qu*). Most rural area counties (*xian*) and county-level cities (*xian ji shi*) are excluded because they are too far from the urban area of the city to be considered as within the same market. We make an exception to this rule if the land parcel in one of these outlying areas is connected to the urban area by a subway line. Table 2 lists the coverage of the database in each city.

Table 3 reports summary statistics on the number of cities covered by year from 2004 to the first half of 2014, along with the number of successful land parcel transactions and the square footage involved. There is complete data dating back to 2004 for 22 markets.³

2. Transaction Prices

³ We also observe cases in which a listed land parcel does not transact, typically because there were no bidders (*liu pai*) in the case of a public auction (either English auction or two-stage auction) or if all bid prices were lower than the local governments' reserve prices in a sealed bid process (*liu biao*). These observations clearly cannot be used in the creation of indexes, but they do allow us to test for sample selection bias. See Deng, Gyourko and Wu (2012) for more on that.

All transaction prices are put in real terms by deflating with the monthly consumer price index (CPI) series published by the local statistical authority in each city. Because it is customary in China to price residential land in terms of the floor area of housing permitted on the parcel, our prices are first put in terms of real yuan per square meter of floor area. However, we do not report these raw real prices in our index creation. They are used to create a constant quality land price index that controls for changes in land quality over time. The attributes used in that process are described in the next subsection.

3. *Land Attributes*

The following land quality characteristics are available for all land parcels in our sample, and are used in the estimation of a constant quality land price index:

- **Distance to city center:** the parcel's distance in kilometers from the center of the city in which it is located (*D_CENTER*); the logarithm of this variable is used in the hedonic equation estimation;
- **Distance to nearest subway station:** the parcel's distance in kilometers to the nearest subway station that was in use on the day the land transaction occurred (*D_SUBWAY*); this variable also is specified in log terms;⁴
- **District dummies:** district fixed effects are included to control for local/neighborhood-level fixed effects not captured by the two previous location controls; these are dichotomous dummies based on the districts reported in Table 2;

⁴ For parcels in cities without a subway system or for parcels sold before any subway line was in operation, we treat the variable as missing, but allow all other parcel traits to affect the hedonic price.

- **Land parcel size:** total land area of the parcel (*SIZE*) in square meters; this variable also is specified in logarithmic term;
- **Floor area ratio:** the ratio of floor area permitted to build to the land area of the parcel (*FAR*) is used as a site density control;
- **Commercial properties:** in some cases a small portion of a residential land parcel is designated for affiliated commercial properties (e.g., retail); in those cases, a dichotomous dummy variable (called *COMMERCIAL*) takes on a value of 1; it is equal to 0 otherwise;
- **Public establishments:** in some cases a small portion of a residential land parcel is designated for affiliated public establishments such as schools or hospitals; in those cases, a dichotomous dummy variable (called *PUBLIC*) takes on a value of 1; it is equal to 0 otherwise;
- **Requirement for public housing:** in select cities like Beijing, a dummy variable (called *PUBLICHOUSING*) is set to reflect whether the purchaser of the parcel is required to provide some public housing units on the parcel.
- **Transaction date:** the specific date of each land sale is observed; after controlling for site quality differences with the aforementioned variables, time dummies are included in the regression; it is their coefficients which form the basis for the constant quality land price indices we report.⁵

⁵ A few of the larger cities such as Beijing typically report other land characteristics such as whether the land parcel is leveled on delivery (*LEVEL*), or whether the local government provides full infrastructure for the parcel (*INFRA*). For these cities, we estimated more extensive hedonic models incorporating these additional variables. While some of these variables were statistically significant, they did not have

4. Hedonic Land Price Estimation

We follow a long literature in economics to estimate constant quality land price indexes based on the hedonic method.⁶ Each city-level index is estimated via ordinary least squares (OLS) as described in equation (1), where all land parcel observations, denoted by the subscript i , are equally weighted:

$$\ln(LP_i) = X_i\alpha + \sum_t \beta_t D_{it} + \varepsilon_i \quad (1)$$

The variable LP is the transaction price of a parcel. On the right-hand side of the equation, the vector α is the set of coefficients on the hedonic attributes X described above (i.e., $\ln(D_CENTER)$, $\ln(D_SUBWAY)$, $\ln(SIZE)$, FAR , $COMMERCIAL$, $PUBLIC$, $PUBLICHOUSING$, and the district dummies). Note that the impacts of these parcel traits do not vary over time, by assumption. The vector D includes dichotomous year dummies which equal 1 in the period when the parcel was sold, and 0 otherwise, with β representing their coefficient vector. The standard error term is denoted by ε .

The price index we report for a given city is then calculated based on the year dummy coefficients β , which capture common time effects on price after controlling for site quality in a given city. Other price indexes could be created, of course. For example, we could have priced a specific quality bundle over time--say for a parcel that was a specific distance from the city center and nearest subway line, had a certain FAR, etc. Not doing so is a matter of choice, and there is no obviously right or wrong way to

a material effect on the overall price index. For consistency across cities, we estimate a common specification that includes only those variable discussed above in the text.

⁶ The hedonic method is widely used in housing and labor economics. See Kain & Quigley (1970) and Thibodeau (1989) for widely cited applications to housing in the U.S. Wu, Deng and Liu (2014) provide an application to China.

proceed. Given our data and the extremely rapid growth and outward expansion observed in many land markets, we believe the index just described is appropriate in the Chinese context.

The same specification and functional form are used for the regional and national indexes reported below, but price indexes are reported at semi-annual frequency for the regions and at a quarterly frequency for the nation. Hence, the time dummies reflect different time periods for those estimations.⁷ When estimating these indexes, each parcel from a relevant series is pooled and equally weighted as described in equation (1). The only difference is that the time dummies reflect either semi-annual or quarterly periods, rather than years. Because we already control for district fixed effects within each city, city dummies are not included in the pooled regressions estimated at the regional or national level.⁸

While we only report the resulting real constant-quality land price indexes, it is worth noting that the coefficients on the quality controls are consistent with expectations and generally statistically significant. There is always meaningful explanatory power for the vector of right-hand side variables. See Wu, Gyourko and Deng (2012) for more detailed analysis of Beijing and Deng, Gyourko and Wu (2012) for the overall performance across the 35 major cities in our sample.

⁷ There is one further complication regarding the national index which is estimated quarterly. Due to a lag in publication of the CPI series, prices for the final quarterly reporting period typically are deflated with the previous period's CPI number. For example, we use the CPI number from the first quarter of 2014 to deflate transactions prices from the second quarter of 2014. This will be changed in subsequent postings when updated inflation data become available.

⁸ Experimentation with doing so introduces collinearity problems in some instances. In any case, it never has a material effect on the price index numbers.

Controlling for changes in land parcel quality is important, as average annualized appreciation in our national aggregate hedonic index is over five percentage points (5.6%) higher than in the unadjusted price series, which indicates that parcel quality has been falling over time on average. This does vary by time and market to some extent, but emphasizes why we report constant quality land price indexes. Not doing so would understate the true price growth in Chinese land markets by a considerable extent (e.g., $1.05^8=1.48$, so constant quality price growth has been about 50% higher than unadjusted real price growth across our 35 markets since 2004).

Finally, we note that updating the indexes over time can lead to revision of past results because we pool all observations from all periods when estimating the hedonic model (whether at the city, regional or national level).⁹ We have analyzed this issue with our present data, and revisions from adding new quarters of data have had at most a moderate impact on past index values. However, there is no guarantee that will always be the case in the future.

⁹ See Deng and Quigley (2008) for more details about this index revision issue.

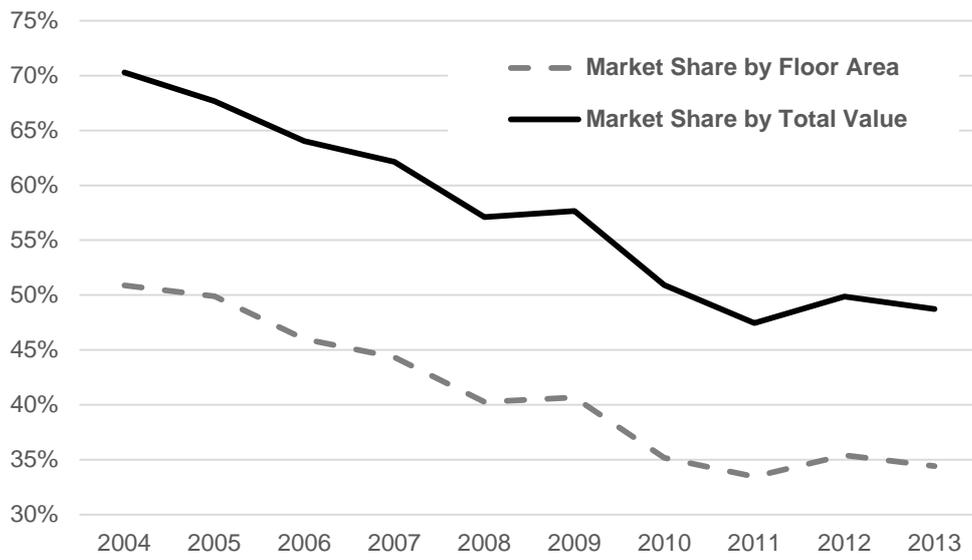
Selected References

- Cai, Hongbin, Vernon Henderson, and Qinghua Zhang (2009). "China's Land Market Auctions: Evidence of Corruption," NBER Working Paper #15067.
- Deng, Yongheng and John Quigley (2008). "Index Revision, House Price Risk, and the Market for House Price Derivatives," *Journal of Real Estate Finance and Economics*, Vol. 37, No. 3: 191-209.
- Deng, Yongheng, Joseph Gyourko and Jing Wu (2012). "Land and House Price Measurement in China", in *Property Markets and Financial Stability*, A. Heath, F. Packer and C. Windsor eds., Reserve Bank of Australia.
- Kain, John and John Quigley. (1970). "Measuring the Value of Housing Quality," *Journal of the American Statistical Association*, Vol. 65, No. 330: 532-548.
- Thibodeau, Thomas. (1989). "Housing Price Indexes from the 1974-1983 SIMSA Annual Housing Surveys," *AREUEA Journal*, Vol. 17, No. 1: 100-117.
- Wu, Jing, Joseph Gyourko and Yongheng Deng (2012). "Evaluating Conditions in Major Chinese Housing Markets," *Regional Science and Urban Economics*, Vol. 42, No. 2: 531-543.
- Wu, Jing, Yongheng Deng and Hongyu Liu (2014). "House Price Index Construction in the Nascent Housing Market: The Case of China," *Journal of Real Estate Finance and Economics*, Vol. 48, No. 3: 522-545.

Figure 1: 35 Major Cities Covered



Figure 2: Market Share of 35 Major Housing Markets in China



Source: authors' calculations based on statistics published by National Bureau of Statistics, China.

Table 1: Components in the Index System

Index		Frequency	Starting Point
National Level		Quarterly	First Quarter, 2004
Regional Level	East	Semi-Annually	First Half, 2004
	Middle	Semi-Annually	First Half, 2004
	West	Semi-Annually	First Half, 2004
City Level	Beijing	Annually	2004
	Tianjin	Annually	2004
	Dalian	Annually	2004
	Shanghai	Annually	2006
	Nanjing	Annually	2004
	Hangzhou	Annually	2004
	Wuhan	Annually	2004
	Changsha	Annually	2004
	Guangzhou	Annually	2007
	Chongqing	Annually	2004
	Chengdu	Annually	2004
	Xian	Annually	2007

Note: the east region includes Beijing, Tianjin, Shijiazhuang, Shenyang, Dalian, Shanghai, Nanjing, Hangzhou, Ningbo, Fuzhou, Xiamen, Jinan, Qingdao, Guangzhou, Shenzhen, and Haikou (16 cities); the middle region includes Taiyuan, Changchun, Harbin, Hefei, Nanchang, Zhengzhou, Wuhan, and Changsha (8 cities); west region includes Hohhot, Nanning, Chongqing, Chengdu, Guiyang, Kunming, Xian, Lanzhou, Xining, Yinchuan, Urumqi (11 cities). The national level index includes all the 35 cities.

Table 2: Coverage of Districts within Each City

City	Districts/Counties Covered
Beijing	Changping, Chaoyang, Daxing, Dongcheng, Fangshan, Fengtai, Haidian, Shijingshan, Shunyi, Tongzhou, Xicheng
Tianjin	Beichen, Dongli, Hebei, Hedong, Heping, Hexi, Hongqiao, Jinan, Nankai, Tanggu, Xiqing
Shijiazhuang	Qiaodong, Qiaoxi, Xinhua, Yuhua, Changan
Taiyuan	Jiancaoping, Jinyang, Wanbolin, Xiaodian, Xinghualing, Yingze
Hohhot	Huimin, Saihan, Xincheng, Yuquan
Shenyang	Dadong, Dongling, Heping, Huanggu, Hunnan, Shenhe, Sujiatun, Tiexi, Yuhong
Dalian	Ganjingzi, Gaoxinyuan, Jinzhou, Kaifaqu, Shahekou, Xigang, Zhongshan
Changchun	Chaoyang, Erdao, Gaoxin, Jingkai, Jingyue, Kuancheng, Lvyuan, Nanguan, Qichekaifa
Harbin	Daoli, Daowai, Dongli, Hulan, Nangang, Pingfang, Songbei, Xiangfang
Shanghai	Baoshan, Changning, Fengxian, Hongkou, Jiading, Luwan, Minhang, Nanhui, Pudong, Putuo, Qingpu, Songhui, Xuhui, Yangpu, Zabei
Nanjing	Baixia, Gulou, Jiangning, Jianye, Pukou, Qinhuai, Qixia, Xiaguan, Xuanwu, Yuhuatai
Hangzhou	Binjiang, Gongshu, Jianggan, Shangcheng, Xiacheng, Xiashan, Xihu, Yuhang
Ningbo	Haishu, Jiangbei, Jiangdong, Yinzhou
Hefei	Baihe, Binhu, Gaoxin, Kaifa, Luyang, Shushan, Xinzhan, Yaohai
Fuzhou	Cangshan, Gulou, Jinan, Taijiang
Xiamen	Haicang, Huli, Jimei, Siming, Xiangan
Nanchang	Donghu, Honggutan, Kaifa, Nanchang, Qingshanhu, Qingyunpu, Wanli, Xihu, Xinchang
Jinan	Gaoxin, Huaiyin, Licheng, Lixia, Shizhong, Tianqiao
Qingdao	Laoshan, Licang, Shibe, Shinan, Sifang
Zhengzhou	Erqi, Gaoxin, Guancheng, Huiji, Jinshui, Zhengdong, Zhongyuan
Wuhan	Donghu, Dongxihu, Hanyang, Hongshan, Jiangan, Jianggan, Kaifa, Qiaokou, Qingshan, Wuchang
Changsha	Furong, Kaifu, Tianxin, Yuelu, Yuhua
Guangzhou	Baiyun, Fanyu, Haizhu, Huangpu, Liwan, Nansha, Tianhe, Yuexiu
Shenzhen	Baoan, Futian, Longgang, Nanshan, Yantian
Nanning	Yongning, Jiangnan, Liangqing, Qingxiu, Xingning, Xixiangtang
Haikou	Longhua, Meilan, Qionghua, Xiuying
Chongqing	Banan, Dadukou, Jiangbei, Jiulongpo, Nanan, Shapingba, Yubei, Yuzhong
Chengdu	Chenghua, Gaoxin, Gaoxin, Jinjiang, Jinniu, Longquan, Pixian, Qingyang, Wuhou
Guiyang	Baiyun, Huaxi, Jinyang, Nanming, Wudang, Xiaohe, Yunyan
Kunming	Guandu, Panlong, Wuhua, Xishan
Xian	Baqiao, Beilin, Changan, Lianhu, Weiyang, Xincheng, Yantan

Lanzhou	Anning, Chengguan, Qilihe, Xigu
Xining	Chengbei, Chengdong, Chengxi, Chengzhong
Yinchuan	Xixia, Jinfeng, Xingqing
Urumqi	Midong, Shayibake, Shuimogou, Tianshan, Toutunhe, Xinshi

Table 3: Sample Volume of the Land Transaction Dataset

	Number of Cities Covered	Number of Land Parcels Sold	Total Floor Area of Land Parcels Sold (million sq.m.)
2004	22	675	70.36
2005	24	770	91.26
2006	33	1,123	178.10
2007	34	1,409	241.27
2008	35	964	168.30
2009	35	1,567	280.20
2010	35	1,752	330.55
2011	35	1,675	274.33
2012	35	1,860	293.85
2013	35	1,871	305.16
First Half, 2014	35	621	109.93
Aggregated	-	14,287	2343.31