Land Conversion Across Cities in China

Shihe Fu — Xiamen University

Xiaocong Xu — Nanjing University

Junfu Zhang* — Clark University

Abstract

The Chinese government has been using annual quotas to control the amount of farmland that can be converted for urban uses. Using an analysis sample of more than 1.5 million land-lease transactions during 2007-2016, we document facts on land conversion for urban development in China. We present evidence that land conversion quotas have been increasingly misallocated across cities in that a growing share of land conversion is occurring in less productive cities. A city-level production function is estimated for counterfactual analysis. Based on estimated parameters, we assess the economic losses from misallocation of land conversion quotas across cities in China and calculate the potential gains from reallocating land quotas to cities where urban land is more productive.

Key words: land conversion, land quota, misallocation, urbanization, China.

JEL classifications: R12, R13, R14, R52, R58.

[Accepted for publication in *Regional Science and Urban Economics*]

*Fu is a professor of economics at Xiamen University; Xu is an assistant professor at Nanjing University; and Zhang is a professor of economics at Clark University (950 Main St., Worcester, MA 01610, USA; email: juzhang@clarku.edu; phone: 508-793-7247).

We thank Adam Storeygard, co-editor of this journal, for his constructive comments and editorial guidance. This paper has benefited from comments by John Anderson, Loren Brandt, Arya Gaduh, Yatang Lin, Yu Qin, Shunfeng Song, and seminar participants at Jinan University, MIT, National University of Singapore, Shanghai Jiao Tong University, University of Nevada–Reno, University of North Dakota, Zhongnan University of Economics and Law, Chinese Economists Society's President Forum in Ningbo, 2018 AREUEA International Conference in Guangzhou, 2018 UEA Annual Conference in New York, Conference on Urban Development and Economics in the Developing World in Shanghai, and Cities and Development Mini-Conference at Harvard. We thank Jan Brueckner for stimulating discussions on this topic. We gratefully acknowledge the financial support of a fellowship from the China Program of the Lincoln Institute of Land Policy. Fu acknowledges support from the Fundamental Research Funds for the Central Universities at Xiamen University (Grant # 20720181046) and the National Natural Science Foundation of China (Basic Scientific Center Project #71988101). All errors are our own.

1 Introduction

During the rapid urbanization in China, a large amount of farmland at the urban edge is converted for urban use. This is taking place in a unique institutional context: The central government specifies the total amount of land to be converted for urban use for a long term and for each year; this quota is divided among different provinces, which in turn is allocated to lower level governments. Under this quota system, city governments acquire land from farmers at low costs and convert it for urban use. While some of this land is allocated to building infrastructure and public facilities, the rest is leased to developers and businesses for very long terms (40, 50, or 70 years depending on the use type). Over the years, local governments have increasingly relied on land lease revenue to finance public spending.

We assembled a large data set on land conversion from a government website. Our data contain all land parcels that have been converted for urban use in China during 2007-2016. We use these data to describe different aspects of land conversion and land finance in China. We find that among prefectural level cities, land revenue amounts to more than one-third of local governments' total revenue. We show that high-land-productivity regions or cities have a declining share of land converted for urban uses, suggesting that urban land has been increasingly misallocated across cities in China. We present further evidence consistent with misallocation of urban land: First, newly converted urban land has a higher market value than agricultural land and this difference varies substantially across cities. Second, economic gains are substantial if land conversion quotas are reallocated from low- to high-land-productivity cities.

Our study contributes to several strands of literature. First, this paper is related to the literature on land-use regulation. Urban land-use regulations are ubiquitous in most countries. They take various forms—such as zoning laws, density restrictions, setback requirements, growth boundaries, etc.—and are extensively studied by urban economists.[1] Yet systematically controlling land conversion with quotas like what the Chinese government does is very unique. The massive scale of this policy is unprecedented. Its potential impact on the urban system, regional balance, and the efficiency of the whole economy is not well understood. Our paper is among the first to use micro data to document and analyze this type of land use regulation in China.[2]

Second, our paper is also related to the small literature on city size distribution in China. Au and Henderson (2006a) find that Chinese cities are typically smaller than the optimal sizes, presumably due to restrictions on internal migration of population. Chen et al. (2017b)

[1]See Gyourko and Molly (2015) for a comprehensive review of the literature on land use regulations.

[2]For studies of other aspects of this policy, see Brueckner et al. (2017), Cai et al. (2017), Chau et al. (2016), and Wang et al. (2020). Lu (2016) provides many insightful observations on this policy.

show that political favoritism affects capital prices faced by Chinese cities and in turn leads to growth differentials across cities. Others (Anderson and Ge 2005, Chen et al. 2013, Fang et al. 2017) examine the Chinese urban system through the lens of power law distributions and find that city size distribution in China is influenced by government policies. While these existing studies have identified the Hukou system, economic reforms, and urban development policies as the main factors that determine the structural characteristics of Chinese cities, our study considers a more recent urban land quota policy that plays a key role in shaping the Chinese urban system.

Third, our findings have implications for understanding the recent dynamics of urban housing markets and economic performance of cities in China. Since larger cities in coastal areas have received less land quotas over time, housing supply has lagged behind the rapidly rising demand, constantly pushing housing prices to new highs in these cities (Fang et al. 2015, Glaeser et al. 2017, Wu et al. 2016). This has a series of side effects on the Chinese economy including for example reduced firm innovation and decreased female labor force participation (Lu 2016, Han and Lu 2017, Fu et al. 2016).

And finally, this paper is related to the growing literature on resource misallocation. There has been a large number of studies on misallocation of resources along various dimensions, some of which investigate the role of land misallocation.[3] Duranton et al. (2015) extend the Olley and Pakes (1996) approach and use the covariance between land share and total factor productivity to measure land misallocation among establishments in India. They find that land misallocation plays an important role in explaining the difference of output per worker. Using a spatial equilibrium model, Fei (2020) seeks to quantify welfare losses from land market distortions in China. She shows that land prices effectively prevent productive firms from locating in large cities, resulting in substantial efficiency losses due to unrealized benefits of agglomeration and spillovers. Other papers study the effect of land misallocation on agricultural productivity (Adamopoulos and Restuccia 2014, 2015; Adamopoulos et al. 2017; Chen et al. 2017a; Restuccia and Santaeulàlia-Llopis 2017). All of these studies are conducted at the firm or farm level; in contrast, our analysis is at the city level due to the unique institutional setting in China.[4]

The rest of the paper is organized as follows. Section 2 briefly describes the institutional context. Section 3 introduces data sources. Section 4 reports several descriptive statistics on land converted to urban uses. Section 5 explores misallocation of urban land across cities. Section 6 summarizes the results with concluding remarks.

[3] See Restuccia and Rogerson (2013, 2017) and Hopenhayn (2014) for comprehensive reviews of this literature.

[4] There are a few studies of misallocation at the city level including Albouy (2009), Hsieh and Moretti (2019), Chen et al. (2017b), and Yang et al. (2017), although none of these focuses on urban land use.

2 Institutional Context

China is experiencing rapid urbanization. In 1982, only 20.9 percent of the Chinese population lived in urban areas; by 2018, this urbanization rate had climbed up to 59.6 percent. At the same time, urban areas expanded at an even faster pace, not only to accommodate the increased urban population, but also to satisfy the rising demand for space by the increasingly richer urban residents. As a result, China's urban area rose from 7,438 square kilometers in 1982 to 43,603 square kilometers in 2011 (Brueckner et al. 2017).[5]

In China, the state by law owns all of the urban land; outside urban areas, rural economic collectives own the agricultural land. Thus the expansion of an urban area involves the urban government acquiring rural land from the local economic collectives and then converting it for urban uses either by allocating the land to urban users or transferring the land use rights to developers through leasehold sales. Government regulations require proper compensation for farmers when their land is taken over for urban development. However, in reality, because the urban government has the administrative authority over the surrounding economic collectives, the compensation for farmers is often far below the market value of urban land.[6] Therefore, city governments often find it lucrative to acquire land at the urban edge and convert it for urban uses.

Two more institutional factors have provided further incentive for city governments to engage in land conversion. First, in 1994, China implemented a tax sharing system that would divide tax revenues between the central and local governments. This reform favored the central government and increased fiscal stress on local governments. In the following years, local governments throughout China had to look for non-tax revenue sources to help finance their expenditures. Before long, they all realized that selling land leases to developers can generate a substantial amount of revenue. Since then, "land finance"—using extra-budgetary land revenues to fund government spending—has become a prominent feature of local public finance in China (Cao et al. 2008). Second, China has a centralized government personnel system in which local leaders are not democratically elected but are promoted by their superiors based on their performance. There is ample evidence that during the

[5]Throughout the paper, we use the term "urban area" to refer to "urban built up area," which includes (contiguous or noncontiguous) developed land areas in a city or town on which buildings and/or infrastructure are present. It is worth noting that in planning and statistical practice in China, a piece of land is considered part of an "urban built up area" once it is acquired by a city government and converted for urban use, although literally building it up may take some time (see, e.g., Tan et al. 2003).

[6]The compensation for farmers is based on the agricultural output of the farmland instead of the opportunity cost or "best use" value of farmland. Land price at the city edge can be 500 times higher than the compensation fees paid to farmers (Wang et al. 2020). Although not our focus here, this low compensation to farmers may also cause welfare loss due to over-conversion of farmland (Tan et al. 2011, Ghatak and Mookherjee 2014).

economic reform era, local economic growth is the most important factor that determines the probability of a local leader being promoted within the Communist Party's cadre system (Li and Zhou 2005). Consequently, local leaders such as city party secretaries and mayors all have strong incentive to develop their local economies. They know that converting and developing land is an important driver of local economic growth: Construction itself contributes to local GDP directly, and better housing and infrastructure attract skilled workers and businesses that lead to long term growth. For these reasons, local governments have all been actively involved in acquiring and converting agricultural land for urban uses.

Land conversion was occurring at such a large scale and such a fast pace that it alerted the central government of a potential threat to the country's food security. To balance between the two goals of achieving economic growth and preserving cultivated land, the Chinese government implemented a top-down urban land quota system. The central government makes the nation's long-term plan to specify the total amount of land that can be used for urban development over a period of time, and then allocates this quota to provincial level governments (provinces, direct-controlled municipalities, and autonomous regions).[7] The provincial-level government then allocates its land quota to the prefectural cities under its jurisdiction, presumably based on a set of factors similar to those used by the central government. Finally, the prefectural city government decides on the scale and location of land conversion and development within the constraints of the land quota it received.

Although the central government has emphasized that an approved land-use plan must be treated as a law, in reality it is not rigidly enforced. When the allocated quota becomes binding, local leaders may petition to the upper level government for some extra quota (Xie 2015, Wang et al. 2020). Nonetheless, such a maneuver costs political capital and can only succeed to a limited extent. Thus the quota system imposes a rather stringent constraint on many local governments, particularly those in the more developed coastal regions. Its side effects have recently drawn the attention of many scholars (e.g., Lu 2016). Since the information on land quota and development at the sub-provincial level is not publicly available, there has never been systematic examination of what is happening at the city level nationwide. We seek to fill this gap in this paper.

[7]See, for example, The Outline of the National Comprehensive Planning on Land Use (2006 - 2020), released in October 2008 (available at: http://www.gov.cn/zxft/ft149/content_1144625.htm). It set the goal of preserving 1.8 billion *mu* of cultivated land nationally in 2010, and allowed the country's total area of developed land to increase from 31.92 to 33.74 million hectares during 2005-2010. This quota of newly developed land is distributed among provincial level governments based on development level, growth trend, resource and environmental conditions, etc. Although the exact quota-allocation formula is not released, differential treatment is evident. For example, the allocated quota would only allow the coastal province Shandong to expand its urban areas by 4.15 percent, but would allow the western Ningxia to expand by 10.44 percent.

3 Data

3.1 China Land Transaction Data

We assembled the China Land Transaction Data by crawling the "China Land Market" website (www.landchina.com), an information portal created and maintained by the Ministry of Land and Resources of China. One of this website's functions is to announce every land-transaction deal in China. As long as a local government handled a parcel of a land, whether it is redevelopment of urban land or conversion of rural land to urban use, it gets posted at this website. For each land transaction, the announcement typically contains information on transaction ID, land parcel address, current use, planned use type, transaction method (through negotiation, English auction, two-stage auction, sealed-bid auction, etc.), land area, price, etc.

In March 2017, we recorded all land transactions from this website with a transaction date prior to January 1, 2017, ending up with a total of 1,941,657 observations. This full data set should have captured the whole universe of land transactions in every year starting from 2007, when the central government began to systematically collect and publicize information on land transactions. After deleting duplicates, years before 2007 (with incomplete coverage), observations with key missing variables, and unreasonable outliers, we constructed an analysis sample of 1,542,279 observations for the period 2007-2016 (see Appendix A for details).[8] Since this is an online data source, we assess its reliability by comparing statistics calculated from these data to those from the China Land and Resources Statistical Yearbooks of different years (see Appendix B).

3.2 China City Statistical Yearbook data

To obtain other prefecture-level information, we use the China City Statistical Yearbooks from 1995-2015 which contain many city characteristics. Specifically, we collect annual data on city level GDP, per capita GDP, employment, fixed assets investment, urban area, budget revenue, and population from this Yearbook. Each yearbook published information from the previous year, thus these variables are used for 1994-2014.

[8]Qin et al. (2016) used the same data source to study the changing distribution of land prices in urban China during 2007-2012; they also dropped all the observations before 2007.

4 Facts on Land Conversion in China

Despite the importance of the land conversion policies in shaping the urban system and the concerns over local governments' reliance on land revenue, little is known about the scale of these issues at the city level. We thus start by documenting some stylized facts on land conversion in China. Our China Land Transaction Data contains information on all land parcels for which the local governments granted the leasehold rights to land users, including both the parcels of land newly converted from agricultural to urban uses in the current year and those already in urban uses previously. During our analysis period 2007-2016, newly converted land generally constitutes more than 80 percent of total land area in our sample, except during the global recession period (2007-2008) when this share falls below 70 percent. Newly converted land generates between 51 and 71 percent of total land revenue over different years; this share of land revenue is smaller than the share of land area in every year because newly converted land is at the urban edge and tends to have a lower market value. In this section, we primarily focus on newly converted land except in the last subsection where we will also consider redeveloped urban land parcels in order to assess the importance of total land revenue in local public finance.

4.1 Land conversion in each year, by use type

Using the land transaction data, we classify land parcels into five different categories based on use type: industrial land, residential land, commercial land, infrastructure land, and other land (including, for example, land used for schools, hospitals, government agencies, and parks). Figure 1 presents total area of and total revenue from newly converted land for each use type during 2007-2016. Notice that residential and commercial land constitute a relatively small share in the total land area converted, yet they generate the bulk of the land revenue in each year. In contrast, industrial and infrastructure land, although constitute the bulk of the land area converted, generate a much smaller share of land revenue for local governments. It is understandable that local governments tend to offer free land for infrastructure construction; after all it is a kind of public good and local governments are not supposed to profit from it. It is rather interesting to see that industrial land is also quite cheap, suggesting that local governments subsidize factories and compete for industrial investors by providing relatively cheaper land.

One might ask why local governments do not allocate more land for residential and commercial uses, given that residential and commercial land are so much more expensive than industrial land. A possible explanation is that allocation across use types is tightly controlled by upper level governments and this leads to misallocation. More plausibly, this

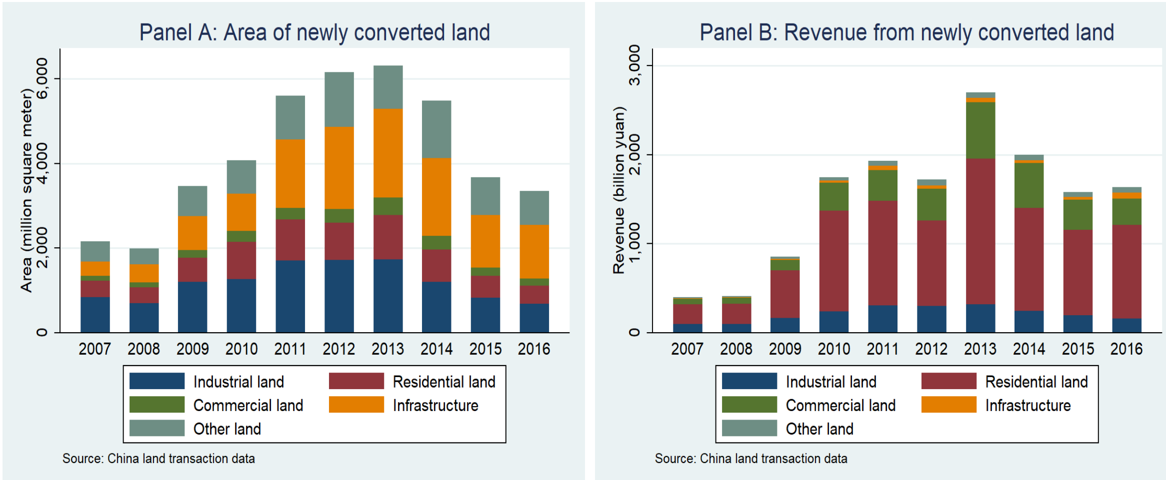


Figure 1: Newly converted land by use type

is a rational decision by local government officials. The higher price of residential land can be explained partly by its longer lease terms (70 years, as opposed to 50 years for industrial land and 40 years for commercial land). Besides, there are no residential property taxes in China, thus the price of residential land should incorporate all of its use value over 70 years. In contrast, industrial land can be used to lure plants and entrepreneurs and create new jobs. New firms and jobs generate future taxes that can justify the lower price of industrial land. Thus it makes economic sense to have differential land prices across use types. And finally, there might be inter-governmental competitions for FDI and industrial enterprises (Zhang 2011), which leads to a "race to the bottom" and thus relatively low prices for industrial land.

4.2 Land conversion in each year, by transaction method

By transaction method, we categorize land parcels in the land transactions data into six groups: by negotiation, English auction, sealed-bid auction, two-stage auction, allocation, and other methods.[9] We sum the area of and revenue from newly converted land over all parcels by transaction method in each year, which are presented in Figure 2. Two patterns are worth noting. First, although the two-stage auction only accounts for 28-47 percent of

[9]Negotiation (*xieyi* in Chinese), English auction (*paimai* in Chinese), and sealed-bid auction (*zhaobiao* in Chinese) are standard and straightforward transaction methods. Two-stage auction (*guapai* in Chinese) goes like this: The local government first posts the information about the land parcel for which the leasehold is to be transferred; potential buyers may submit their bids over a specified period of time, which is the first stage; if more than one bidder participated in the first stage, they are allowed to revise their bids in a standard English style auction at the end of the specified period, which is the second stage. Allocation refers to the free transfer of urban land to users such as schools and hospitals.

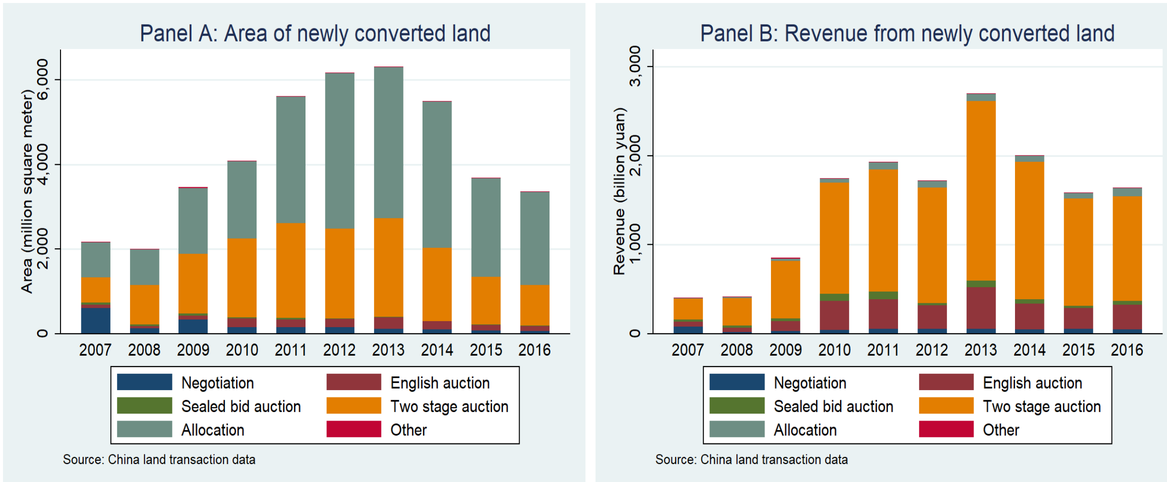


Figure 2: Newly converted land by transaction method

land area converted, it generates 59-77 percent of land revenue over different years. Many cities in our sample primarily use two stage auctions to allocate residential and commercial land. Cai et al. (2013) suggest that two-stage auctions are subject to manipulation, and show that these auctions tend to be noncompetitive and end up with lower prices. Second, a very large amount of converted land is transacted by the non-market "allocation" method, mainly for public uses, which hardly generates any land revenue.

4.3 Land conversion in coastal and inland regions

Following common practice, we define Liaoning, Beijing, Tianjin, Hebei, Shandong, Shanghai, Jiangsu, Zhejiang, Fujian, Guangdong, and Hainan as coastal regions, and the rest of mainland China as inland regions. Figure 3 graphs total area of and total revenue from newly converted land for coastal and inland regions in each year. The results show that more land area is converted for urban uses in inland regions even though less land revenue is generated in these regions. This is suggestive evidence of land misallocation since more land quotas appear to be allocated to cities in less productive regions. We will focus more intensively on this issue in section 5.

Note that coastal regions' share in total area of newly converted land was higher during 2007-2009 than later years. Although there is no reliable micro data prior to our analysis period, one can use the officially published provincial level data to calculate this share in earlier years. Indeed, one only needs to look at the early 2000s to see that coastal regions' share in land area converted used to be much higher than inland regions. As observed by scholars (e.g., Lu 2016, Han and Lu 2017), there was a dramatic change around 2003 when

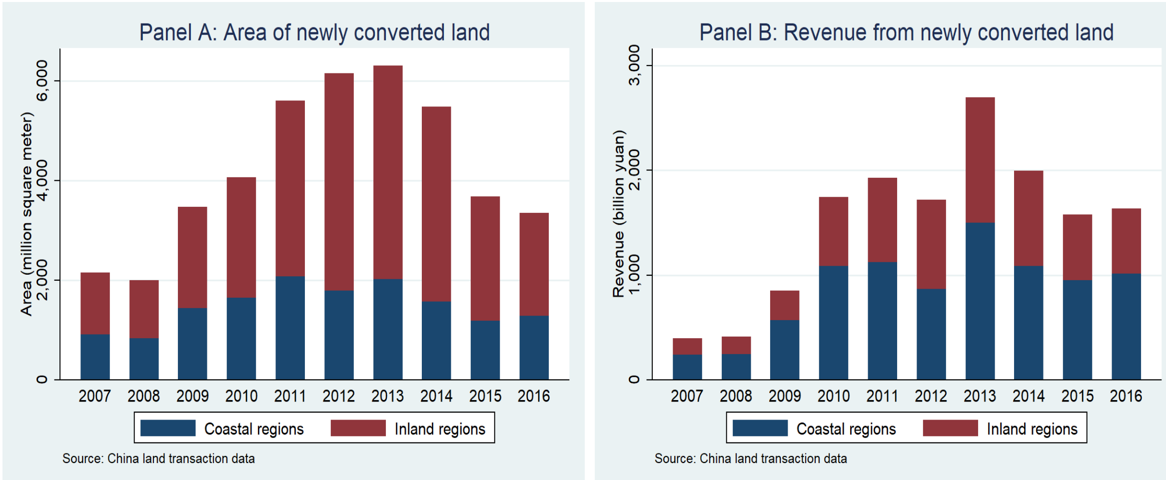


Figure 3: Newly converted land by regions

the central government started to allow an increasingly higher share of land converted in inland regions.

4.4 Land revenue and its importance in local public finance

The China City Statistical Yearbook reports each prefecture's "budget revenue" (mainly from taxes and fees) in each year, for the central city (including city districts only) and the whole prefecture (including city districts, as well as rural counties and county-level cities surrounding the central city). From the China Land Transaction Data, we can calculate total revenue from land leases for each central city or prefecture in each year. While we have been focusing on the revenue generated from newly converted land, here we also calculate the revenue from leases on redeveloped urban land; we add them together and refer to the sum as "land revenue." To examine the importance of land revenue in local public finance, we present the share of land revenue in total revenue (land revenue plus budget revenue) at both the central city and whole prefecture levels.

Table 1 reports the summary statistics for the share of land revenue. The sample mean is calculated in two ways: (1) as a weighted average of the share in each city-year, using the city's nominal total revenue in the year as weights; and (2) as a weighted average of the share in each city (over 2007-2014), using the city's real total revenue (over 2007-2014) as weights.[10] It turns out that the sample means calculated in these two ways are almost identical. For central cities (columns (1)-(2)), land revenue on average amounts to 35 percent

[10] For method (2), we use CPI (consumer price index, downloaded from http://www.stats.gov.cn/tjsj/ndsj/) to deflate land and budget revenues. Cities with at least one year of missing data are excluded from this calculation.

Table 1: Summary statistics for share of land revenue in total revenue

Statistics	Central Cities		Whole Prefectures	
	Annual shares in total revenue	2007-2014 shares in total revenue	Annual shares in total revenue	2007-2014 shares in total revenue
	(1)	(2)	(3)	(4)
Mean	0.3516	0.3515	0.3407	0.3401
Std. Dev.	0.184	0.1397	0.139	0.0993
Median	0.3614	0.3897	0.3245	0.3578
Minimum	0.0006	0.0861	0.0008	0.0864
Maximum	0.901	0.7677	0.944	0.5872
Observations	2,237	252	2,237	252

Notes: The mean is a weighted average across cities, using each city's total revenue as weights.

of the total revenue. At the whole prefecture level (columns (3)-(4)), the average share is 34 percent, implying a slightly lower share at jurisdictions outside of the central cities.

Figure 4 presents the distribution of land revenue's share in total revenue, for central cities and for whole prefectures, calculated using method (2). While the average share is lower than 0.4, the distribution is highly dispersed. Many central cities have a land revenue share higher than 0.5. That is, these jurisdictions derived more revenue from land leases than from taxes and fees. They tend to be smaller cities, which have a relatively small influence on the sample mean.

Figure 5 plots the average of land revenue share by year, for central cities and whole prefectures. We show the trend for cities/prefectures in the whole sample, in the inland provinces, and in the coastal provinces. For central cities, the share is noticeably higher in inland provinces than coastal provinces every year during 2007-2014. Also, this share fluctuates substantially from year to year, at both the central city and the whole prefecture levels. For example, the share of land revenue in total revenue for central cities were only 0.25 in 2008, a year when economic growth slowed down and urban development slacked off amid a global recession; two years later, the ratio jumped to 0.43.

5 Misallocation of Land Quotas across Cities

In this section, we first present some indicative evidence of land misallocation across cities in China. We then present a simple graphical model to provide a framework for detecting

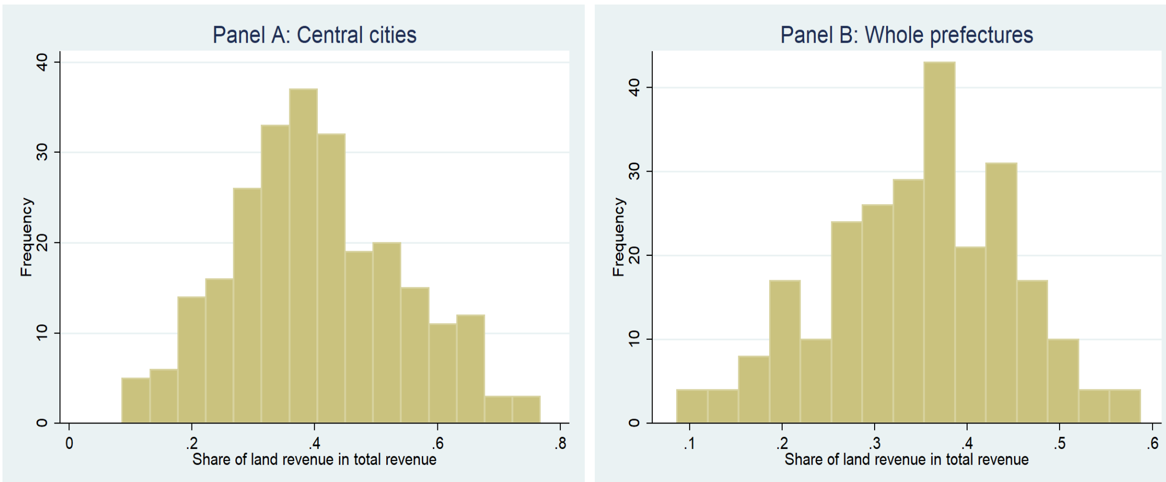


Figure 4: Distribution of land revenue's share in total revenue

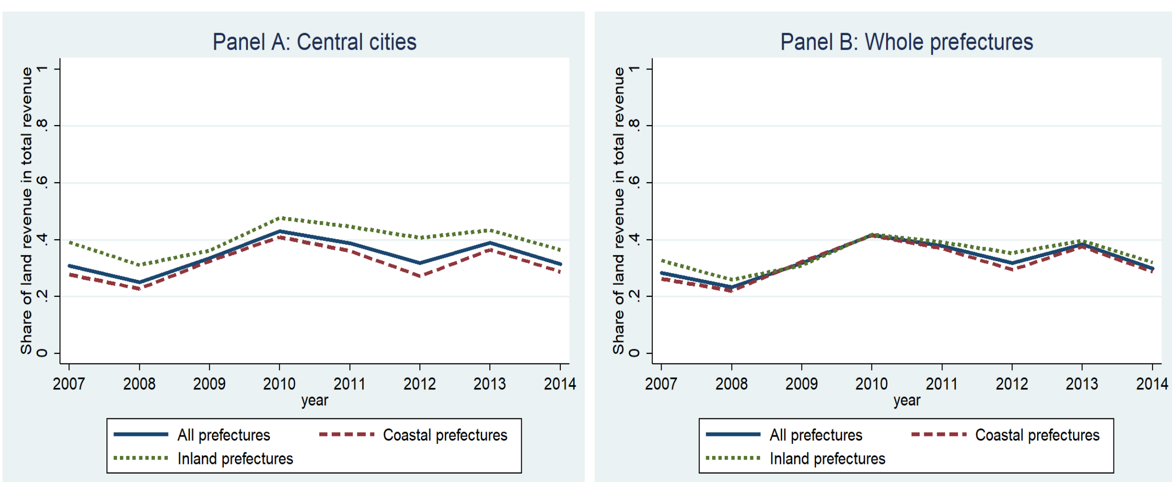


Figure 5: Share of land revenue in total revenue 2007-2014

land misallocation across cities. Using the model as a guide, we present evidence that at the urban edge, the price gap between urban and agricultural land and the marginal productivity gain from converting land for urban uses both vary a lot across cities, implying misallocation of land across cities. We perform some counterfactual analysis to show that the economic gain is substantial if the Chinese government can reallocate some land conversion quotas from low- to high-land-productivity cities. From this point on, our analysis sample will focus exclusively on newly converted urban land and will necessarily drop land parcels with missing price or area data.

5.1 Indicative evidence on land misallocation across cities

5.1.1 Quota restrictions on regions/cities with high land productivity

Some economists (e.g., Lu 2016) have pointed out that since 2003 a smaller and smaller share of land conversion quotas has been allocated to coastal provinces. They argue that this is a misallocation because urban land is much more valuable in coastal than inland provinces. Using the China Land Transaction Data, we calculate the share of newly converted urban land for coastal provinces during 2007-2016 (Panel A in Figure 6). Indeed this share was declining during 2007-2014 and only started to increase in the last two years. According to Lu (2016) and Han and Lu (2017), who calculated the share using officially published provincial level data, the declining trend started earlier in 2003. We replicated their calculation using the provincial-level data from the China Land and Resources Statistical Yearbook and confirmed this claim: Indeed, the share of land converted by coastal provinces was rising during 2000-2003; it then decreased steadily for a whole decade until 2014.

In Panel B of Figure 6, instead of using the coastal-inland classification, we directly divide cities into high- and low-land-productivity cities, and then examine the share of newly converted land in high productivity cities. Specifically, we first calculate the average productivity of land (APL) for each city by dividing the city's real GDP by its total land area and averaging this measure over 2007-2014 (the yearbook data only available up to 2014 when this project was started). Next, we classify all cities into low APL cities and high APL cities by splitting the sample roughly in between (in terms of newly converted land): High APL cities include all cities with an APL higher than the median areal unit of converted land in terms of city APL, and low APL cities include all the rest. This approach classified 199 cities into the group of low APL cities and 80 cities into the group of high APL cities. Using the China Land Transaction Data, we calculate the share of newly converted land (in central cities and whole prefectures) allocated to the high APL cities. It appears that these two shares both declined substantially from 2007 to 2012; high APL cities (relative

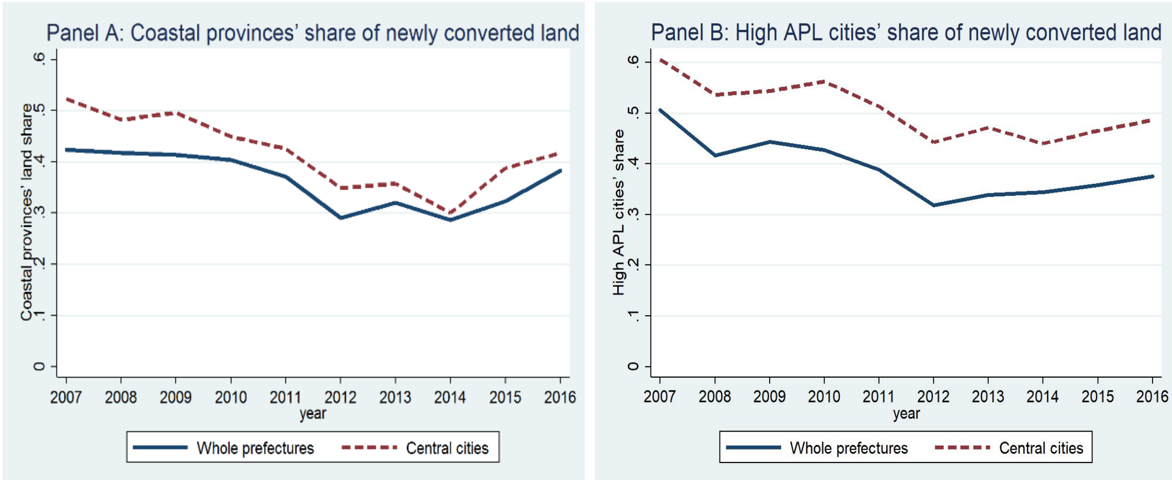


Figure 6: Share of newly converted urban land for coastal provinces and high productivity cities

to low APL cities) converted less and less land for urban uses over time, suggesting that land misallocation got worse during this period. Starting in 2012, the trend leveled off and reversed slightly.[11]

5.1.2 Indicators of land misallocation across cities

Following standard practice in the literature on misallocation, we next examine a few commonly used misallocation indicators. Using the China City Statistical Yearbook data, we calculate the average productivity of land for all prefectural-level cities in each year during 1994-2014. We then look at the dispersion of land productivity. Higher dispersion is indicative of land misallocation across cities.[12]

Panel A of Figure 7 plots the difference between the 10th and 90th percentile of land productivity among cities over time. Panel B similarly plots the difference between the 25th and 75th percentile of land productivity among cities over time. Panel C plots the standard deviation of city-level land productivity over time. In each panel, we draw a vertical line to indicate the year 2003, when the Chinese government started to allocate more land quotas to cities in inland provinces. Each of the three dispersion measures has an increasing trend,

[11]We also tried regressing land converted to urban use on land productivity at the city-year level and find a negative (though not statistically significant) coefficient.

[12]We draw intuition from Hsieh and Klenow (2009), who show that the efficiency of factor allocation across firms is related to the variance of total factor productivity among firms. In our case here, one might argue that the dispersion of marginal (rather than average) productivity of land is a more relevant indicator of misallocation. However, if city production function is of the Cobb-Douglas form, as will be assumed below, then marginal productivity is proportional to average productivity and their dispersions should follow the same trend.

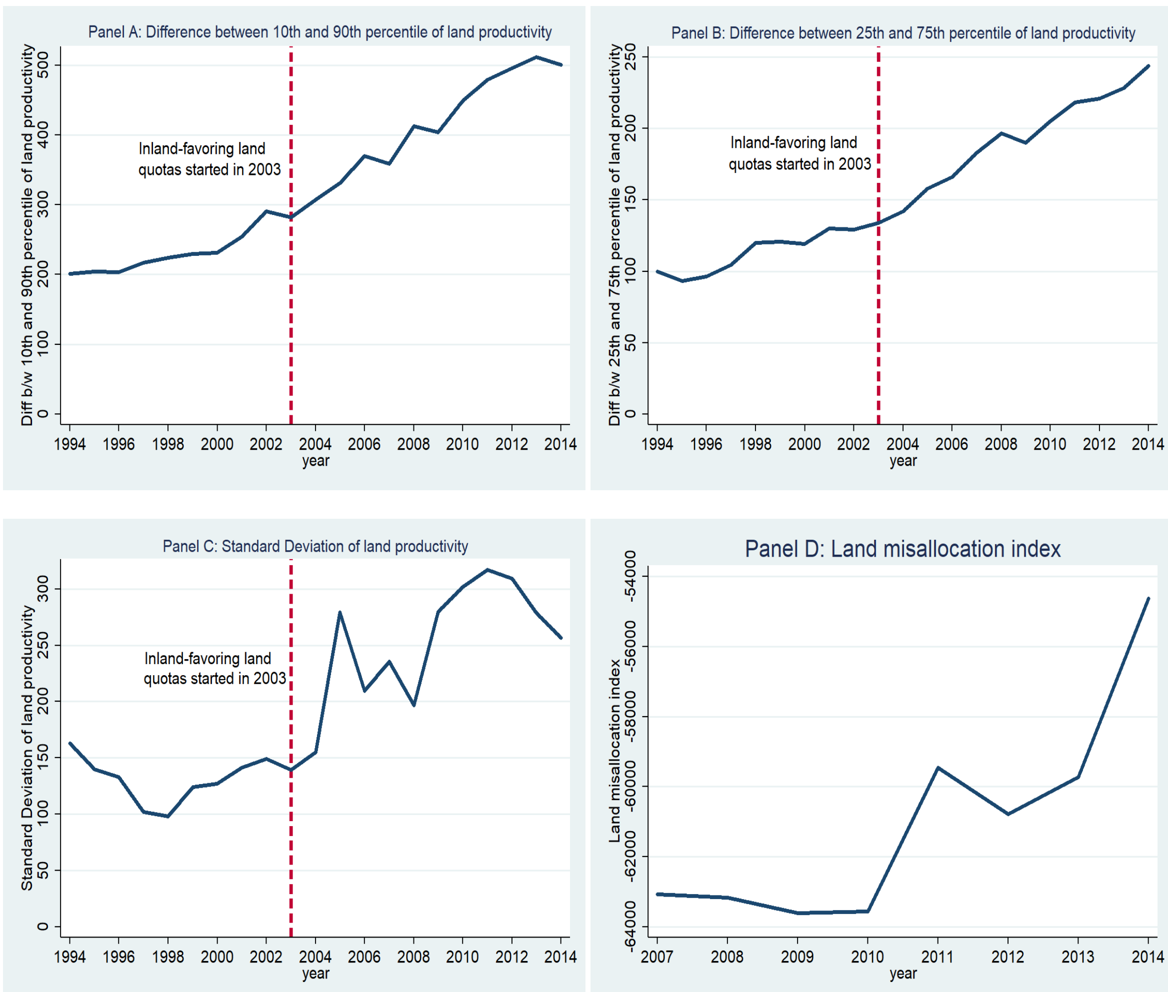


Figure 7: Indicators of land misallocation over time

suggesting that land misallocation had become more serious over time. The trend of the standard deviation (Panel C) clearly shows 2003 as a break point.

Instead of the *ad hoc* productivity dispersion measures, we next look at a regression-based measure of misallocation. Following Duranton et al. (2015), we define misallocation of land across cities in year t as follows:[13]

$$M_t = -n_t * cov_t\left(s_{it}, A_{it}\right), \tag{1}$$

where n_t is the number of cities in year t; s_{it} is city i's share of all urban land in year t; A_{it} is the total factor productivity (TFP) in city i in year t. The land misallocation index is a rescaled covariance between land share and TFP, the latter of which is calculated as the residual of a city-level Cobb-Douglas production function in land, labor, and capital (see Appendix C for detail). This measure is very intuitive: If the more productive cities have increasingly larger land shares, then there's little misallocation of urban land; otherwise, if the more productive cities have decreasing land shares, then there is misallocation of land across cities. The misallocation index, plotted in Panel D of Figure 7, also suggests that misallocation has become more serious over time, although the trend is less clear with only eight years of data.

5.2 A simple model

To illustrate the economic intuition, we present a simple model of land misallocation across cities.

Consider an urban system with two monocentric cities, 1 and 2 (see Figure 8). In each city, production all takes place at the central business district (CBD). Workers live around the CBD, trading off between higher commuting costs and lower land rents. $r_i^u(d)$, $i = 1, 2$, is the bid rent curve in the urban sector, decreasing with d, the distance from the CBD. r_i^a, $i = 1, 2$, is the bid rent curve in the agricultural sector, assumed to be constant in either area.

Without government intervention, in each city land will be used by the sector that can afford a higher bid rent. Thus city i ends at d_i, where the urban bid rent equals the agricultural bid rent. Total urban area in this economy is $\pi(d_1^2 + d_2^2)$. Land allocation is efficient in that there is no way to gain from rearranging land use within or across cities.

Suppose that for some reason, the government decides to control urban land supply.

[13] This measure of misallocation is equal to the difference between the simple and share-weighted average productivity. Olley and Pakes (1996) first used this measure to study firm productivity; Duranton et al. (2015) modified it to measure misallocation along different dimensions.

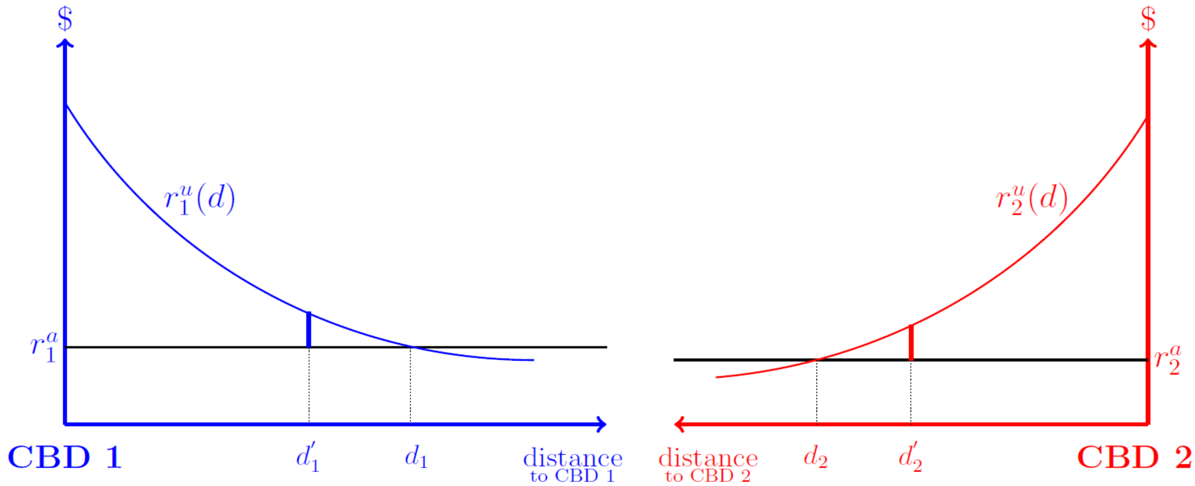


Figure 8: A simple graphic model

Say, it only allows $\pi(d_1^2 + d_2^2) - \theta$ units of land to be used by the urban sector, where θ is a predetermined constant. Suppose that the government uses land quotas to stop urban development at d_i' so that $\pi[(d_1^2 + d_2^2) - (d_1'^2 + d_2'^2)] = \theta$. In this case, land allocation across cities is efficient under the following condition:

$$r_1^u(d_1') - r_1^a = r_2^u(d_2') - r_2^a. \tag{2}$$

If $r_i^u(d_i') - r_i^a > r_j^u(d_j') - r_j^a$, then the government can reallocate some land quota from city j to i to improve efficiency.

Let R be the discount rate. At the edge of city i, the price of land in the agricultural sector is $p_i^a = \sum_{t=1}^{\infty} \frac{r_i^a}{(1+R)^t} = r_i^a/R$; similarly, the price of land in the urban sector is $p_i^u = r_i^u(d_i')/R$. Thus, $p_i^u - p_i^a = [r_i^u(d_i') - r_i^a]/R$. That is, if we have land price or rent in both the urban and agricultural sectors at the urban edge, we can test whether land allocation across cities is efficient. In this study, we have land price when it is converted for urban use, and we can estimate land rent in the agricultural sector by its productivity. Thus we can test efficiency by calculating the following urban-rural land price gap for each city i:

$$\kappa_i \equiv p_i^u - r_i^a/R. \tag{3}$$

If κ_i is significantly different across cities, then $p_i^u - p_i^a = [r_i^u(d_i') - r_i^a]/R$ is significantly different across cities, implying inefficient allocation of land quotas across cities.

Alternatively, unlike in equation (3) where we calculate the gap between sales prices of urban and rural land, we can also examine the gap between rental prices of urban and rural land at the urban edge. With the estimated rural land rent (r_i^a), we can proxy urban land rent ($r_i^u(d_i^{'})$) by the estimated marginal productivity of urban land in city i (MPL_i). If $MPL_i - r_i^a > MPL_j - r_j^a$, then there is misallocation of land quotas across cities and the government can improve efficiency by reallocating some land quota from city j to i. For every unit of land quota reallocated, the gain is $(MPL_i - MPL_j) - (r_i^a - r_j^a)$.

5.3 Differential urban-rural land price gaps across cities

Guided by the simple model, we detect misallocation of land quotas by testing equality of urban-rural land price gaps across cities.

5.3.1 Construction of variables

We use the China Land Transaction Data to calculate the price of urban land (p_{it}^u) for each city in each year as follows:

$$p_{it}^u = \frac{\sum_{k \in B_{it}} Sold\_Price_{itk}}{\sum_{k \in B_{it}} Area_{itk}}, \tag{4}$$

where the subscripts are city i, year t, and land parcel k. B_{it} is the set of newly converted land parcels in city i in year t. That is, for each city-year, we divide the total land revenue by the total land area for parcels newly converted for urban uses.[14] We use the consumer price index, collected from the China Statistical Yearbook, to deflate land price.

Since there is no rental or sales market for rural land, we use the China City Statistical Yearbook data to calculate rural land rent outside the city as follows:

$$r_{it}^a = \frac{First\_Sector\_GDP_{it}}{Cultivated\_Land\_Area_{it}}, \tag{5}$$

where $First\_Sector\_GDP_{it}$ is the GDP in the agricultural sector in city i in year t. $Cultivated\_Land\_Area_{it}$ is the total cultivated land area in city i in year t. From the China City Statistical Yearbook, we have the cultivated land area for each city in 2007. To obtain cultivated land area in each year during 2008-2014, we subtract the area of converted

[14]Our calculation is based on all land parcels newly converted for urban uses, including those whose leasehold rights are transferred to users at very low prices. One might argue that this measure underestimates the market value of urban land, because local governments have incentives to charge low prices for certain land parcels (e.g., those for industrial uses) in exchange for nonpecuniary gains or future benefits (e.g., employment opportunities or tax revenue). We want to emphasize that this potential underestimation does not affect the validity of our test as long as our measure is proportional to the true market value of urban land.

Table 2: Summary statistics for urban land price and agricultural land rent

Variable	Obs.	Mean	Std. Dev.	Min	Max
Urban land price (yuan/m^2)	2,186	333.7	359.6	3.012	6,128
Rural land rent (yuan/m^2)	2,186	5.199	5.076	0.356	111.8

land in the year from the previous year's cultivated land area.[15] We again use the consumer price index to deflate the GDP in the agricultural sector. Table 2 reports the summary statistics of urban land price and imputed rural land rent. We find that urban land price is over 60 times of average rural land rent. Note that urban land price is transaction price, which should be a sum of discounted revenue streams it can generate over many years; rural land rent, however, is calculated based on the value of one year's output rather than the price one pays to acquire the ownership of the land.

5.3.2 Estimating urban-rural land price gap

We now have urban land price (p^u_{it}) and rural land rent (r^a_{it}) for each city in several years, but the discount rate R in equation (3) is not observable. Thus we cannot directly calculate the urban-rural land price gap (κ_i). Instead, we will simultaneously estimate R and an average κ_i based on the following version of equation (3):

$$p^u_{it} = \delta r^a_{it} + \bar{\kappa}_i + \varepsilon_{it}, \tag{3'}$$

where $\delta = \frac{1}{R}$ and the city specific constant $\bar{\kappa}_i$ represents the average urban-rural land price gap for city i. Since the yearly variation of the urban-rural land price gap for city i ($\kappa_i - \bar{\kappa}_i$) is absorbed in the error term ε_{it}, we estimate equation ($3'$) by clustering standard errors at the city level. This generates an estimate $\hat{\delta} = 16.1876$, which is marginally significant with a standard error of 9.385. Figure 9 illustrates the variation in estimated coefficients for city dummies ($\hat{\bar{\kappa}}_i$, i.e., the estimated urban-rural land price gaps). To make the figure more legible, we show only the top and bottom 30 cities separated by the median city (which is Liupanshui in Guizhou province). Beijing and Shanghai are clear outliers in that urban land is much more valuable than average rural land in surrounding areas. There is substantial variation among other cities too: Whereas all of the top 30 cities have an estimated urban-

[15]We implicitly ignored factors other than urbanization that could affect a city's total cultivated land area from year to year. For example, some farmland may be converted to forestry to achieve ecological balance; some rural family homesteads may be reclaimed into farmland. Whereas these factors may decrease or increase a city's total cultivated land area, we believe that they are relatively unimportant given that the central government has continuously cited urbanization as a threat to maintaining farmland and food security.

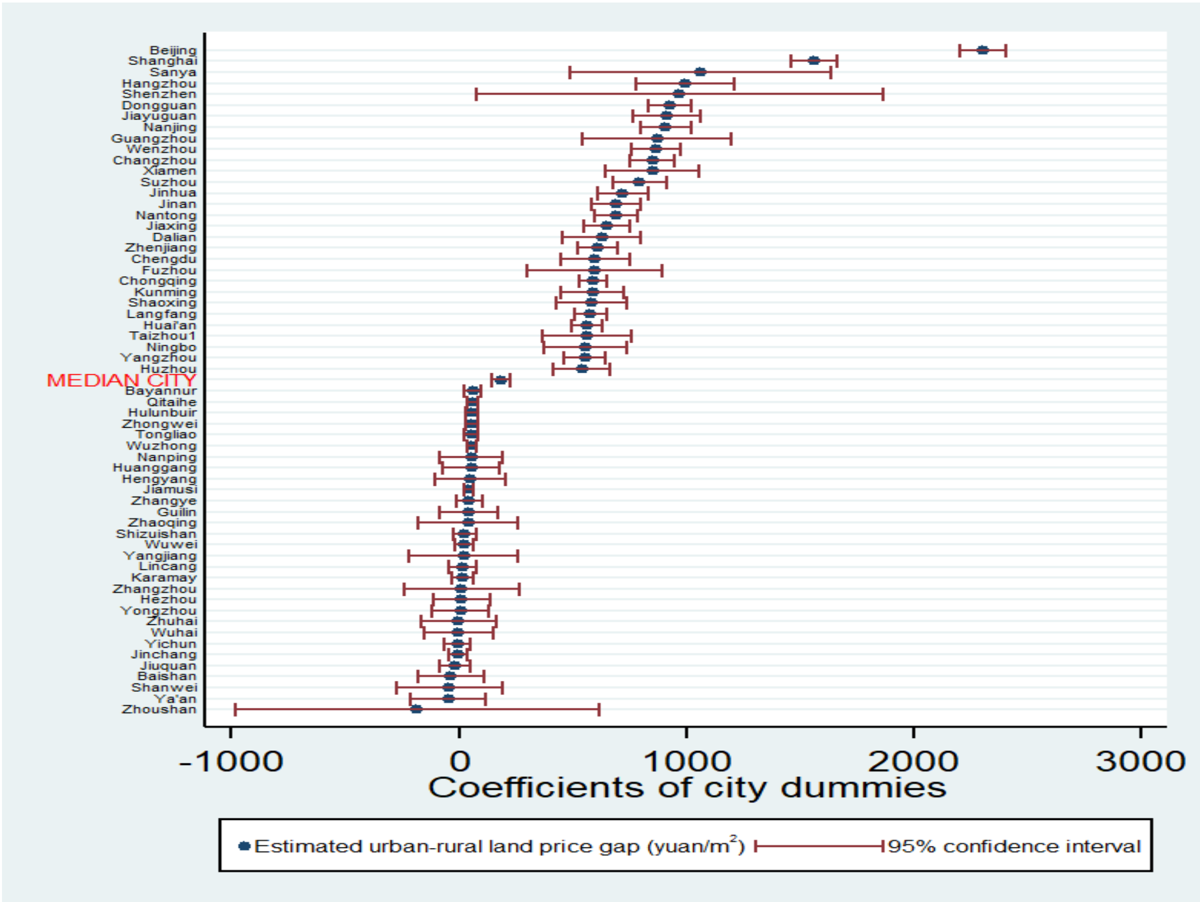


Figure 9: Top and bottom 30 cities by estimated urban-rural land price gap

rural land price gap higher than 540 $yuan/m^2$, none of the bottom 30 cities has a gap higher than 62 $yuan/m^2$.[16] That is, if land conversion quotas are reallocated from the bottom to the top cities, there will be substantial gains without affecting the total amount of land converted for urban use in the whole country. Figure 9 also shows that indeed coastal provinces tend to have larger urban-rural land price gaps. Among the top 30 cities with the largest urban-rural land price gaps, 26 are in coastal provinces; in contrast, among the bottom 30 cities, only 7 are in coastal provinces.[17]

Equation (3′) assumes a single δ and thus the same discount rate R everywhere. Given the work by Chen et al. (2017b), one might suspect that different cities have different interest rates and therefore different discount rates. In alternative specifications, we tried the following: (1) allow provincial-level and province capital cities to have a different discount

[16]For many of the bottom 30 cities, the estimated urban-rural land price gap is not significantly different zero, which would satisfy the boundary condition in a standard monocentric city model.

[17]In an alternative specification, we tried regressing the ratio of urban land price to rural land rent on city dummies. The results also show that there is substantial variation in the urban-rural land price gap across cities and cities in coastal provinces tend to have larger urban-rural land price gaps.

rate than other cities; and (2) divide cities into three size groups (small, medium, and large, each constituting one third of the sample) and allow different groups to have different discount rates. The results indeed suggest lower discount rates for larger cities and cities higher in the political hierarchy, consistent with the notion that such cities face lower interest rates. However, these alternative specifications do not change the fact that the estimated urban-rural land price gaps vary a great deal across cities.

5.3.3 Explaining urban-rural land price gap

We further explore what kind of cities tend to have a higher gap. We regress the urban-rural land price gap on city characteristics as follows:

$$p_{it}^{u} - 16.1876 * r_{it}^{a} = \alpha + \beta * X_{it} + \varepsilon_{it}, \tag{6}$$

where city characteristics X_{it} include coastal-province dummy, inland-province dummy, provincial capital dummy, population, urban area, per capita GDP, and per capita government revenue. Table 3 reports the results. To account for the fact that the left hand side of equation (6) is estimated, we report bootstrapped standard errors in addition to "regular" standard errors calculated from the asymptotic covariance matrix. The results in column (1) show that the average urban-rural land price gap for cities in coastal provinces is more than twice as big as that in inland provinces. In other words, converting one unit of land for urban use generates a much larger value premium in coastal than inland provinces, suggesting that land quota puts a more stringent constraint on cities in coastal provinces. Results in columns (2)-(8) suggest that the urban-rural land price gap is larger for provincial capitals, cities with more population, cities with larger urban areas, and cities with higher per capita government revenues. There has been speculations that the Chinese government uses land quotas to control population growth in large cities and balance cross-region inequalities in government revenue. Our regression results are consistent with such notions.

One might wonder whether the variation of urban-rural land price gaps simply reflects measurement errors. Indeed, this issue is a common concern in the broader literature on misallocation of productive resources.[18] Following Hsieh and Klenow (2009), we perform two informal checks. First, we re-estimate equations (3') and (6) after dropping the top and bottom 1 percent outliers of urban land prices and those of rural land rents. The idea is that compared to data points in the interior of the distribution, these outliers are more likely

[18]See, for example, the influential study of Hsieh and Klenow (2009). After showing potential gains from better resource allocation in China and India, they admit that "our potential efficiency gains could be a figment of greater measurement error in Chinese and Indian data than in the U.S. data" (Hsieh and Klenow 2009, p.1426).

Table 3: Explaining urban-rural land price gap

	Dependent variable: $p^u_{it} - 16.1876 * r^a_{it}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coastal provinces	388.9							
	(11.79)***							
	[15.34]***							
Inland provinces	172.8							
	(8.67)***							
	[6.63]***							
Province capital		357.9					81.56	97.67
		(22.36)***					(25.77)***	(27.43)***
		[45.26]***					[37.35]**	[35.54]***
Population			108.8				89.10	51.21
			(10.40)***				(11.73)***	(13.74)***
			[17.96]***				[24.58]***	[27.53]*
Urban area				164.7			39.63	34.71
				(7.72)***			(13.64)***	(13.78)**
				[16.66]***			[12.91]***	[17.01]**
Per-capita GDP					171.8		-11.07	51.58
					(9.77)***		(21.13)	(22.43)**
					[15.68]***		[23.60]	[30.08]*
Per-capita government revenue						137.7	114.4	41.16
						(6.71)***	(15.11)***	(16.94)**
						[13.64]***	[18.71]***	[23.03]**
Province fixed effect		N	N	N	N	N	N	Y
Constant		210.1	-389.9	-466.9	-1510.7	-784.3	-1197.0	-466.4
		(7.337)***	(61.44)***	(33.77)***	(100.3)***	(50.81)***	(147.9)***	(200.2)**
N	2,183	2,183	2,183	1,908	2,183	2,183	1,908	1,908
R^2	0.405	0.105	0.048	0.193	0.124	0.162	0.254	0.381

Regular standard errors (calculated from the asymptotic covariance matrix) are in parentheses and bootstrapped standard errors are in square brackets. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

to come from recording errors (e.g., mistakenly adding or dropping a digit, misplacing the decimal point, or entering a wrong unit) and one should be concerned if they are driving our results. We find that the results in Figure 9 and Table 3 are all robust to dropping outliers. Second, we correlate our constructed price of urban land and the urban-rural land price gap with the average productivity of land at the city level. The price of newly converted urban land in theory is determined by the marginal productivity of land and thus should be highly correlated with the average productivity of land in the city. Similarly, because the urban-rural land price gap is mainly driven by the price of urban land, this gap should also be related to the average productivity of land in the city. Simple correlations and univariate regressions show that both the price of urban land and the urban-rural land price gap are indeed highly correlated with the average productivity of urban land, suggesting that both measures contain real information instead of just random errors.[19] Although neither of these informal checks can rule out the possibility that measurement errors have played a role in creating the dispersion of urban-rural land price gaps, we find these results reassuring.

5.4 Potential gains from more efficient land allocation

In this subsection, we quantify the potential gains if some land conversion quotas are reallocated from low- to high-productivity cities. As suggested by the model above, we need to estimate the productivity of rural land right outside of city boundaries and newly converted urban land for each city. We will continue to use first-sector GDP per unit of land as rural land productivity. Whereas rural land right outside city boundaries may be more productive due to its proximity to a large urban market, negative externalities from the city (e.g., pollution) may also make such land less productive than rural land further away. Other factors affecting land productivity, such as soil quality and climate variability, are likely to be uncorrelated with distance from the city. Thus it seems reasonable to estimate the productivity of rural land right outside of city boundaries using the average productivity of rural land in the surrounding rural areas.

For newly converted urban land, it is more complicated. Newly converted land parcels are at the edges of existing cities and thus their productivity tends to be much lower than the average parcel of urban land. Using our land transaction data, we provide some evidence that newly converted urban land has a much lower value than redeveloped urban land, the latter of which has typically been developed for urban use many years ago and should be closer to the urban center. Specifically, we select all residential, commercial, and industrial land

[19] Even with the outliers included, the slope coefficient in a city-level univariate regression of average land productivity on urban land price has a t-statistic of 18.2; the slope coefficient in a city-level univariate regression of average land productivity on urban-rural land price gap has a t-statistic of 15.6.

parcels from our data, and regress log land price on a redevelopment dummy together with log land area, city fixed effects, and year fixed effects. We find that the price of redeveloped urban land is 61.1% higher than newly converted land.[20] Therefore, we must make some adjustment when using average urban land productivity to estimate the productivity of newly converted urban land.

We treat the productivity of newly converted urban land as the marginal productivity of land in a city. Our estimation below is based on a crucial assumption—city production function takes the Cobb-Douglas form, implying that the marginal productivity of land is proportional to its average productivity. Although this function form is commonly used in the literature for its convenient properties, we must admit upfront that it is a very strong assumption. In particular, it requires that the output elasticity of labor, capital, and land are all constant across cities and over time. This is unlikely to be true in a large, diverse, and rapidly-growing economy like China. Here we proceed by maintaining this assumption but will, at the end of this subsection, explore how the assumption may have affected our results.

5.4.1 Specification

Consider the following city-level production function:

$$Y_{it} = A_{it} N_{it}^{\alpha} K_{it}^{\beta} L_{it}^{\gamma}, \tag{7}$$

where Y_{it} is the output level; A_{it} is a productivity parameter; N_{it} is the number of workers; K_{it} is the capital stock; and L_{it} is the quantity of urban land, all indexed by city i and year t. Taking log of equation (7), we have:

$$\ln Y_{it} = \ln A_{it} + \alpha \ln N_{it} + \beta \ln K_{it} + \gamma \ln L_{it}. \tag{8}$$

For empirical implementation, we further assume that the total factor productivity can be decomposed as follows: $\ln A_{it} = C_i + \tau_t + \varepsilon_{it}$, where C_i is a city-specific time-invariant component that captures the effect of local fundamentals; τ_t is a year fixed effect that captures common macroeconomic shocks; and ε_{it} is an idiosyncratic error term. Therefore, we have

[20] Further investigation reveals that this difference is mainly driven by residential and commercial land: Redeveloped residential land is 77.7% more expensive than newly converted residential land; for commercial land, this difference is 26.7%. In contrast, redeveloped industrial land is 2.5% cheaper than newly converted industrial land, perhaps because the value of industrial land depends more on its proximity to highways and locations near highways are equally available at the urban edges.

the following empirical specification:

$$\ln Y_{it} = \alpha \ln N_{it} + \beta \ln K_{it} + \gamma \ln L_{it} + C_i + \tau_t + \varepsilon_{it}. \quad (9)$$

We need an estimate of γ, based on which we can perform some counterfactual analysis.

5.4.2 Key variables

We estimate equation (9) using city-level data during 2007-2014. For output Y, we use city GDP; for labor N, we use total employment in the city; both are from the China City Statistical Yearbook.

For capital K, we use the perpetual inventory method to estimate it since we only have fixed assets investment data from the China City Statistical Yearbook. The initial capital is calculated as

$$K_{i,2007} = \frac{FAI_{i,2007}}{g + \phi}, \quad (10)$$

where $FAI_{i,2007}$ is city i's fixed assets investment in year 2007; g is the annual growth rate of fixed assets investment during 2007-2014; ϕ is the annual depreciation rate, assumed to be 5%. We obtain the fixed assets investment during 2007-2014 from the China City Statistical Yearbook, so we can calculate its annual growth rate g. Using equation (10), we obtain $K_{i,2007}$. The capital in the following years is calculated by

$$K_{it} = (1 - \phi) K_{it-1} + FAI_{it} \quad (11)$$

For urban land area L, we combine the China Land Transaction Data with the China City Statistical Yearbook data. Specifically, from the China Land Transaction Data, we calculate the total area of land converted for urban uses in each city in each year during 2008-2014. For urban land area in 2007, we use the 2007 urban area in each city from the China City Statistical Yearbook. To obtain a city's land area in each year from 2008 to 2014, we add its total area of newly converted land in each year to its land area in the previous year. Table 4 reports the summary statistics of regression variables.

5.4.3 Estimating city-level production function

To consistently estimate equation (9), we need to confront a major econometric issue due to potential simultaneity bias. That is, observed inputs (land, labor, and capital) may be correlated with unobserved inputs or productivity shocks. For example, a young mayor may have more incentive to promote growth in her city in order to be promoted within the Communist Party. She may thus negotiate with upper level government officials for

Table 4: Summary statistics of variables for estimating city production function

Variable	Obs.	Mean	Std. Dev.	Min	Max
Log output ($\ln Y$)	2,153	15.08	1.153	12.21	19.06
Log employment ($\ln N$)	2,153	3.338	1.020	0.761	7.339
Log capital ($\ln K$)	2,153	16.07	1.234	11.81	20.14
Log land ($\ln L$)	2,153	4.339	0.816	1.946	7.366

Units of measurement: Output—10,000 yuan; employment—10,000 persons; capital—10,000 yuan; and land—square kilometers.

more land quotas and at the same time use a few other pro-growth measures, consequently introducing an upward bias in the estimated land coefficient. On the other hand, facing a negative local productivity shock, a mayor may request more land quotas with the hope that new development projects will help boost the local economy. This, in contrast, leads to a downward bias in the estimated land coefficient. The city fixed effects specification can be thought of as a partial solution to this simultaneity problem, in that the fixed effect term can absorb the effect of time invariant unobserved inputs or productivity shocks. However, it does not solve the problem if the unobserved shocks vary over time, e.g., with two consecutive mayors having different motives to develop the local economy.

To deal with this issue, following Levinsohn and Petrin (2003), we use intermediate inputs (water, gas, or both) to control for unobserved productivity heterogeneities.[21] Table 5 reports the regression results of equation (9). Column (1) is the OLS results, and columns (2)-(4) report the Levinsohn-Petrin (LP) results, using water, gas, or both water and gas to control for unobserved productivity.[22] The OLS estimate of γ, the coefficient of log land and our key parameter of interest, is 0.135. The LP estimates are slightly higher, ranging from 0.155 to 0.177. These LP estimates imply that the marginal productivity of urban land is about one sixth of the average productivity, which seems reasonable.[23] The fact that all LP

[21] Given that the Levinsohn-Petrin method was initially designed for firms, there are two differences when we apply it to cities. First, since the number of cities is relatively small, our sample size here is many orders of magnitude smaller than that of a typical data set for firms. As remarked below, some extensions of this method may not be directly used for cities because of a small sample size. Second, when applying this method to firms, there is an "endogenous exit" issue that firms dropping out of the sample are not random. This is not a concern in our context because our sample of cities is rather stable over the relatively short period of time.

[22] Electricity is a commonly used input to control for unobserved productivity in the empirical IO literature. However, there are too many missing values for the electricity variable in the Yearbook data, making it not useful in our case here.

[23] The coefficients for labor and capital, on the other hand, seem too small. We suspect that there are some unaccounted factors of production, such as infrastructure financed by central and provincial governments, which may be missing from local statistics. The contribution of such unaccounted factors to local GDP should be partially captured by the city and year fixed effects. Also, we cannot rule out the possibility that these coefficients might have suffered from collinearity and attenuation bias. Note that our analysis below does not rely on consistent estimation of these two parameters.

Table 5: Regression results for city-level production function

Dependent variable: Log city GDP ($\ln Y$), 2007-2014				
	OLS	LP-water	LP-gas	LP-water & gas
	(1)	(2)	(3)	(4)
Log Employment ($\ln N$)	0.169***	0.173***	0.187***	0.170***
	(0.0147)	(0.0182)	(0.0308)	(0.0419)
Log Capital ($\ln K$)	0.186***	0.190***	0.191***	0.189***
	(0.0133)	(0.0171)	(0.0247)	(0.0278)
Log Land ($\ln L$)	0.135***	0.155***	0.161***	0.177**
	(0.0273)	(0.0121)	(0.0792)	(0.0809)
City fixed effects	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y
N	2,153	2,119	1,901	1,900
R^2	0.992	—	—	—

Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$. "LP" indicates the use of intermediate inputs (water, gas, or both) to control for unobserved productivity shocks, following Levinsohn and Petrin (2003).

estimates are larger than the OLS estimate suggests a downward simultaneity bias in the OLS regression.

As well-known in the literature on estimating production-function parameters, one also worries about potential bias due to measurement errors in land area. A couple of recent studies extend the LP method to deal with this problem (see Kim et al. 2016). However, estimators proposed in these studies are based on strong assumptions about error structures and observed inputs. More importantly, such estimators have very slow convergence rates and thus require large sample size data, which are available at the firm level but not for cities. We thus proceed with the understanding that our land coefficient might be biased due to measurement errors. To be cautious, in our counterfactual analysis below, we will use the smallest LP estimate from Table 5: $\hat{\gamma} = 0.155$.[24] This is obtained from the specification in column (2) using a sample with fewer missing values of inputs.

5.4.4 Counterfactual analysis: Reallocating thirty percent of low-productivity cities' land conversion quotas to high-productivity cities

Recall from section 5.1.1 that we classified cities into two groups based on their average productivity of land (APL): There are 199 low APL cities and 80 high APL cities. By design, these two groups converted roughly the same amount of rural land to urban use

[24] Unlike classical measurement errors in standard linear regressions that cause attenuation bias, measurement errors in our LP regressions could bias estimates in either direction.

Table 6: Urban land productivity and rural land rent in low and high APL cities, 2007-2014

Year	High APL cities			Low APL cities		
	(1)	(2)	(3)	(4)	(5)	(6)
	APL_t^H	$MPL_t^H = \hat{\gamma} \cdot APL_t^H$	r_t^H	APL_t^L	$MPL_t^L = \hat{\gamma} \cdot APL_t^L$	r_t^L
2007	776.84	120.41	3.37	332.28	51.50	1.33
2008	864.17	133.95	4.31	372.65	57.76	1.56
2009	946.23	146.67	4.69	403.51	62.54	1.78
2010	1066.14	165.25	5.35	453.70	70.32	1.99
2011	1191.63	184.70	6.27	420.97	65.25	2.32
2012	1267.89	196.52	6.86	549.72	85.21	2.48
2013	1327.91	205.83	7.55	587.41	91.05	2.83
2014	1323.97	205.22	7.77	597.90	92.68	2.84

Unit: ten thousand yuan/hectare. $\hat{\gamma} = 0.155$. APL stands for average productivity of urban land; MPL stands for marginal productivity of urban land; r stands for rural land rent.

(in their respective central cities) from 2007-2014. Note that the average low APL city is much smaller than the average high APL city. In this counterfactual analysis, we reallocate 30 percent of low APL cities' land quotas to high APL cities, which is a rather reasonable scenario.[25]

When we reduce the amount of land converted in low APL cities by 30 percent and reallocate the land conversion quota to the high APL cities, there are two consequences. First, in each year, the converted land in low APL cities decreases by 30 percent, and the converted land in high APL cities increases by the same amount in absolute terms. In relative terms, this reallocation increases high APL cities' annual converted land by 19.53-38.22% in the central cities and by 29.28-64.36% in the whole prefectures. Second, in each year, agricultural land area in low APL cities increases by the same amount in absolute terms, and agricultural land area in high APL cities decrease by the same amount in absolute terms. Thus, the total gain in year t is as follows:[26]

$$\left[(MPL_t^H - MPL_t^L) - (r_t^H - r_t^L)\right] * 0.3 * LC_t^L,$$

[25] We point out here that it is unreasonable to use the condition for efficient allocation in equation (2) as a policy goal. In reality, marginal gains of land conversion may differ across cities for many reasons including, for example, random shocks to land productivity, adjustment costs of land use, and measurement errors of land productivity. These issues are well understood in the literature on resource misallocation among firms (Restuccia and Rogerson 2017).

[26] This calculation assumes that the changes are marginal. While 30% of the land converted is not an insignificant amount, it is very small relative to the whole urban area. If we focus on the central cities, reallocating 30% of the land converted would reduce low APL cities' urban area by 0.95-2.37% and increase high APL cities' urban area by 0.82-2.13%; if we consider the whole prefectures, reallocating 30% of the land converted would reduce low APL cities' urban area by 2.12-6.35% and increase high APL cities' urban area by 1.83-5.61%. Thus it seems reasonable to treat these changes as "marginal." We tried an alternative calculation by continuously adjusting the marginal productivity of land along with reallocation, the results are almost identical.

where MPL_t^H and MPL_t^L are marginal productivity of urban land in year t in high- and low-land-productivity cities, respectively; r_t^H and r_t^L are rural land rent in year t in high- and low-land-productivity cities, respectively; LC_t^L is the total area of land converted for urban use in year t in low-land-productivity cities. Also, it is important to note that the potential gains (or actually loss because the gains were not realized) in each year are not a one-shot deal. We should expect a similar loss in each of the following years due to the original misallocation. Thus the cumulative gain in year t is calculated as

$$\left[(MPL_t^H - MPL_t^L) - (r_t^H - r_t^L)\right] * 0.3 * \sum_{i=2007}^{t} LC_i^L.$$

From equation (7), we know that the marginal productivity of land (MPL) in a city is proportional to the average productivity of land (APL):

$$MPL = \frac{\partial Y_{it}}{\partial L_{it}} = \gamma A_{it} N_{it}^{\alpha} K_{it}^{\beta} L_{it}^{\gamma-1} = \gamma \frac{Y_{it}}{L_{it}} = \gamma * APL. \tag{12}$$

From Table 5, we use the estimated land coefficient $\hat{\gamma} = 0.155$. From the China City Statistical Yearbook, we obtain city-level GDP and total urban land area for low APL cities and high APL cities in each year, so we can calculate the average productivity of land and the marginal productivity of land, presented in columns (1)-(2) and (4)-(5) of Table 6. Note that the differences in marginal productivity of land are much smaller than the differences in average productivity between high and low APL cities because the land coefficient γ is much smaller than one.

Rural land rents in high and low APL cities are calculated according to equation (5) as first-sector GDP per unit of cultivated land, which are shown in columns (3) and (6) of Table 6. For both high and low APL cities, rural land rent is much lower than urban land productivity. As a result, the difference in rural land rent is much lower than the difference in urban land productivity between high and low APL cities. Thus the gain from reallocation of land quotas will be driven primarily by the difference in urban land productivity between high and low APL cities.

The calculated potential gains are presented in Table 7. We consider two cases: Apply the reallocation to all newly converted land in the whole prefectures (columns (1)-(2)) and only to newly converted land in central cities (columns (3)-(4)). In each case, two sets of estimates are calculated, including both annual and cumulative gains. Estimates in column (1) suggest that the reallocation can generate an annual gain equivalent to 0.07-0.21 percent of the country's GDP. The results in column (2) show that the cumulative gains are nearly 1 percent of China's GDP eight years later. Comparing columns (3)-(4) with (1)-(2), we see that about 40 percent of the annual and cumulative gains occur at central cities. Overall,

Table 7: Gains from reallocating 30 percent of land quotas from low to high APL cities

Year	Whole prefectures		Central cities	
	Annual	Cumulative	Annual	Cumulative
	(1)	(2)	(3)	(4)
Total gains (hundred million yuan)				
2007	180.06	180.06	85.20	85.20
2008	229.50	428.34	94.32	188.40
2009	414.78	887.42	191.25	399.13
2010	556.54	1558.15	219.18	669.67
2011	1033.12	2996.76	423.40	1267.35
2012	1110.42	3895.11	420.52	1598.18
2013	1188.79	5195.26	417.93	2061.80
2014	929.65	6020.39	355.01	2375.34
Total gains as percentage of China's GDP				
2007	0.07%	0.07%	0.03%	0.03%
2008	0.07%	0.14%	0.03%	0.06%
2009	0.12%	0.26%	0.06%	0.12%
2010	0.14%	0.38%	0.05%	0.16%
2011	0.21%	0.62%	0.09%	0.26%
2012	0.21%	0.73%	0.08%	0.30%
2013	0.20%	0.88%	0.07%	0.35%
2014	0.15%	0.95%	0.06%	0.37%

these calculations suggest that the cumulative effects from a few years of misallocation can be substantial, which might be a reason why the Chinese economy has slowly regressed to "a new normal" in recent years.

We need some benchmarks to assess how important these welfare gains are. One such benchmark is available in the international trade literature. Krugman (1979) builds a model to show that new varieties are an important source of gains from trade. Broda and Weinstein (2006) find the variety gains to be 0.1 percent of GDP in the U.S., and Chen and Ma (2012) show that the welfare gain from new import varieties amounts annually to 0.4 percent of GDP in China. Since the potential welfare gain from land reallocation across cities in China is in the same order of magnitude as these import variety gains, it is quite large and economically significant. A second benchmark comes from a more recent study of the United States. Using data from 220 U.S. metropolitan areas from 1964 to 2009, Hsieh and Moretti (2019) calibrate a spatial equilibrium model to quantify the effect of spatial misallocation due to housing supply constraints. They find that if the housing supply in New York, San Jose, and San Francisco increased by relaxing land use restrictions to the level of the median U.S. city, it would increase U.S. GDP by 3.7 percent. This implies an annual gain of 0.08 percent of GDP, which lies in the lower range of our results.

Finally, we discuss how the way we estimate the city-level production function might have affected the results of the counterfactual exercise. We assumed the Cobb-Douglas function form to allow for a straightforward calculation of marginal productivity of land from its average productivity. In addition, we imposed that the coefficient of log land for all cities is the same. This may be violated. In particular, in cities like Shenzhen and Xiamen where the expansion of the urban area is geographically constrained, marginal productivity of land may have a different relationship with the average productivity of land.[27] To explore this issue, we first examine how many cities are seriously constrained by limited availability of convertible land. For each city, we calculate the ratio of total converted land area (2007-2016) to the total area of cultivated land available in 2007 in the whole prefecture. It appears that for most cities, this ratio is very small, with an average of 7.34 percent. That is, there is still a lot of land that can be easily converted for urban use, even for megacities such as Beijing and Shanghai. We tried dropping ten cities with the highest values of this ratio (i.e., the most land-constrained cities) and reestimating equation (9). The estimated coefficients of log land are similar to those in Table 5. The LP estimate using water to control for unobserved productivity is now 0.178, compared to 0.155 in column (2) of Table 5; the smallest LP estimate, from the specification in column (3) of Table 5, is now 0.151. Thus our main results are not driven by the most geographically constrained cities.

[27] We thank a referee for raising this point.

We next tried adding to equation (9) an interaction term between log land and a city characteristic, allowing the land coefficient to vary with city characteristics. We tried, one by one, to interact log land with provincial capital dummy, log population, log per capita GDP, log per capita government revenue, urbanization rate of land (share of urban area in total land area), availability of convertible land (total area of converted land from 2007-2016 divided by area of cultivated land in 2007), and log average urban land productivity. Note that one cannot easily adapt the LP method to allow for an interaction term, so we focus on OLS regressions in this exercise.[28] Instead of obtaining consistent estimates, our goal here is to explore the significance of heterogeneity. In all but two cases, the coefficient of the interaction term between log land and the city characteristic is statistically significant. The two exceptions are the interaction with availability of convertible land and average land productivity. However, the effect on the direction of bias is rather ambiguous. For example, the interaction terms suggest that province capital, cities with larger population, and cities with higher per capita government revenue all tend to have larger land coefficients. Recall from Table 3 that these cities also have higher urban-rural land price gaps. Together these imply that imposing the same value of land coefficient could underestimate potential gains from reallocating land quotas to such cities. On the other hand, we find the interaction term with per capita GDP has a negative coefficient, suggesting a smaller land coefficient for these cities and overestimation of gains from reallocating land quotas to richer cities. Overall, our extensive experimentation suggests that estimating a more flexible city-level production function would improve our counterfactual analysis. However, this requires not only better econometric methods but also more and higher-quality data, which we leave for future work.

5.5 Why does the Chinese government allocate so much land quotas to low-productivity cities?

In this subsection, we first investigate whether cities with higher land revenues tend to have less land converted in the following year over the period 2007-2016. We hypothesize that upper level governments cannot easily transfer land revenue from one jurisdiction to another, and thus they use land quotas to rein in growing inter-jurisdictional inequalities of land revenue that cannot be easily justified. This implies that if land revenue is growing too fast in a city, relative to its own or other cities' growth path, the city tends to get a lower quota next year.

From the land transaction data, we obtain land areas converted for urban use (LC_{it}) and

[28] Results from this exercise are available from the authors upon request.

Table 8: Area of land converted and lagged land revenue

Dependent variable: $\triangle$ ln(area of land converted for urban use)					
	Full sample	Com., Res. & Ind.	Industrial	Residential	Commercial
	(1)	(2)	(3)	(4)	(5)
Lagged $\triangle$ ln(land revenue)	-0.167***	-0.202***	-0.314***	-0.221***	-0.272***
	(0.021)	(0.016)	(0.017)	(0.014)	(0.016)
City fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Constant	0.373***	0.382***	0.477***	0.412***	0.305***
	(0.045)	(0.037)	(0.048)	(0.049)	(0.067)
N	2,766	2,766	2,681	2,632	2,608
R^2	0.150	0.245	0.248	0.239	0.222

Standard errors are in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

land revenue (LR_{it}) for each city in each year. To test our hypothesis, we regress change in log land area converted on lagged change in log land revenue at the city level:

$$\triangle \ln LC_{it} = \alpha + \beta \triangle \ln LR_{it-1} + C_i + \tau_t + \varepsilon_{it} \tag{13}$$

where C_i is a city fixed effect; τ_t is a year fixed effect; and ε_{it} is the error term.

Table 8 reports the results. In column (1), we include the full sample of land parcels. In columns (2)-(5), we use the subsamples of commercial, residential, and industrial land, either together or separately. Our results show consistent negative coefficients of lagged log land revenue. We also tried dropping the city fixed effects, estimating the coefficients with cross-city variations; the results are similar in that lagged log land revenue has a negative and statistically significant coefficient in every column. This is suggestive evidence that the Chinese government is trying to curb rising inequalities of local land revenue: If land revenue grows faster in a city, less land will be converted for urban use in the following year so that land revenue is not growing out of proportion relative to other cities.[29] Given that high-productivity cities/regions experienced faster housing and land price growth in this period, this way of adjusting quotas would allocate less land to those cities.

Another possible reason for land misallocation is that land quota is used as a policy tool to control population growth in larger cities. The Chinese government has a long standing policy to control population growth in large cities and at the same time invest more resources in large cities (Xing and Zhang 2017).[30] As a result, land tends to be more productive in

[29]This policy operation assumes that local land demand is elastic enough so that a reduced land supply will not increase land revenue, which seems reasonable.

[30]Starting in the 1980s, China officially pursued a policy to contain the scale of large cities, which has repeatedly featured in central government plans. For recent examples, see the Twelfth Five-

larger cities. For example, using data from 2014, the correlation coefficient between city population and average productivity of urban land is 0.405. Given this, if land supply is more tightly controlled in larger cities to contain population growth, it leads to land misallocation across cities. Recall from Table 3 that cities with a larger population tend to have a bigger urban-rural land price gap, which is consistent with the notion that land quota is used to control the growth of larger cities.

Although both explanations may be true in reality, it is important to figure out the main reason behind the misallocation of land across cities in China, because it has implications for policy solutions to the problem. For example, if indeed the central government allocated more land quotas to inland provinces only to guarantee certain amount of land revenue for those provinces, then misallocation of land can be easily avoided by some cap-and-trade type of system that allows for buying and selling land quotas among local jurisdictions. Indeed this cap-and-trade type of policy was experimented in Zhejiang province during the early 2000s (see Chau et al. 2016). A policy with similar features was also tried in Chongqing and Chengdu (Xiao 2015). Finally, starting in March 2018, the central government decided to allow trading of land conversion quotas across provinces, which should help alleviate the misallocation problem. However, if the misallocation is a result of controlling growth in larger cities or promoting urbanization in less developed regions, then the welfare loss is inevitable unless the government reverses the policy.

6 Conclusion

Using a large data set of land transactions, we document various facts about land conversion for urban uses in China. We find that revenue from selling land leaseholds amounts to more than one-third of local governments' total revenue during 2007-2014. An increasingly larger share of land is converted for urban uses in low productivity regions or cities. There is evidence for land misallocation in this period. We find that the urban-rural land price gap and the marginal productivity of urban land varies substantially across cities. Our counterfactual analysis shows that the potential gains from reallocating land quotas from low- to high-productivity cities are economically significant.

Year Plan of China that was passed in 2011 (available at http://www.china.com.cn/policy/txt/2011-03/16/content_22156007.htm). It clearly stated the policy goal of controlling the population size of the largest cities and regulating urban land use to prevent "over-expansion" of such cities. The National Plan for a New-Style Urbanization 2014-2020, released in 2014, reiterated the same principles (available at http://www.gov.cn/zhengce/2014-03/16/content_2640075.htm). These principles are followed by lower level governments. For example, according to the Beijing City Master Plan (2004-2020), in 2020 the population of Beijing would be controlled to under 18 million and its urban built-up area would be controlled to under 1,650 square kilometers.

Our analysis takes the allocation of other productive factors across cities as given and focuses exclusively on land. It is possible that there are also other sources of misallocation. For example, China has long used the Hukou system to control internal migration. There might be serious misallocation of workers both between rural and urban areas and across cities (Au and Henderson 2006a, 2006b). Moreover, there is also evidence that as a result of political favoritism and place-based subsidies, different cities in China face different prices of capital, leading to a misallocation of physical capital across cities (Chen et al. 2017b; Yang et al. 2017). Solving all of these misallocation problems simultaneously should generate even higher welfare gains (Henderson et al. 2020).

One limitation of our research is that our discussion and calculation has ignored non-market benefits and costs. For example, one might argue that rapid urbanization in coastal regions poses a threat to the environment and that the land quota system helps slow down the development and preserve the ecological balance in such regions. One might also believe that the land quota system helps maintain regional balance in urbanization that is socially desirable. It is a methodological challenge to incorporate such non-market benefits and costs in our analysis. If indeed there are worthy causes for the allocation of more land quotas to inland regions and less productive cities, then our estimated losses should be interpreted as the economic costs of such policy goals.

Appendix A: Sample construction using the China Land Transaction Data

There are 1,941,657 observations in the full data set. Table A.1 describes the steps we followed to create our analysis sample. First, since there are no city-level variables in the land transaction data, we merge it with a data set that contains city characteristics, resulting in 1,914,927 observations. Second, if two observations have identical province, city, district, transaction ID, project name, contract date, land price, and land area, we consider them as duplicates and only keep one observation. Third, we drop 132 observations that the contract year in these observations is missing. Fourth, since the website was launched in early 2008, the coverage of pre-2007 deals is very incomplete, we thus drop all the observations before 2007. Fifth, we drop 401 observations in which price is negative or land area is non-positive. Sixth, we drop a total of 15,427 top 1 percent outliers based on land price in each city-year. For each city-year, the top 1 percent are separately identified and excluded from our empirical analysis. While this may lead to an underestimation of totals, it is necessary because some prices are unbelievably high (likely a result of wrong units used in data recording). Seventh,

Table A.1: The steps to create our analysis sample

Sample Selection	Number of Obs.
Full data set	1,941,657
Successfully merged with administrative unit identifiers	1,914,927
231,031 duplicate observations deleted	1,683,896
132 observations without contract year deleted	1,683,764
125,629 pre-2007 observations deleted	1,558,135
401 obs. with negative price or non-positive land area deleted	1,557,734
15,427 top 1% outliers in each city-year deleted	1,542,307
28 outliers (price > 500,000 yuan/m^2 or area > 20,000 ha anywhere, or price > 80,000 yuan/m^2 in Xinjiang in 2009) deleted	1,542,279

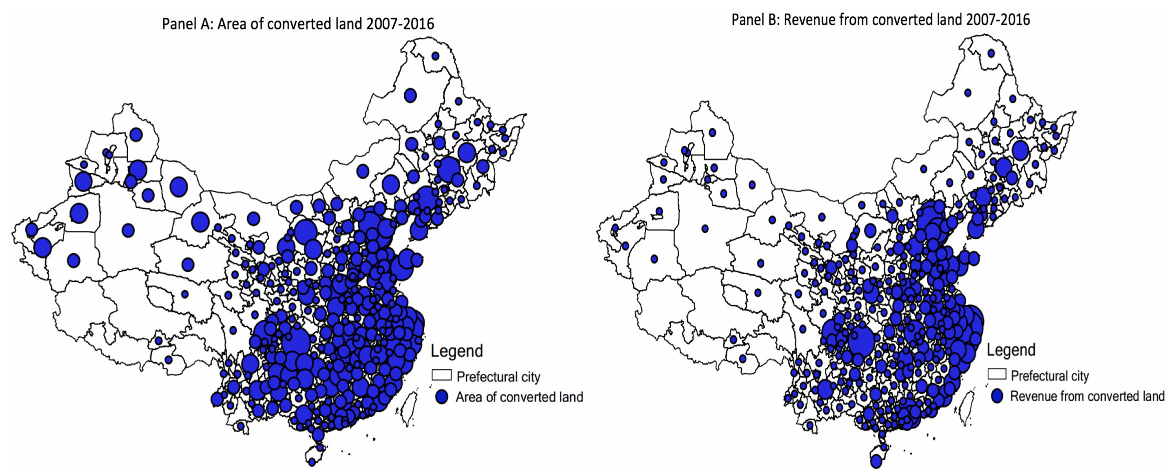


Figure A.1: Land conversion across prefectures in China, 2007-2016

we drop 24 outliers in which price is larger than 500,000 yuan per square meter or land area is larger than 20,000 hectares. Finally, we drop 4 more outliers in Xinjiang in 2009 where the land price is unbelievably high (by local standard). We end up with 1,542,279 land transaction deals after these steps.

Figure A.1 shows where land conversion had happened, based on the China Land Transaction data. Panel A shows the total area of converted land in each prefecture during 2007-2016. Land conversion was clearly occurring all over the country, except in the most mountainous, sparsely-populated regions. Panel B shows the total revenue generated from leases on converted land in each prefecture during 2007-20016. Prefectures in coastal provinces generated much more revenue.

Appendix B: Assessing the China Land Transaction Data

The quality of the China Land Transaction Data is crucial for the reliability of our empirical findings. We here try to assess the data by comparing it with the only alternative data source available: provincial-level data from the China Land and Resources Statistical Yearbook compiled by the Ministry of Land and Resources of China. The yearbook data also break down their statistics by transaction type, which indicates that allocated land is not included. Recall that allocated land is directly granted by the local government for purposes of building schools, hospitals, parks, and so on. They are not included in the yearbook data perhaps because they do not generate any government revenue. In our comparison here, we also drop allocated land from our China Land Transaction Data.

In Panel A of Figure A.2, we compare log total area of land transaction in each year from the two data sources. We see that the two series follow the same trend and are very close to each other. In Panel B we compare log total land revenue in each year calculated from the two data sources. They also follow the same trend, but the total from our China Land Transaction Data is always smaller. We suspect that this is a result of dropping top outliers based on price. However, we find that adding back the outliers only makes this difference smaller, but cannot eliminate this difference completely. In other words, our data may underestimate land revenue. In Panels C and D, we compare province-year level log land area and log land revenue respectively. Total area or revenue from our China Land Transaction Data is on the vertical axis and the corresponding total from the statistical yearbook is on the horizontal axis. If they perfectly coincide, all of the dots should be on the (solid) 45-degree line. In Panel C, there is no discernible difference between the fitted (dash) line and the 45-degree line. In Panel D, the fitted (dash) line is slightly below the 45-degree line, suggesting that provincial-level total land revenues calculated from the China Land Transaction Data are smaller than the aggregates published by the government.

In Figure A.3, we break down the total land area and land revenue series from each data source by coastal and inland regions. To make the relative scales clearer, we present the aggregates here in levels (instead of logs). Again, we see that the series calculated from our China Land Transaction Data (solid lines) follow those from the yearbook data (dash lines) rather closely. The largest discrepancy occurs in 2016, the last year in our analysis period, where total areas and revenues from our China Land Transaction Data are considerably smaller. Since our data were scraped from the website in early 2017 and the yearbook data were published with much longer lags, the latter might have included more updates for the year 2016. Given that the yearbook data started in 1999, we also included in this figure the yearbook data from those earlier years. Panel A clearly shows that coastal regions' share of total area of land transaction peaked in 2003 and then continuously declined for a whole

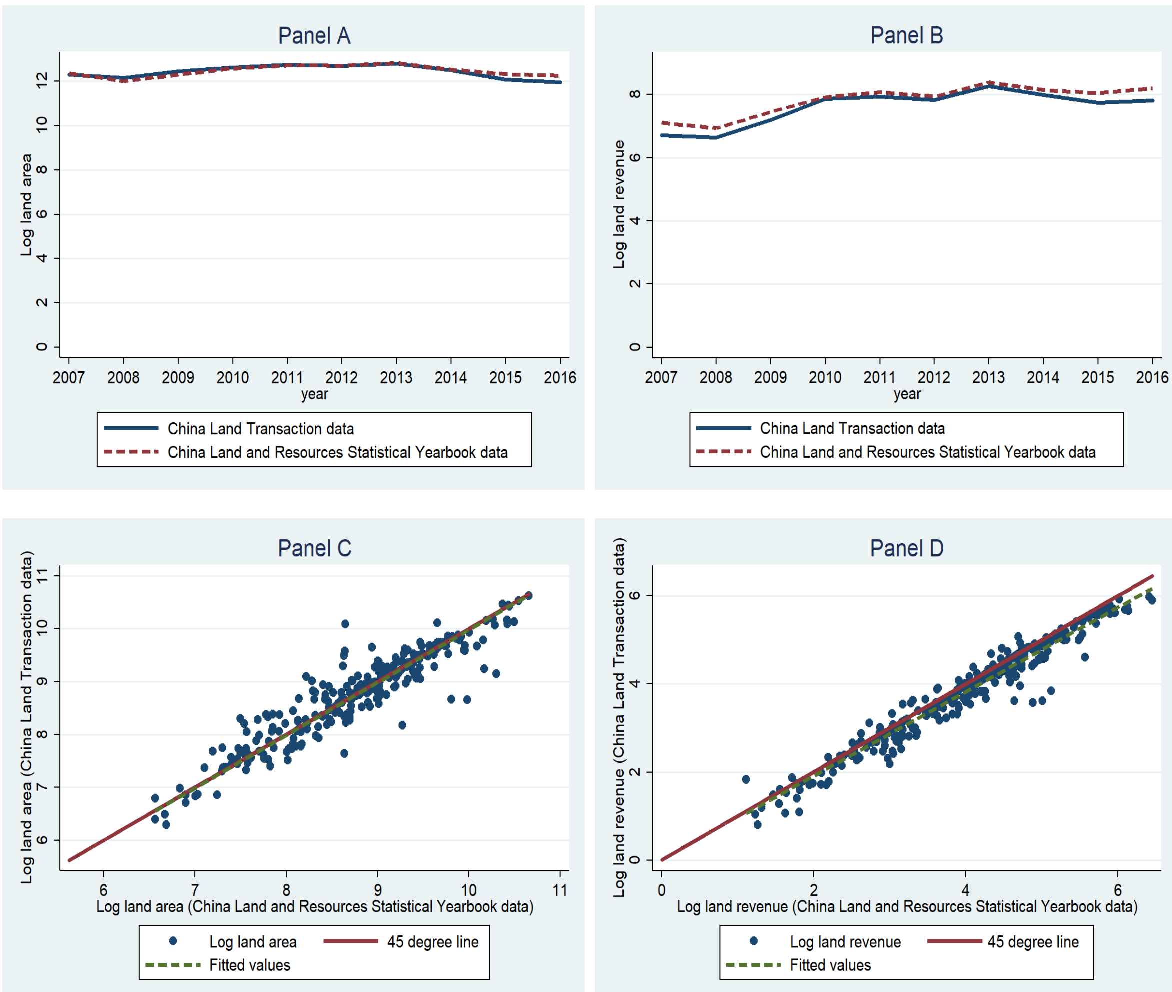
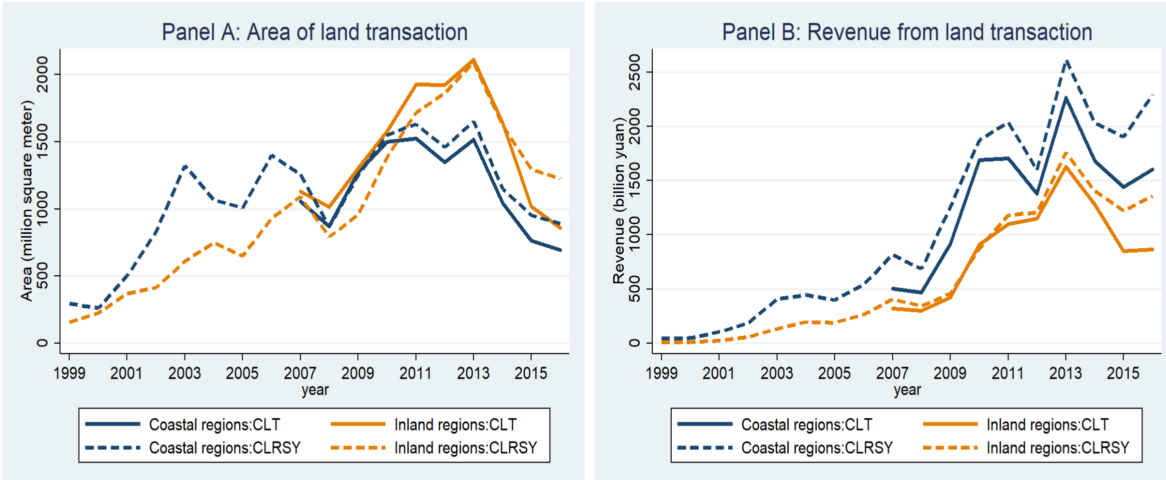


Figure A.2: Comparing land transaction data with alternative data source, 2007-2016



Note: Solid lines are series calculated from our China Land Transaction (CLT) data, and dash lines calculated from the China Land and Resources Statistical Yearbook (CLRSY) data.

Figure A.3: Comparing land transaction data with alternative data source, by coastal and inland regions

decade, a fact emphasized by Lu (2016) and Han and Lu (2017).

Overall, our analysis here indicates that if the aggregate data from the China Land and Resources Statistical Yearbook can be trusted, then the China Land Transaction Data are reasonably good. The data we collected produce identical trends as the publicly available yearbook data; they give very similar aggregate statistics on land area; they may underestimate land price and revenue, which one should keep in mind when interpreting our results.

Appendix C: Calculation of the misallocation index

The misallocation index in equation (1) is defined as $M_t = -n_t * cov_t\left(s_{it}, A_{it}\right)$, where n_t is the number of cities in year t.

To calculate s_{it}, city i's share of all urban land in year t, we first obtain each city's built-up area in 2007 from the China City Statistical Yearbook. From the land transaction data, we calculate the total area of land converted for urban uses in each city in each year. These newly converted land areas are added to the 2007 built-up area to obtain land area in each city in each year, which are then used to calculate each city's land share in each year.[31]

To measure the total factor productivity A_{it} for each city, we estimate the following

[31] An alternative way to calculate this share is to use each city's built-up area, directly from the yearbook data. However, in recent years many cities expanded by "redistricting": Many small towns in surrounding counties are "acquired" by city districts and become part of the central city. This will introduce artificial changes to the built-up area of a city, making the calculation of the misallocation index imprecise. Thus we do not use the annual data on built-up area from the yearbooks.

production function for all cities using the yearbook data:

$$\ln Y_{it} = \ln A_{it} + \alpha \ln N_{it} + \beta \ln K_{it} + \gamma \ln L_{it},$$

where Y_{it} is the output level; A_{it} is a productivity parameter; N_{it} is the number of workers; K_{it} is the capital stock; and L_{it} is the quantity of urban land. We estimate the parameters of the production function in a two-way fixed effects model and use the method pioneered by Levinsohn and Petrin (2003) to control for unobserved productivity shocks, which is detailed in section 5.4.3. Total factor productivity for each city in each year is calculated from the residuals not explained by the three factors of production:

$$A_{it} = \exp\left(\ln Y_{it} - \hat{\alpha} \ln N_{it} - \hat{\beta} \ln K_{it} - \hat{\gamma} \ln L_{it}\right),$$

where the estimated parameters, $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$, are from column (2) of Table 5.

References

[1] Albouy, David. 2009. "The Unequal Geographic Burden of Federal Taxation." *Journal of Political Economy* 117(4), 635-667.

[2] Adamopoulos, Tasso, Loren Brandt, Jessica Leight, and Diego Restuccia. 2017. "Misallocation, Selection and Productivity: A Quantitative Analysis with Panel Data from China." NBER Working Paper No. 23039.

[3] Adamopoulos, Tasso and Diego Restuccia. 2014. "The Size Distribution of Farms and International Productivity Differences." *American Economic Review* 104(6), 1667-1697.

[4] Adamopoulos, Tasso and Diego Restuccia. 2015. "Land Reform and Productivity: A Quantitative Analysis with Micro Data." University of Toronto Working Paper.

[5] Anderson, Gordon and Ying Ge. 2005. "The Size Distribution of Chinese Cities." *Regional Science and Urban Economics* 35(6), 756-776.

[6] Au, Chun-Chung and J. Vernon Henderson. 2006a. "Are Chinese Cities Too Small?" *Review of Economic Studies* 73(3), 549–576.

[7] Au, Chun-Chung and J. Vernon Henderson. 2006b. "How Migration Restrictions Limit Agglomeration and Productivity in China." *Journal of Development Economics* 80, 350–388.

[8] Broda, Christian and David E. Weinstein. 2006. "Globalization and the Gains from Variety." *Quarterly Journal of Economics* 121(2), 541-585.

[9] Brueckner, Jan K., Shihe Fu, Yizhen Gu, and Junfu Zhang. 2017. "Measuring the Stringency of Land-Use Regulation: The case of China's Building-Height Limits." *Review of Economics and Statistics* 99(4), 663-677.

[10] Cai, Hongbin, J. Vernon Henderson, and Qinghua Zhang. 2013. "China's Land Market Auctions: Evidence of Corruption?" *RAND Journal of Economics* 44(3), 488-521.

[11] Cai, Hongbin, Zhi Wang, and Qinghua Zhang. 2017. "To Build above the Limit? Implementation of Land Use Regulations in Urban China." *Journal of Urban Economics* 98, 223-233.

[12] Cao, Guangzhong, Changchun Feng, and Ran Tao. 2008. "Local 'Land Finance' in China's Urban Expansion: Challenges and Solutions." *China & World Economy* 16(2), 19-30.

[13] Chau, Nancy H., Yu Qin, and Weiwen Zhang. 2016. "Leader Networks and Transaction Costs: A Chinese Experiment in Interjurisdictional Contracting." IZA Discussion Paper No. 9641.

[14] Chen, Bo and Hong Ma. 2012. "Import Variety and Welfare Gain in China." *Review of International Economics* 20(4), 807-820.

[15] Chen, Chaoran, Diego Restuccia, and Raül Santaeulàlia-Llopis. 2017a. "The Effects of Land Markets on Resource Allocation and Agricultural Productivity." NBER Working Paper No. 24034.

[16] Chen, Ying, J. Vernnon Henderson, and Wei Cai. 2017b. "Political Favoritism in China's Capital Markets and its Effect on City Sizes." *Journal of Urban Economics* 98, 69-87.

[17] Chen, Zhihong, Shihe Fu, and Dayong Zhang. 2013. "Searching for the Parallel Growth of Cities in China." *Urban Studies* 50(10), 2118-2135.

[18] Duranton, Gilles, Ejaz Ghani, Arti Grover Goswami, and William Kerr. 2015. "The Misallocation of Land and Other Factors of Production in India." Policy Research Working Paper 7221, The World Bank.

[19] Fang, Hanming, Quanlin Gu, Wei Xiong, and Li-An Zhou. 2015. "Demystifying the Chinese Housing Boom." in *NBER Macroeconomics Annual* vol. 30, pp. 105-166, edited by Martin Eichenbaum and Jonathan Parker. University of Chicago Press.

[20] Fang, Li, Peng Li, and Shunfeng Song. 2017. "China's Development Policies and Urban Dynamics: An Analysis based on the Zipf Law." *Urban Studies* 54(12), 2818-2834.

[21] Fei, Xuan. 2020. "The Misallocation in the Chinese Land Market." BOFIT Discussion Papers 23/2020.

[22] Fu, Shihe, Yu Liao, and Junfu Zhang. 2016. "The Effect of Housing Wealth on Labor Force Participation: Evidence from China." *Journal of Housing Economics* 33, 59-69.

[23] Ghatak, Maitreesh and Dilip Mookherjee. 2014. "Land Acquisition for Industrialization and Compensation of Displaced Farmers." *Journal of Development Economics* 110, 303-312.

[24] Glaeser, Edward, Wei Huang, Yueran Ma, and Andrei Shleifer. 2017. "A Real Estate Boom with Chinese Characteristics." *Journal of Economic Perspectives* 31(1), 93-116.

[25] Gyourko, Joseph and Raven Molloy. 2015. "Regulation and Housing Supply." In: Gilles Duranton, J. Vernon Henderson and William Strange (eds.) *Handbook of Regional and Urban Economics*, Vol. 5, Elsevier, 1289-1337.

[26] Han, Libin and Ming Lu. 2017. "Housing Prices and Investment: An Assessment of China's Inland-favoring Land Supply Policies." *Journal of the Asia Pacific Economy* 22(1), 106-212.

[27] Henderson, J. Vernon, Dongling Su, Qinghua Zhang, and Siqi Zheng. 2020. "The Costs of Political Manipulation of Factor Markets in China." CEPR Discussion Paper No. DP15247.

[28] Hopenhayn, Hugo. 2014. "Firms, Misallocation, and Aggregate Productivity: A Review." *Annual Review of Economics* 6, 735-770.

[29] Hsieh, Chang-Tai and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics* 124(4), 1403-1448.

[30] Hsieh, Chang-Tai and Enrico Moretti. 2019. "Housing Constraints and Spatial Misallocation." *American Economic Journal: Macroeconomics* 11(2), 1-39.

[31] Kim, Kyoo il, Amil Petrin, and Suyong Song. 2016. "Estimating Production Functions with Control Functions When Capital Is Measured with Error." *Journal of Econometrics* 190(2), 267-279.

[32] Krugman, Paul R. 1979. "Increasing Returns, Monopolistic Competition, and International Trade." *Journal of International Economics* 4, 469-479.

[33] Levinsohn, James and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *Review of Economic Studies* 70(2), 317-341.

[34] Li, Hongbin and Li-An Zhou. 2005. "Political Turnover and Economic Performance: The Incentive Role of Personnel Control in China." *Journal of Public Economics* 89(9-10), 1743-1762.

[35] Lu, Ming. 2016. *Great State Needs Bigger City*. Shanghai People's Press. (In Chinese)

[36] Olley, G. Steven and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunication Equipment Industry." *Econometrica* 64(6), 1263-1297.

[37] Qin, Yu, Hongjia Zhu, and Rong Zhu. 2016. "Changes in the Distribution of Land Prices in Urban China during 2007–2012." *Regional Science and Urban Economics* 57, 77-90.

[38] Restuccia, Diego and Richard Rogerson. 2013. "Misallocation and Productivity." *Review of Economic Dynamics* 16(1), 1-10.

[39] Restuccia, Diego and Richard Rogerson. 2017. "The Causes and Costs of Misallocation." *Journal of Economic Perspectives* 31(3), 151–174.

[40] Restuccia, Diego and Raül Santaeulàlia-Llopis. 2017. "Land Misallocation and Productivity." NBER Working Paper No. 23128.

[41] Tan, Minghong, Xiubin Li, and Changhe Lv. 2003. "An Analysis of the Driving Forces Behind Urban Land-Use Expansion in Our Country." Economic Geography 23(5), 635-639. (In Chinese)

[42] Tan, Rong, Futian Qu, Nico Heerink, and Evy Mettepenningen. 2011. "Rural to Urban Land Conversion in China—How Large Is the Over-Conversion and What Are Its Welfare Implications?" *China Economic Review* 22(4), 474-484.

[43] Wang, Zhi, Qinghua Zhang and Li-An Zhou. 2020. "Career Incentives of City Leaders and Urban Spatial Expansion in China." *Review of Economics and Statistics* 102(5), 897-911.

[44] Wu, Jing, Joseph Gyourko, and Yongheng Deng. 2016. "Evaluating the Risk of Chinese Housing Markets: What We Know and What We Need to Know." *China Economic Review* 39, 91-114.

[45] Xiao, Yuan. 2015. "Making Land Fly: Land Quota Markets in Chinese Urbanization." Working Paper #22, NYU Marron Institute of Urban Management.

[46] Xie, Zhikui. 2015. *Administration upon Spring: Institutional Analysis of the Land Administration in China*. The Commercial Press. (In Chinese)

[47] Xing, Chunbing and Junfu Zhang. 2017. "The Preference for Larger Cities in China: Evidence from Rural-Urban Migrants." *China Economic Review* 43, 72-90.

[48] Yang, Shangming, Yuming Fu, and Yi Lu. 2017. "The Spatial Misallocation of Capital and Labor in China: A Welfare Analysis Based on Urban Accounting." Unpublished manuscript, Department of Real Estate, National University of Singapore.

[49] Zhang, Junfu. 2011. "Interjurisdictional Competition for FDI: The Case of China's 'Development Zone Fever.'" *Regional Science and Urban Economics* 41(2), 145-159.