Land Offers and Fiscal Competition Between City Governments in China*

Wending Liu†

July, 2023

Abstract

I analyze the fiscal competition between city governments in China by structurally estimating a Bertrand pricing game model. The model characterizes the land pricing strategy of city governments as they use land sales discounts to attract industrial firms. The estimation results imply that city governments can generate a huge amount of fiscal revenue from landing industrial firms, which is around 45% of the firm's yearly output. By counterfactual experiments, I show that the impact of this kind of fiscal competition on resource allocation is small. Simulation results also show that fiscal centralization and increasing urban wages would result in a modest average land price increase.

Keywords: Fiscal competition, land market, Bertrand game, structural estimation

JEL codes: C57, H71

*I am indebted to my supervisor Fedor Iskhakov for his continuous guidance on this paper. I acknowledge comments from Xin Meng, Bob Gregory, Ruitian Lang, and John Stachurski at the applied microeconomics seminar at ANU.

†Australian National University `Wending.Liu@anu.edu.au`

1 Introduction

In this paper, I analyze the fiscal competition between city governments in China as they use industrial land sales discounts to attract industrial firms to land in their geographical jurisdictions. Understanding this process is important since the unique "regionally decentralized authoritarian regime" (Xu, 2011) plays an important role in China's economic growth. This regime is unusual in combining a high degree of political centralization and economic decentralization, and it gives local officials great autonomy in economic and fiscal issues (Kroeber, 2020; Xiong, 2018). Thus, local officials have strong incentives to develop the local economy as well as maximize their fiscal revenues. One of the most salient characteristics of the system is the competition among city governments, which use special deals to attract businesses (Bai et al., 2020). A major type of special deals is selling industrial lands at low prices to firms (Su and Tao, 2017). Thus, my study helps to understand the economic system of China, and it also sheds light on broader public economics issues, especially the mechanism and impacts of fiscal competition caused by the decentralization of economic and fiscal power.

Local governments can generate fiscal revenue by landing industrial firms. The fiscal revenue includes tax revenue, promotion of local businesses and housing market, political benefits, etc. However, most of them except the tax revenue cannot be measured directly from the data. Thus, the total fiscal revenue generated by landing industrial firms is difficult to be measured directly. I address this issue by structurally estimating a Bertrand pricing game model that characterizes the fiscal competition between Chinese city governments. In this model, city governments maximize their fiscal revenue by providing special industrial land price offers to attract firms, and they need to consider competitor city governments' strategies when they make land price offers. The basic trade-off in the model is between the higher probability of getting the firm and higher land-selling revenue. I solve the continuous Bertrand pricing game using the Gauss-Seidel algorithm, and structurally estimate the key parameters of this model by the method of simulated moments (MSM), which is numerically implemented by a polyalgorithm combining the grid search and Quasi-Newton method.

The estimation results show city government can generate a huge amount of fiscal revenue from landing industrial firms. The total fiscal revenue is around 45% of the firm's per year output level, which is far beyond the official tax revenue. This finding shows the fiscal spillover effects of landing industrial firms for city governments are enormous, and that's why local governments in China compete against each other so fiercely to attract industrial firms. I use counterfactual experiments to study the impacts of this fiscal competition on resource allocation, and the results suggest the impact is small assuming the total output level is fixed. I also simulate the impacts of fiscal power re-centralization and rising wages on this kind of fiscal competition. The simulation results show that both fiscal power re-centralization and the rising urban wage will raise industrial land prices moderately.

A large strand of literature studies the decentralization of economic and fiscal power in China, particularly, Cheung (2014), Su and Tao (2017), Bai et al. (2020) and Liu and Xiong (2020) all notice that local governments in China use land sales discounts to attract firms. Chen et al. (2017) constructs a land price index in China and finds that price of industrial land in China is significantly lower than commercial and residential land. Bai et al. (2020) also confirms the fact by regression analysis. But these papers don't build any formal economic model to explain this issue. For general theoretical discussion on the impacts of fiscal competition, literature built on Tiebout (1956) emphasizes the welfare improvement effect caused by competition for mobile capital, which creates efficient equilibrium. However, literature on the tax competition models (Keen and Marchand, 1997; Wilson, 1999) emphasizes the downward pressure on fiscal revenue induced by fiscal competition and the possibilities of "race to the bottom". See Wilson (1999) for a literature survey for this research area.

For the structural estimation, my work is closest to Mast (2020), which estimates a Bayesian game between towns in the U.S. as they use tax breaks to bid for firms. Two key differences are that I explicitly model the cost minimization problem of the firms, and I solve the game on a continuous space rather than on grids. The numerical methods for solving the game and estimation are both fast.[1]

The outline of the paper is as follows: In Section 2 I examine the evolution of fiscal competition between local governments in China and its relation with the land market. Section 3 presents my model and the numerical method to solve the model. Section 4 and Section 5 describe the data and my estimation method. Section 6 shows the estimates and the fit of my model. Section 7 shows the counterfactual analysis based on the estimates. And Section 8 concludes.

2 Fiscal Competition and Land Market in China

As discussed in Section 1, city governments in China use low land prices to attract industrial firms to produce in their jurisdictions. In this section, I briefly review the evolution of this kind

[1]It costs around 15 minutes to get the main estimation results in Section 6 on the author's PC with Core i7-1165G7 (2.80GHz) CPU and 16 GB RAM.

of fiscal competition and its impact on China's land market from a historical perspective, which will provide a basis for the model introduced in Section 3.

2.1 The evolution of fiscal competition between local governments

The reason why Chinese local governments compete fiercely against each other to attract industrial firms is deeply rooted in China's fiscal system, especially the fiscal revenue sharing scheme between the central government and local governments. China began its economic reform as a very poor country in 1978, and the central government used "fiscal contracting" system in the 1980s to incentivize local governments to develop the local economy. Under this system, local governments got an increasing marginal share of the fiscal revenue they collected. Moreover, they often colluded with local state-owned enterprises (SOEs), which generated most of the tax revenue in the 1980s, to hide the fiscal revenue from the central government (Su and Tao, 2017).

The "fiscal contracting" system led to a continuous decline of the central government's share of total fiscal revenue and threatened the stability of China's macro economy though it enriches the local governments (Liu and Xiong, 2020). The Tax-sharing Reform in 1994 changed this situation drastically by reconstructing the tax system. Under this new system, fiscal revenue sharing based on fiscal contracting is abolished, and the central government takes the larger part of tax revenue collected by local tax bureaus, which are controlled by the central government directly since then. The tax revenue left to local governments comprises 25% of value-added tax (VAT), business tax, and income tax.[2] Meanwhile, local governments had to accept more obligations for public spending. To make the matter worse, low-efficient local SOEs were hard to survive under the competition with private and foreign firms, which are thriving in China since the mid-1990s. As a result of these dramatic changes in China's fiscal system, local governments had to find new ways to generate more fiscal revenues to make up for their fiscal shortfall. The natural choice for local governments is to attract private (including foreign) firms to land in their jurisdictions. In addition to the benefits of tax revenue generated by landing private firms, huge fiscal spillover effects include the promotion of local businesses since more workers come to the city, and the land sale revenue promoted by the increasing demand in the residential and commercial housing market, which will be discussed in the next subsection. The strong incentives for attracting firms lead to fiscal competition between local governments. Selling industrial land at low prices is the most common method of attracting firms between competing cities.

[2] The central government also takes 60% of income tax since 2003.

2.2 The land market in China

All urban land in China legally belongs to the state, however, local governments, especially city and county governments have de facto ownership over land in their jurisdictions.[3] Thus, local governments are the de facto monopolist of land supply in their jurisdictions though the upper bound of land supply is dependent upon the land quota set by the central government.

For residential and commercial land, as the monopolist, the local governments tend to rise the price by restricting the supply of land of these two types. According to the national land price index constructed by Liu and Xiong (2020), the price of commercial land in China rose from 1 in 2004 to 6.11 in 2015, and which of residential land rose from 1 to 4.75 during the same period.[4]

But for industrial land, the story is different. According to Liu and Xiong (2020), the price index of industrial land in China was just 1.5 in 2015 (starting from 1 in 2004). Nevertheless, suppressing industrial land prices is also a rational strategic choice of local governments. By making special deals with firms, the local governments use low land prices to attract firms to their jurisdiction.[5] And because a firm is always attracted by several cities, the firm will often get an upper hand in bargaining with city governments. Thus, the land price will become lower as the competition becomes fiercer.

Workers always live in the city where they work, so they don't have too much freedom to choose where to buy houses.[6] And since the service industry (local businesses) thrives in cities that have many industrial firms and a large consumption market (which is brought by the large working population), commercial enterprises don't have too much freedom to choose their locations, either. Thus, local governments can extract consumer surplus in both the residential and commercial land markets, which may not only compensate for their loss in industrial land sales but also bring huge additional profits since the demand for residential and commercial land increases sharply if more industrial firms produce in their jurisdictions. Local governments also get huge fiscal revenue from the promotion of local businesses if more workers live in their jurisdictions.

[3] The Land Management Law passed in 1998 authorizes local governments to sell usufruct rights over the land. See Liu and Xiong (2020).

[4] The land price index in Liu and Xiong (2020) is based on the data and calculation of Chen et al. (2017).

[5] The industrial lands are sold through case-by-case negotiations and open auctions in China. In the first case, the special deal is easily achieved. In the second case, which is promoted by the central government to prevent corruption and fiscal revenue waste by selling lands at low prices, the special deals can be also realized by sending signals in the first stage of the auction to deter the entry of other bidders (Cai et al., 2013).

[6] Actually, only a part of workers like advanced engineers and managers buy houses in cities, most manufacturing workers coming from rural areas don't buy houses in cities due to low wage level and institutional discrimination (the hukou system). And that's why the high residential housing price doesn't raise the wage cost of industrial firms and influence their location choices.

Having introduced the background of fiscal competition in China, I build a model of the fiscal competition in the next section.

3 The Model

3.1 General settings

I use a Bertrand pricing game model to characterize the fiscal competition between Chinese city governments. In this model, each firm signs a contract of producing a given amount of output at given prices with the buyers on the global market.[7] With fixed output levels, each firm makes a short list of potential cities for building its factories, then the firm solves the production cost minimization problem and builds its factories in the city that is in its choice set and has the lowest production cost.[8]

The production cost of each firm is composed of four parts in the model:

(i) Capital cost: I assume that the price of capital is the same across the country.

(ii) Labor cost: I assume that each city has a wage level determined by the local labor market. An individual firm is just a price-taker in the labor market.

(iii) Land cost: The industrial land price is completely controlled by the city government, and the city government makes a special land price offer for each firm considering landing in its jurisdiction.

(iv) Other costs: all other costs (dependent on both firm and city characteristics) are included in the error term of the model, the distribution of which is common knowledge for all city governments.

The city governments maximize their expected fiscal revenue by attracting firms to land in their jurisdictions. The term "fiscal revenue" here is most broadly defined, which includes but is not limited to tax revenue, promotion of local businesses and housing market, political benefits, etc., i.e. all the potential benefits generated by landing firms.

The channel by which city governments attract firms is to provide special industrial land price offers to the firms. Since the output level of each firm is common knowledge for all city governments in the firm's choice set, the city governments can adjust the production cost of a firm by changing the land price offered to the firm. This will change the probability of the firm

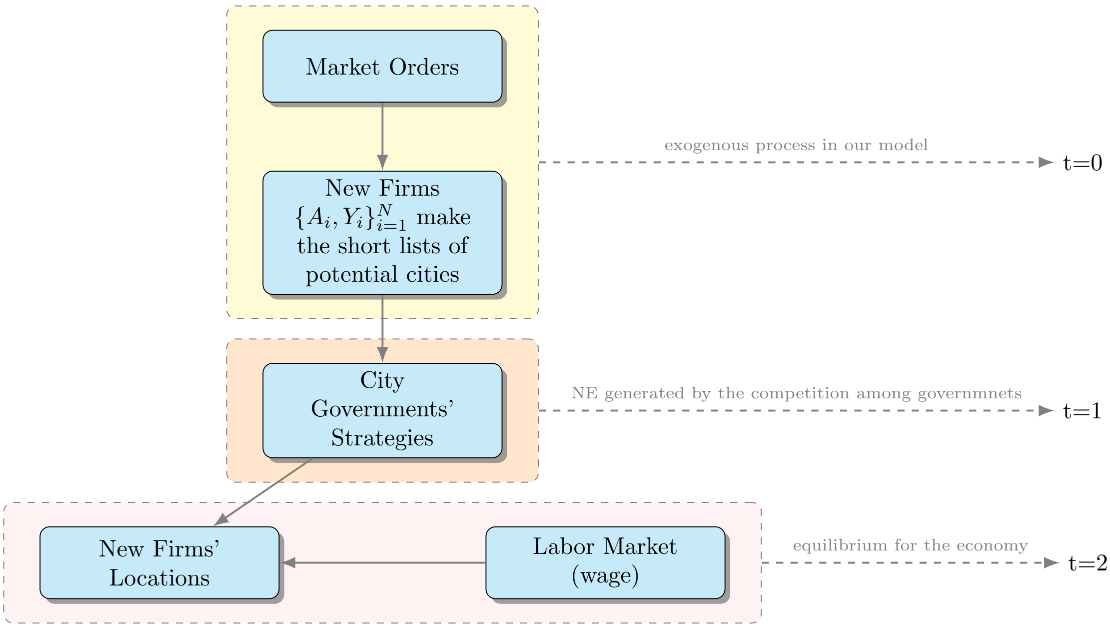
[7]The "firms" defined in this paper are private firms (Chinese private firms and foreign firms). State-owned firms are excluded from analysis since the goal of state-owned firms cannot be simply characterized as profit maximization, and they don't have absolute freedom to choose their locations.

[8]The assumption is valid since an individual firm has only limited information and political resources, so it can only deal with a limited number of city governments.

to land in this city as well as the expected fiscal revenue for the city government to attract the firm. And to determine the land price, a city government must consider the land prices offered by other city governments, since the other city governments' price choices will also impact the probability of the city getting the firm as well as the city's expected fiscal revenue. Thus, I use a Bertrand pricing game model to characterize the strategic interaction of city governments when they make the land offers. At the Nash equilibrium of this game, each city government has no incentive to change its land offer since the land price at the equilibrium has already maximized its expected fiscal revenue given other cities' land offers in the equilibrium strategy profile.

The main structure of the model implied in the discussion above can be described using Figure 1 below:

Figure 1: Main structure of the model



Notes: The figure presents the main structure and timing assumptions of the model. Solid arrows represent causal channels in the model. I use the productivity of firm i, A_i and the output level Y_i to characterize firm i in this figure.

Next, I describe the cost minimization and location choice problems faced by firms, then I write down the expected fiscal revenue maximization problem faced by city governments as well as the Bertrand pricing game between city governments. Finally, I define the pure strategy Nash equilibrium (NE) in this game and prove the uniqueness of pure NE.

3.2 Firms

I index new private firms by $i = 1, \ldots, N$; cities and city governments by $k = 1, \ldots, M$. Each firm i is characterized by its output level Y_i as well as its production function F_i, which is a mixture of Cobb-Douglas and Leontief technology:

$$F_i(K, L, T) = \min\left\{A_i K^{\alpha} L^{1-\alpha}, g_i(T)\right\}.$$

Here A_i is the productivity of firm i, K is the amount of capital the firm uses in production, L is the number of workers, T is the area of industrial the firm i uses, and $g_i(\cdot)$ is a strictly increasing function which reflects that land is a special input, which restricts the maximum output level of capital and labor. Notice that in this production function, capital and labor can substitute each other, and the land is the pure complement for capital and labor. I also make the no-waste assumption, i.e., firms make full use of the lands.

Given the output level Y_i firm i needs to produce, it chooses a city k in its choice set $\mathbf{C}_i$ (the short list of candidate cities) to build the factory, as well as the amount of capital input K_i per year, number of workers L_i and the area of industrial land T_i. Thus, firm i solves the following cost minimization problem:

$$\begin{aligned} &\min_{k \in \mathbf{C}_i, K_i, L_i, T_i} \left\{ rK_i + w_k L_i + \frac{p_{ik}}{m} T_i - \varepsilon_{ik} \right\} \\ &\text{s.t. } Y_i = \min\left\{A_i K_i^{\alpha} L_i^{1-\alpha}, g_i\left(T_i\right)\right\}, \end{aligned} \tag{1}$$

where r is the price of capital, which is assumed to be a constant in the whole country; w_k is the wage level in city k, which corresponds to the setting that firm i is just a price taker in the labor market; p_{ik} is the industrial land price city government k offers to firm i; m is the term of land lease.[9] I divide p_{ik} by m so that the time scales for all variables are unified to a year, and the firm is minimizing the production cost per year. The term ε_{ik} in the production cost is an error term, which is assumed to have type I extreme value distribution with scale parameter σ. It represents all other unobserved costs and benefits for the firm i to produce in the city k.

Furthermore, the price of one unit of product is normalized as 1 yuan (RMB) in the global market. Finally, I assume the size of the candidate set $|\mathbf{C}_i| = l(Y_i)$ is a nondecreasing function in Y_i since larger firms have more information and political resources to deal with a higher number of local governments.

Given the structure of the constrained minimization problem faced by firms, firm i always

[9] I drop the subscript of m since the terms of the land lease in my data set are all 50 years.

chooses T_i s.t. $Y_i = g_i(T)$. And the firm's cost minimization problem can be rewritten as:

$$\begin{aligned} &\min_{k \in \mathbf{C}_i, K_i, L_i} \left\{ rK_i + w_k L_i + \frac{p_{ik}}{m} g_i^{-1}(Y_i) - \varepsilon_{ik} \right\} \\ &\text{s.t. } Y_i = A_i K_i^{\alpha} L_i^{1-\alpha}. \end{aligned} \tag{2}$$

I use a two-step method to solve (2), first I solve the cost minimization problem for each city $k \in \mathbf{C}_i$; then I characterize the firm's location choice problem as finding the city in $\mathbf{C}_i$ with the lowest production cost for firm i :

(i) If firm i chooses to land in city k, then it should choose K_i and L_i to minimize its production cost. F.O.Cs of this problem show:

$$L_i = \left(\frac{1-\alpha}{\alpha}\right)^{\alpha} \left(\frac{r}{w_k}\right)^{\alpha} \frac{Y_i}{A_i} = B_i w_k^{-\alpha}, \tag{3}$$

and

$$K_i = \frac{\alpha}{1-\alpha} \cdot \frac{w_k}{r} \cdot L_i = \frac{\alpha}{r(1-\alpha)} B_i w_k^{1-\alpha}, \tag{4}$$

where $B_i := \left(\frac{(1-\alpha)r}{\alpha}\right)^{\alpha} \frac{Y_i}{A_i} = L_i w_k^{\alpha}$ only depends on the firm i's characteristics.

(ii) With a little abuse of notations, I denote the optimal level of land area for firm i by T_i in all the remaining parts of this paper, i.e. $T_i := g_i^{-1}(Y_i)$. Plug (3) and (4) into the cost function in (1) to get the total cost c_{ik} of producing in city k for firm i:

$$\begin{aligned} c_{ik} &= w_k \cdot L_i + r \cdot K_i + \frac{p_{ik}}{m} \cdot T_i - \varepsilon_{ik} \\ &= \frac{1}{1-\alpha} B_i w_k^{1-\alpha} + \frac{p_{ik}}{m} T_i - \varepsilon_{ik}. \end{aligned} \tag{5}$$

Now the firm just needs to choose the best location $k^* = argmin_{k \in \mathbf{C}_i} c_{ik}$. Notice in the model, the exact value of ε_{ik} is the private information of firm i, but the distribution of ε_{ik} is common knowledge among all city governments. Thus, city governments in firm i's choice set $\mathbf{C}_i$ can calculate the probability that firm i successfully lands in their jurisdictions though they don't know the actual location choice of the firm in advance. I denote $P_{ik}(p_{ik}, p_{i(-k)}) := Pr(i \text{ lands in } k | p_{ik}, p_{i(-k)})$ then:

$$\begin{aligned} P_{ik}(p_{ik}, p_{i(-k)}) &= Pr(c_{ik} < c_{ij} \ \forall j \in \mathbf{C}_i \setminus \{k\}) \\ &= \frac{exp\left[(-\frac{1}{1-\alpha} B_i w_k^{1-\alpha} - \frac{p_{ik}}{m} T_i)/\sigma\right]}{\sum_{j \in \mathbf{C}_i} exp\left[(-\frac{1}{1-\alpha} B_i w_j^{1-\alpha} - \frac{p_{ij}}{m} T_i)/\sigma\right]}, \end{aligned} \tag{6}$$

where (6) is the logit formula derived by McFadden (1974) and $p_{i(-k)}$ is the vector of land prices offered by all cities in $\mathbf{C}_i \setminus \{k\}$.

3.3 City governments

Now I turn to the government's expected revenue maximization problem. I assume the city government's revenue from attracting new firms to land in its jurisdiction is composed of two parts: (i) the fiscal revenue generated by the new firm itself, which includes tax revenue, promotion of local businesses and housing market, political benefits, etc, as discussed in Section 2, and (ii) the fiscal revenue generated by selling the industrial land. More specifically, I set:

$$\begin{aligned} v_{ik} &= P_{ik}(p_{ik}, p_{i(-k)}) \cdot \big(\beta Y_i + p_{ik} T_i\big) \\ &= \frac{exp\big[(-\frac{1}{1-\alpha} B_i w_k^{1-\alpha} - \frac{p_{ik}}{m} T_i)/\sigma\big]}{\sum_{j \in \mathbf{C}_i} exp\big[(-\frac{1}{1-\alpha} B_i w_j^{1-\alpha} - \frac{p_{ij}}{m} T_i)/\sigma\big]} \cdot \big(\beta Y_i + p_{ik} T_i\big), \end{aligned} \tag{7}$$

where v_{ik} is the expected fiscal revenue for city government k to attract firm i; the first term in the RHS of (7) is the probability of firm i to land in city k; the second term in the RHS of (7) is the city government's revenue from attracting new firms to land in its jurisdiction. More specifically, βY_i is the fiscal revenue generated by the new firm itself, where $\beta \geq 0$ can be interpreted as the revenue share of city governments in the output of the firm, and $p_{ik} T_i$ is the revenue from selling the land.[10]

Equation (7) shows the basic trade-off between the higher land-selling revenue and the higher probability of getting the firm. The higher land price can bring a city government higher land-selling revenue if it can get the firm, but the higher land price will also lower the probability of the firm choosing this city.

Notice that v_{ik} is not only dependent on city k's land price offer p_{ik}, but also dependent on other competitor cities' land price offers $p_{i(-k)}$, which implies each city government should consider other city governments' land pricing strategy when it tries to solve its own expected fiscal revenue maximization problem. And this observation leads us to model the interaction between city governments in the game-theoretic framework.

3.4 Bertrand pricing game

As mentioned at the end of Section 3.3, I model the interaction between city governments by a Bertrand pricing game model. More explicitly, for each new firm i, the cities in its choice set $\mathbf{C}_i$

[10] Since the "fiscal revenue" here is most broadly defined as all the revenue the firm brings to the city government directly and indirectly in the whole term of the current officials, $\beta > 1$ is possible. In other words, β can also be interpreted as a multiplier of the output of firms, and the fiscal revenue from attracting a firm to land in a city is just the multiplier times the output level of the firm. I assume the fiscal revenue is positively proportional to the scale of the firm measured by its output level per year.

play a Bertrand pricing game to maximizing their expected revenue by affecting the probabilities of the firm choosing each of these cities. At the Nash Equilibrium (NE) of the game, each city has no incentive to change its land price offer given other cities' prices. Now I formally define the game and the NE:

The Bertrand pricing game for attracting firm i is denoted by $\langle \mathbf{C}_i, (S_{ik}), (v_{ik})\rangle$, where $\mathbf{C}_i$ is the set of players, i.e. the city governments in firm i's choice set; S_{ik} is the action space for each player, where I stipulate that $S_{ik} \equiv S_i = [p^i_{min},\ p^i_{max}]$ for $\forall i, k$;[11] and the payoff function v_{ik} is just the expected fiscal revenue function I defined in (7). I also denote $N_i \equiv |\mathbf{C}_i|$ as the number of players in this game. The pure strategy Nash equilibrium of the Bertrand game is defined as below:[12]

Definition 1 (**Nash Equilibrium**). For all i, a strategy profile (price vector) $p^*_i \in S_i^{N_i}$ is pure strategy Nash Equilibrium (NE) if for all $k \in \mathbf{C}_i$ and $p_{ik} \in S_i$,

$$v_{ik}(p_{ik},\ p^*_{i(-k)}) \leq v_{ik}(p^*_{ik},\ p^*_{i(-k)}).$$

Now I consider the properties of NEs in the Bertrand Game. First, I analyze the best response functions in this game. Next, I prove the existence and uniqueness of pure NE.

Without loss of generality, I consider a city government k in $\mathbf{C}_i$ trying to attract firm i, and analyze the city government k's best response given other competitor cities' land prices.

First I derive the partial derivative of v_{ik} w.r.t. p_{ik}:

$$\frac{\partial v_{ik}}{\partial p_{ik}} = T_i P_{ik}(p_{ik}, p_{i(-k)}) \cdot \big[1 - \frac{1}{\sigma m}\big(1 - P_{ik}(p_{ik}, p_{i(-k)})\big)\big(\beta Y_i + p_{ik} T_i\big)\big]. \tag{8}$$

Notice that the first term in (8): $T_i P_{ik}(p_{ik}, p_{i(-k)}) > 0$, I fix $p_{i(-k)}$ and define:

$$f(p_{ik}) := 1 - \frac{1}{\sigma m}\big(1 - P_{ik}(p_{ik}, p_{i(-k)})\big)\big(\beta Y_i + p_{ik} T_i\big).$$

The derivative of $f(p_{ik})$ is:

$$f'(p_{ik}) = \underbrace{-\frac{1}{\sigma m}}_{<0} \cdot [\underbrace{-\frac{\partial P_{ik}(p_{ik}, p_{i(-k)})}{\partial p_{ik}}}_{>0} \underbrace{(\beta Y_i + p_{ik} T_i)}_{>0} + \underbrace{(1 - P_i(k))\, T_i}_{>0}] < 0. \tag{9}$$

[11] I use p^i_{min} to reflect the participation constraint of city governments in $\mathbf{C}_i$ (governments can't set the land price smaller than the opportunity cost of transferring the land), p^i_{max} to reflect the participation constraint of firm i (firms will not build factories if the land price is extremely high).

[12] I don't consider mixed strategy Nash equilibrium in this paper.

And if the domain of f is extended to $\mathbb{R}$, it can be shown:[13]

$$\begin{aligned} &\lim_{p_{ik}\to-\infty} f(p_{ik}) = 1, \\ &\lim_{p_{ik}\to+\infty} f(p_{ik}) = -\infty. \end{aligned} \tag{10}$$

Combining (8), (9) and (10), it is clear that for any $p_{i(-k)}$, $\exists\zeta \in \mathbb{R}$ s.t. $\frac{\partial v_{ik}}{\partial p_{ik}} > 0$ for all $p_{ik} < \zeta$, $\frac{\partial v_{ik}}{\partial p_{ik}} > 0$ for all $p_{ik} > \zeta$, and $\frac{\partial v_{ik}}{\partial p_{ik}}\Big|_{p_{ik}=\zeta} < 0$. Based on this observation, I have the following theorem:

Theorem 1 (**Uniqueness of Best Response**). *For any Bertrand Game $\langle \mathbf{C}_i, (S_{ik}), (v_{ik})\rangle$ and any city government k in $\mathbf{C}_i$, given other city governments' land price profile $p_{i(-k)}$, k's best response p^*_{ik} is uniquely given by:*

$$p^*_{ik} = \begin{cases} p^i_{min}, & \text{if } f(p^i_{min}) \le 0 \\ p^i_{max}, & \text{if } f(p^i_{max}) \ge 0 \\ \text{the root of } f(p_{ik}) = 0, & \text{if } f(p^i_{min}) > 0 \text{ and } f(p^i_{max}) < 0 \end{cases} \tag{11}$$

The proof of Theorem 1 is in Appendix A.

To prove the existence and uniqueness of pure NE, I first define a mapping G. For any Bertrand Game $\langle \mathbf{C}_i, (S_{ik}), (v_{ik})\rangle$, I define $G : S^{N_i} \to S^{N_i}$ s.t. $G(p) = p'$:[14] for $k = 1$, p'_k is the best response of city k given other cities' land price profile $(p_2, \dots, p_N)$; for $2 \le k \le N$, p'_k is the best response of city k given other cities' land price profile $(p'_1, \dots, p'_{k-1}, p_{k+1}, \dots, p_N)$; for $k = N$, p'_k is the best response of city k given other cities' land price profile $(p'_1, \dots, p'_{N-1})$. I have the following theorem:

Theorem 2 (**Uniqueness of Pure NE**). *Every Bertrand Game $\langle \mathbf{C}_i, (S_{ik}), (v_{ik})\rangle$ has a unique pure strategy Nash Equilibrium p^*. Moreover $G^n(p) \to p^*$ for any $p \in S_i^{N_i}$ as $n \to \infty$, where G^n refers to n-th composition of G with itself.*

The proof of Theorem 2 is in Appendix A.

3.5 Numerical method of solving the model

Now I consider the numerical method used to solve the Bertrand Game.

[13] This extension of the domain is just for the convenience of analyzing the property of f. The action space S_i is still a closed interval.

[14] The subscript i is dropped here for brevity.

Theorem 1 and Theorem 2 immediately provide a way to numerically solve the game based on the Gauss-Seidel algorithm (Algorithm 1): I can start with an initial strategy profile (price vector), and for each city, I calculate the city government's best response given others' land prices, then update the city's strategy. I repeat this process until the price vector converges, i.e. NE is found. The algorithm is described on the next page.

Notice that the inner loop of Algorithm 1 is the mapping G defined in Section 3.4. Thus, the outer loop is to compute the fixed point of the contraction mapping G using the successive approximation algorithm implemented with Gauss-Seidel iterations.

Algorithm 1: Gauss-Seidel algorithm to solve the Bertrand Game $\langle \mathbf{C}_i, (S_{ik}), (v_{ik}) \rangle$

Input: any initial price vector p_i;
action space $[p^i_{min}, p^i_{max}]$;
$tol \leftarrow 1e - 10$; `/* Convergence criteria */`
Result: a price vector sufficiently close to NE p^*
repeat
  $p_i^0 = p_i$
  for k *in* $1 : N_i$ **do**
    if $f(p_{ik} = p^i_{min}) \leq 0$ **then**
      $p_{ik} \leftarrow p^i_{min}$
    end
    else if $f(p_{ik} = p^i_{max}) \geq 0$ **then**
      $p_{ik} \leftarrow p^i_{max}$
    end
    else
      $p_{ik} \leftarrow$ the root of $f(p_{ik}) = 0$
    end
  end
until $max_k(abs(p^0_{ik} - p_{ik})) < tol$;
return p_i

4 Data

In this section, I briefly introduce the data used for estimation. I use the firm-land-city data in the year 2012 since my model is static, and the number of observations in 2012 is the largest. Another reason I choose 2012 data is that after 2013, the autonomy of local governments decreases due to the change in the political environment. Thus, data after 2013 may not be fully suitable for my model.

4.1 Data sources

First, I use the 2013 Chinese industrial firm survey to get the output level,[15] the number of workers, the area of industrial land usage of the industrial firms newly established in 2012, and I also know the cities where these new firms choose to locate in. It is worth noticing that the firms recorded in the Chinese industrial firm survey are "enterprises above designated size", which means the firms with output levels greater than or equal to five million yuan per year. Thus, my estimation is restricted to these large industrial firms in China, which are the focus of the model.

Second, I use the land selling data from *www.landchina.com* to get the industrial land selling data in 2012. The website is operated by the Ministry of Natural Resources of China, which is a reliable data source. And the land-selling data includes the areas, prices, and firms which buy the land.[16]

Third, I use the China city statistical yearbook to get the average wage of each city.

Finally, I match these data and get the final data set, which includes 1019 observations. An observation is a tuple of a firm's output level, labor usage, area of industrial land bought by the firm, the city where the firm chooses, and the average wage in this city.

4.2 Data descriptions

I visualize the spatial distribution of the observations in my data set in Figure 2.

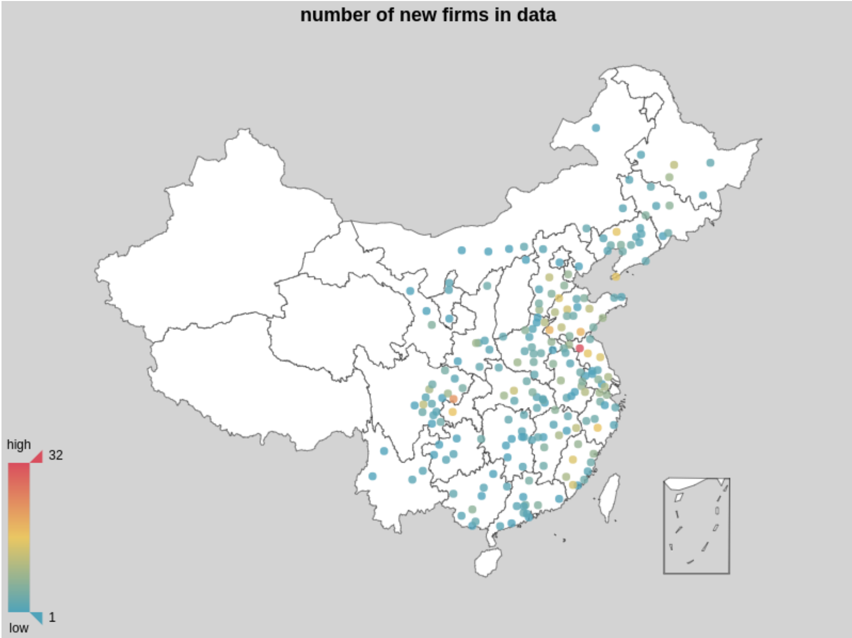
I draw the map of China in Figure 2. Each point in the map represents a city that at least one firm chooses. The color of the point represents the number of firms which the corresponding city lands. There are 212 such cities (points in the map) in my data set, and there are 4 direct-administered municipalities (Beijing, Shanghai, Tianjin, Chongqing) and 293 prefecture-level cities in China, thus my data set covers roughly two-thirds of Chinese major cities,[17] though several cities have far more observations than others.

[15] I use the operation revenue of a firm as its output level since the price of the product is normalized as 1 yuan in my model. And I delete some problematic observations in the survey data, i.e. firms with zero operation revenue and firms with less than ten workers.

[16] I delete the outliers in the land selling data, i.e. the lands with top 1% land price and bottom 1% land price (22 observations are deleted).

[17] There is no observation in several regions of China like Xinjiang, Tibet, Qinghai, and Hainan. However, there are few large industrial firms in these regions due to their geographic characteristics, thus, my data set still covers enough cities.

Figure 2: Number of firms each city lands in data



Notes: Each point in the map represents a city which lands at least one firm in my data set, and there are 213 such cities in my data set. The color of the point denotes the number of firms a city lands, and the number increases as the color changes from blue to red.

I also show the descriptive statistics of the variables used for estimation in Table 1:

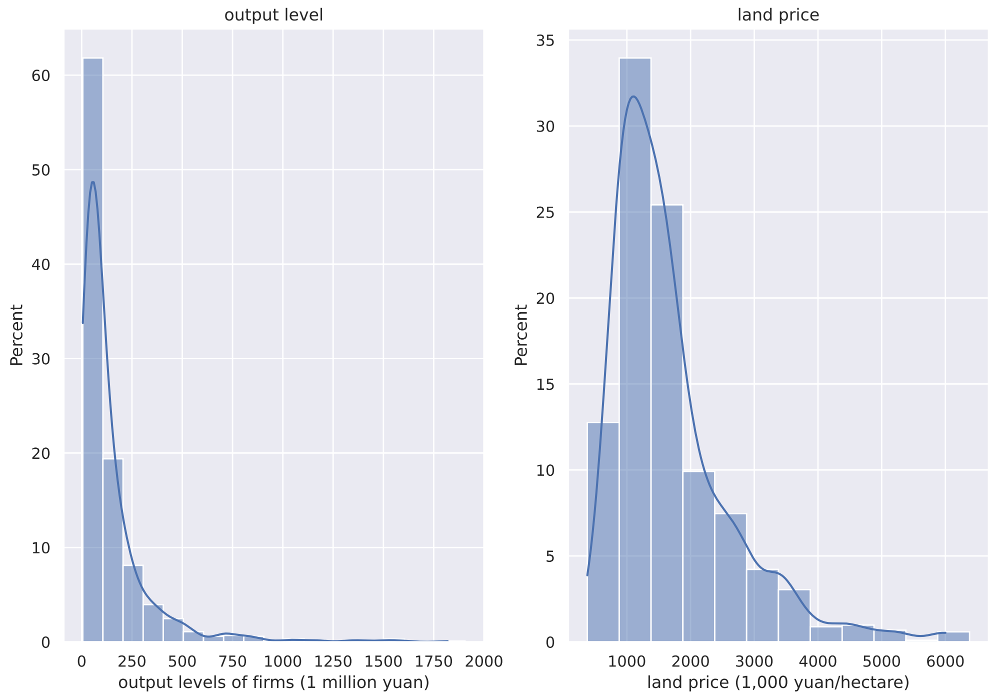
Table 1: Descriptive statistics of variables in model

	land price (1,000 yuan/hectare)	output (1,000 yuan/year)	number of workers (person)	area of land (hectare)	city wage (1,000 yuan/year)
mean	1685.10	170733.30	228.16	4.59	36.96
std	934.95	409499.05	143.57	5.65	18.69
min	379.50	5662.00	10.00	0.03	17.21
25%	978.62	38354.00	125.00	1.60	30.49
50%	1440.01	72909.00	223.00	2.83	34.73
75%	2015.52	164553.00	324.50	5.40	39.83
max	6001.56	7282607.00	2326.00	65.48	320.63
N	1019	1019	1019	1019	288

Notes: There are 1019 observations (new firms established in 2012) in our data set. The last column is the city average wage for 288 cities, which is very close to the total number of prefecture-level cities and direct-administered municipalities (297 cities) in China. And 25%, 50%, 74% in the table mean 25% percentile, median, 75% percentile respectively.

Figure 3 shows the distribution of land price and firms' output level, which are the two most important variables in my model.

Figure 3: Sample distribution of output level and land price



Notes: The blue line in the graphs are kernel density plots. To make the first histogram clearer, I just draw the distribution of firms with output levels smaller than 2 billion yuan/year. There are several firms with output levels of more than 2 billion yuan/year, which makes the actual tail of the distribution of output fatter.

According to Figure 3, both the distributions of the output level of firms and the industrial land price are left-skewed, moreover, more than 80% firms have output levels between 5 million and 250 million yuan per year, and most of the firms buy land with a price below 3 million yuan per hectare. This implies most of the industrial land is sold at relatively low prices in my data set, which is consistent with my observations in Section 1 and Section 2.

5 Estimation Method

5.1 Settings for calibration and estimation

I calibrate the parameter of capital income share α in the model, then use the method of simulated moments (MSM) (Pakes and Pollard, 1989; McFadden, 1989; Gouriéroux and Monfort, 1996) to estimate the parameter of governments' revenue share β and the scale parameter σ of the distribution of the unobserved term in production cost.

More specifically, I choose the capital income share $\alpha = 0.5$ according to Zhu (2012) and Brandt et al. (2008). I also do estimations for $\alpha = 0.33$ and 0.67, which generate similar estimates of β and σ. This may indicate a potential identification problem for α.

I bound the parameters space as $\beta \in [0, 2]$ and $\sigma \in (0, 1000]$ and perform the initial grid search for parameters. The parameters space is large enough since two times of yearly output of firms is huge as the city governments' revenue share, and σ should not be too large to make the location choice irrelevant of the city governments' land-selling strategies.

I also specify the action space of cities for the game of attracting firm i as $[p^i_{min},\ p^i_{max}]$, where $p^i_{min} = \max\{0,\ p^i_{obs} - 2000\}$, $p^i_{max} = p^i_{obs} + 2000$, p^i_{obs} is the observed land price offered to firm i in data. p^i_{min} is used to roughly characterize the participation constraints of cities since the cities will not sell the land to the firm if the cost of land cannot be covered by the land-selling revenue and fiscal revenue. And p^i_{max} is used to characterize the participation constraint of the firm since if the land price (a part of production cost) is too high, the firm will suffer loss. The action space is neither too large nor too small (2000 is roughly one-third of the range of observed land prices). See Appendix B.2 for robustness checks of other settings of the action space.

To estimate β and σ, I need to solve the Bertrand pricing games for each firm under different values of parameters. All the parameters and variables (inputs) needed for solving the Bertrand Game $\langle \mathbf{C}_i, (S_{ik}), (v_{ik}) \rangle$ are listed in Table 2:

Table 2: Inputs in the Bertrand Game model

	Name	Interpretation	Source	Value/Range/Formula
Parameters	β	governments' revenue share	estimation	$[0, 2]$
	σ	scale parameter in the distribution of ε_{ik}	estimation	$(0, 1000]$
	α	capital income share	calibration	$\{0.33, 0.5, 0.67\}$
Variables	Y_i	output level of firm	data	
	L_i	number of workers in the firm	data	
	T_i	area of land usage	data	
	w_k	city average wage	data	
	B_i	firm's characteristic	data	$L_i w^{\alpha}_{k^*}$
	$\mathbf{C}_i$	firm's choice set	random draws	
	p^i_{obs}	observed land price for the firm	data	
	p^i_{min}	the minimum of possible land price	data	$\max\{0,\ p^i_{obs} - 2000\}$
	p^i_{max}	the maximum of possible land price	data	$p^i_{obs} + 2000$

Notes: In this table, w_k is the abbreviation of $\{w_k\}_{k \in \mathbf{C}_i}$. k^* is the city the firm actually lands in (and I observe) in data. The specification of $\mathbf{C}_i$ is written in the text. $[p^i_{min},\ p^i_{max}]$ is the action space for the cities in the game of attracting firm i, which is also described in the text.

If I have all the inputs defined in Table 2, the equilibrium price vector p_i can be solved by using Algorithm 1 for all observations (all firm i) in data. Thus, I need to specify $\mathbf{C}_i$ (the choice set of the firm) in Table 2 explicitly to obtain the parametric model.

I simulate $\mathbf{C}_i$ by repeatedly drawing $|\mathbf{C}_i|$ cities randomly from ten cities that are nearest to the city where the firm lands in data (the observed chosen city).[18]

To fully specify the generating process of the choice set, the size of the set $|\mathbf{C}_i|$ has to be determined. In Section 3 I set $|\mathbf{C}_i| = l(Y_i)$, and $l(Y_i)$ is defined explicitly here:

$$l(Y_i) = \begin{cases} 3 & \text{if } Y_i \leq 33rd \text{ percentile of } Y \\ 4 & \text{if } 33rd \text{ percentile of } Y < Y_i \leq 67th \text{ percentile of } Y \\ 5 & \text{if } Y_i > 67th \text{ percentile of } Y \end{cases} \tag{12}$$

In other words, there are 33% firms that have 3, 4, and 5 candidate cities respectively. And what I assume here is the number of candidate cities of a firm is increasing in the scale (output level) of a firm. This is consistent with the observation that large firms always have more political resources and are more valuable in the eyes of city governments, and are often attracted by more city governments. I use five as an upper bound for the number of candidate cities based on the observation that the political resources of a firm to deal with city governments are limited.[19] For robustness checks, I also do estimation under the two settings that (i) one-half of the firms have 3 candidate cities, the other half of the firms have 4 candidate cities. (ii) one-half of the firms have 4 candidate cities, and the other half of the firms have 5 candidate cities respectively, and the results are reported in Appendix B.1.

5.2 Moment conditions

Next, I construct the MSM estimator for the parameter β and σ. I define all the information of firm i as $Q_i \equiv \{Y_i, T_i, B_i\}$. All the parameters are denoted by $\theta = (\alpha, \beta, \sigma)$, and the true parameters are denoted by θ_0. From the discussion in Section 3 it is known that if the distribution of the choice set $\mathbf{C}_i$ is given, the distribution of land price at NE for firm i will also be determined, thus the expectation of the land price will be determined. I define $x_i \equiv [p_i, p_i^2]'$,[20] and denote the conditional first and second-order moments of land price by $h(x_i;\theta_0) \equiv E(x_i|Q_i;\theta_0)$. By the

[18] By "nearest" I mean the ten cities which have the shortest straight-line distances (between city centers) to the chosen city. The chosen city is included in the ten cities.

[19] It's difficult if not impossible for researchers to know the exact number of cities attracting each firm, thus, I choose the numbers in (12) based on the media coverage of location choices of some large firms. For example, four cities Zhengzhou, Kaifeng, Nanyang, Hebi in Henan province tried to attract Foxconn, a Taiwanese multinational electronics contract manufacturer to build its factory in 2010 (`http://www.chinanews.com.cn/cj/2010/07-22/2419622.shtml`). Five cities (Xi'an, Beijing, Chongqing, Wuxi, Suzhou) were reported to attract Samsung, a Korean multinational manufacturing conglomerate to build its factory in 2012 (`http://jingji.cntv.cn/20120417/109763.shtml`).

[20] p_i is scaled by 10,000 for estimation since $p_{max} = 6001.56$ in data.

definition of conditional land price, I have:

$$E[x_i - h(Q_i;\theta_0)|Q_i] = 0 \tag{13}$$

Since the structural residual $x_i - h(Q_i;\theta_0)$ is orthogonal to any function of Q_i, I choose the instrument variables vector $Z_i = [1, Y_i, T_i, Y_i \times T_i]'$ for simplicity.[21] The moment conditions are:

$$E[Z_i \otimes \big(x_i - h(Q_i;\theta_0)\big)] = 0 \tag{14}$$

where the expectation of Kronecker product is zero vector, i.e., each IV is orthogonal to the two-by-one structural residuals vector.

Since calculating the theoretical moment $h(Q_i;\theta)$ is computationally expensive,[22] I use the simulated moment $\frac{1}{S}\sum\limits_{s=1}^{S}\tilde{h}(Q_i,\mathbf{C}_i^s;\theta)$ to replace $h(Q_i;\theta_0)$, where $\tilde{h}(Q_i,\mathbf{C}_i^s;\theta) \equiv E(x_i|\mathbf{C}_i^s, Q_i;\theta_0)$, the number of simulations $S = 10$, and $\mathbf{C}_i^s$ is randomly drawn as described in Section 5.1. The MSM estimator is defined by:

$$\hat{\theta} = \underset{\theta}{argmin}\Big(\sum_{i=1}^{N} Z_i \otimes \big(x_i - \frac{1}{S}\sum_{s=1}^{S}\tilde{h}(Q_i,\mathbf{C}_i^s;\theta)\big)\Big)' W \Big(\sum_{i=1}^{N} Z_i \otimes \big(x_i - \frac{1}{S}\sum_{s=1}^{S}\tilde{h}(Q_i,\mathbf{C}_i^s;\theta)\big)\Big) \tag{15}$$

$$s.t. \quad \theta = (\alpha_0, \beta, \sigma),\ 0 \leq \beta \leq 10,\ 0 < \sigma \leq 1000$$

where α_0 is the calibrated α, and W is the weighting matrix described in Section 5.3. As the estimation results in Section 6 show, the constraints of β and σ are not binding at the solution.

5.3 Numerical algorithm for estimation

In this section, I discuss the numerical algorithm used to estimate β and σ. Showing that a global minimum of the MSM objective is found in nonlinear models is generally very difficult or impossible (Iskhakov and Keane, 2021), but I try to reach the global minimum as close as possible by combining the grid search and Quasi-Newton algorithm.

I do a two-stage MSM estimation. As mentioned in the last section, I calibrate α to 0.33, 0.5, and 0.67 respectively. For each α, in the first stage of estimation, I set the weighting matrix to the identity matrix. Next I calculate the MSM objective for all the (β, σ) pairs on the grids $\{0, 0.1, \ldots, 2.0\} \times \{100, 200, \ldots, 1000\}$ and find the $(\tilde{\beta}, \tilde{\sigma})$ with the smallest MSM objective, which should be near the global minimizer. Then I use Quasi-Newton (L-BFGS) algorithm to

[21] Y_i and T_i are also scaled to the interval $[0, 1]$ for estimation.

[22] For example, if each firm only has three candidate cities, then to calculate the conditional moments for one firm at any given parameters θ, I need to solve $\binom{10}{3} = 120$ games.

do a finer search of the two parameters starting from $(\tilde{\beta}, \tilde{\sigma})$ to get the first stage estimates $(\hat{\beta}_0, \hat{\sigma}_0)$.

To do the second-stage estimation, I update the weighting matrix W as the sample analog of $\{\mathrm{Var}[Z_i \otimes (x_i - h(Q_i; \hat{\theta}_0))] + \frac{1}{S}\mathrm{Var}[Z_i \otimes (\tilde{h}(Q_i, \mathbf{C}_i^s; \hat{\theta}_0) - h(Q_i; \hat{\theta}_0))]\}^{-1}$, where $\hat{\theta}_0 = (\alpha_0, \hat{\beta}_0, \hat{\sigma}_0)$, and I use the average of 30 simulations to approximate the theoretical moments $h(Q_i; \hat{\theta}_0)$ for each observation i. I use Quasi-Newton (L-BFGS) algorithm to search the second-stage estimates $(\hat{\beta}, \hat{\sigma})$, which minimizes the new criteria function with the updated W, staring from the first stage estimates $(\hat{\beta}_0, \hat{\sigma}_0)$.

I treat W as the optimal weighting matrix, and the estimated asymptotic covariance matrix of the second-stage estimates are $\frac{1}{N}(\hat{D}'W\hat{D})^{-1}$, where $\hat{D}$ is the sample analog of the gradient D of moment conditions evaluated at $(\hat{\beta}, \hat{\sigma})$.

For the whole estimation procedure, I parallelize the calculation of the MSM objective by splitting the tasks of solving all the games into several chunks. And I let each CPU to solve a chunk of the games simultaneously to save the time of estimation.

6 Estimation Results

6.1 Structural estimates

The estimation results are given in Table 3.

Table 3: Estimates of Parameters ($3 \leq |\mathbf{C}_i| \leq 5$)

Calibrated α	Parameters	Estimates	Standard Error
0.33	β	0.512	0.043
	σ	172.246	24.058
0.5	β	0.453	0.032
	σ	178.301	25.102
0.67	β	0.462	0.034
	σ	182.761	24.920

The estimates of governments' revenue share β and the scale parameter of the error term σ are stable. $\hat{\beta}$ varies from 0.45 to 0.51, $\hat{\sigma}$ varies from 172 to 182 for all three different capital income shares. Appendix B.1 presents the estimation results under the two settings of $3 \leq |\mathbf{C}_i| \leq 4$ and $4 \leq |\mathbf{C}_i| \leq 5$, and $\hat{\beta}$ varies from 0.53 to 0.58 for the first case, $\hat{\beta} \approx 0.4$ for the second case. Thus, even if each firm has 4 or 5 candidate cities, i.e., the fiscal competition is very fierce, the estimates of β don't vary far from 0.45. And $\hat{\beta} = 0.45$ is a conservative estimate for the governments' revenue share since if I decrease the sizes of the choice sets, $\hat{\beta}$ will be higher than 0.5. Appendix B.2 presents the estimation results under the different settings of action spaces of city governments, and it also shows $\hat{\beta} = 0.45$ is a conservative estimate.

To conclude, $\alpha = 0.5$ is a valid calibrated value, and $\hat{\beta} = 0.45$, $\hat{\sigma} = 178$ are my preferred estimates. I will use these estimates in the remaining analysis of the paper.

6.2 Discussion of the estimation results

As Table 3 shows, the city governments' revenue share of firms' output is quite high ($\hat{\beta} \approx 0.45$), which is consistent with my assumption that attracting industrial firms to land in a city's jurisdiction will generate huge potential benefits for the city government by the fiscal externality or spill-over effects.

To clearly illustrate this point, a thought experiment can be considered: I assume the only benefits city governments can get from landing new firms is the value-added tax (VAT) revenue and the city government officials think they can get the benefits in five years after the new firm landed.[23]

I also assume the labor income is 70% of firms' output, which might be an overestimation since $\alpha = 0.3$ in this case. Since the average profit ratio of industrial firms in 2012 is 6%,[24] I can assume the taxable added value is 80% (greater than 70% + 6%) of the firms' output, which is also an overestimated figure.

The VAT rate is 17%, the share of city government in VAT revenue is 25%, and the back-of-envelop calculation shows that the city government's revenue share is $5 \times 80\% \times 17\% \times 25\% = 0.17$, which is 38% of $\hat{\beta} \approx 0.45$.[25] Thus, attracting new firms brings city governments a huge amount of benefits such as promoting local business, boosting the housing market, etc. besides the official tax revenue. The fiscal externality of attracting industrial firms is huge.

6.3 Fit of the model

To check whether my model describes the fiscal competition between city governments properly, I discuss the fit of the model in this section.

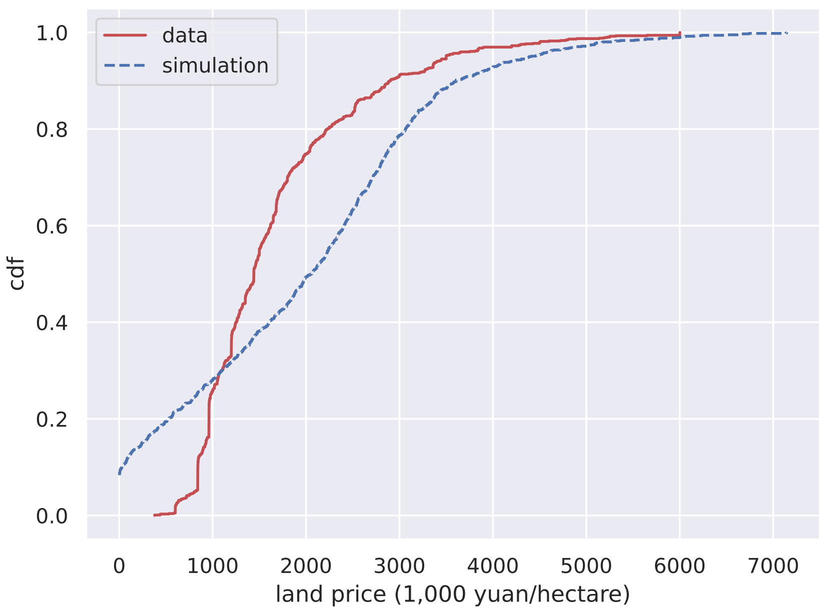
[23] Five years is an overestimated term since the average term for Chinese city communist party secretary and mayor are 3.6 years and 3.2 years respectively during 2000-2010, see `https://www.yicai.com/news/3106338.html`.

[24] This profit ratio (6%) is calculated by the National Bureau of Statistics of China, see `http://www.stats.gov.cn/tjsj/zxfb/201301/t20130127_12932.html`.

[25] According to a survey conducted by Chinese Academy of Fiscal Sciences, the ratio of firm tax to added value is around 6.6% in 2013, see `http://www.cf40.com/news_detail/7374.html`, which is much lower than $80\% \times 17\% = 13.6\%$, thus, my comparison is very conservative.

I compare the distributions of observed land prices and simulated land prices in Figure 4 and Table 4.

Figure 4: Comparison of empirical CDFs of observed land prices and simulated land prices



Notes: I use the preferred estimates to do 30 simulations, and use the average of the expected land price in these simulations as the simulated land price.

Figure 4 shows the empirical CDFs of both observed land prices and simulated land prices. And the two empirical CDFs are close though there are more high prices in simulation.

Table 4: Descriptive Statistics of Observed Prices and Simulated Prices

	observed price	simulated price
mean	1685.097	1988.439
std	934.952	1396.583
min	379.505	0.000
25% percentile	978.617	832.676
50% percentile	1440.007	2058.919
75% percentile	2015.520	2863.554
max	6001.559	7145.852

Notes: $N = 1019$. I use the preferred estimates to do 30 simulations, and use the average of the expected land price in these simulations as the simulated land price.

Table 4 shows the descriptive statistics of the observed prices and simulated prices. The mean and 25% percentile of simulated prices are matched well to observed prices. But the standard deviation, median, and 75% percentile of simulated prices are not matched well to observed prices. This is because my model is parsimonious, and may not explain the second-order moment of land price very well. Another reason for the relatively poor matches of the median and 75% percentile of land prices is that I don't use the higher-order moments in the estimation, since higher-order moments are difficult to be precisely computed within a finite number of observations (Adda and Cooper, 2003). However, since the matches of the mean and 25% percentile are good, I proceed to use the preferred estimates for the counterfactual analysis.

7 Counterfactual Analysis

In this section, I discuss the implications of the model by constructing three counterfactuals. First, I examine whether the fiscal competition characterized by the model can improve allocation efficiency by changing the locations of firms. Second, I analyze the potential impacts of suppressing the fiscal competition between city governments on industrial land prices and fiscal revenues. Third, I discuss the potential impacts of rising urban wages on the fiscal competition.

7.1 Allocation efficiency

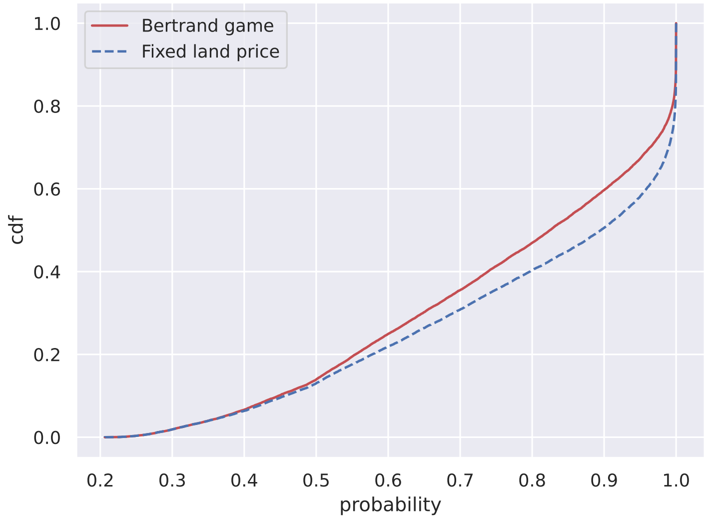
As I discussed in Section 1 and Section 2, the potential benefits of fiscal competition between cities is the improvement of resource allocation between cities. More specifically, cities that are disadvantaged in attracting firms (wage levels are high in these cities) can increase the probability of landing the firms by decreasing the land price more sharply than their competitors. But the extent of these benefits is unclear, since the sharp decreases of prices by disadvantaged cities may also promote other cities to decrease their prices, and probabilities of getting the firm across cities may not change by a lot.

To measure the exact impact of fiscal competition, I consider the counterfactual that the central government freezes the industrial land market by fixing the industrial land price across the whole country. If the land price is fixed, the city with the lowest wage level among the candidates of a firm will have the highest probability of landing the firm. And I denote these cities as advantaged cities.

I use the change of the probabilities of the advantaged cities landing the firms to measure the impact of fiscal competition on allocation efficiency. If fiscal competition decreases these probabilities significantly, the disadvantaged cities will have higher chances to get the firms, i.e., fiscal competition may be helpful to improve allocation efficiency.

I also do thirty simulations for both fiscal competition and fixed land price cases, and the empirical distribution of the probability of getting the firm in the fixed land price case is shown with the baseline case together in the Figure 5.

Figure 5: The empirical CDF of the probability of getting the firm for advantaged cities



Notes: I use the preferred estimates to do thirty simulations, and calculate all the choice probabilities in the $N \times S$ (1019×30) location choice problems to draw the empirical distribution.

I compare the empirical CDF of the highest probability of getting the firm in the Bertrand game with which in the case of nationwide fixed land price (see Figure 5), and they are very similar, though the empirical CDF curve of the Bertrand game is slightly higher than which of the fixed land price case, i.e. the highest probability in the Bertrand game is stochastically dominated by that in the fixed land price case. This implies that the impact of fiscal competition on the locations of firms (allocation efficiency) is very small though fiscal competition indeed gives some disadvantaged cities higher chances of landing the firms.

I also show the descriptive statistics of the probability of getting the firms for advantaged cities in Table 5 below. It shows that fixing the land price (banning the fiscal competition) will increase this probability by 2.9% on average and 7% for the median, which verifies again that the impact of fiscal competition on allocation efficiency is small.

Table 5: Empirical distribution of probability of getting the firms for advantaged cities

	fixed land price	Bertrand game	difference
mean	0.801	0.773	0.029
std	0.218	0.215	0.002
min	0.206	0.206	0.000
25% percentile	0.635	0.601	0.035
median	0.895	0.825	0.071
75% percentile	0.996	0.984	0.012
max	1.000	1.000	0.000

7.2 Potential impacts of fiscal centralization

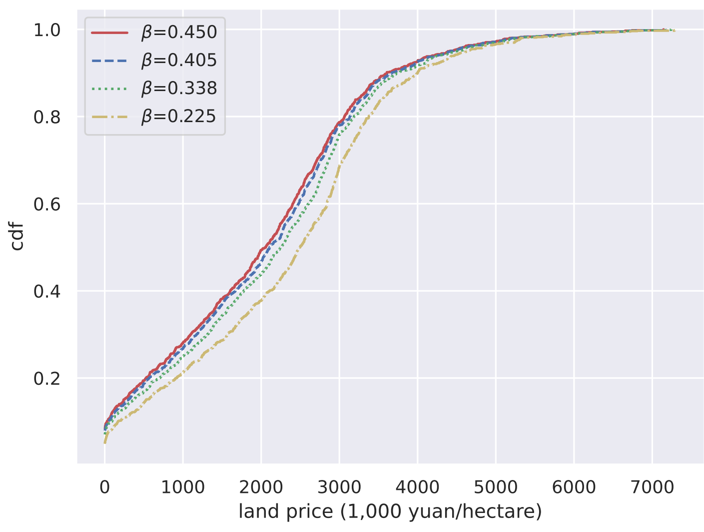
As I discussed in Section 6.2, the estimates of $\hat{\beta} = 0.45$ is much greater than the official tax revenue share. Such large city governments' revenue share of firms' output reflects not only the strong incentive of local officials to attract firms, but also the significant economic and fiscal power of local officials, which can convert the spillover effects of industrial firms on local businesses, the housing market, etc., to their fiscal revenue. Moreover, as Kroeber (2020) observed: *"The ability of these leaders (local officials) to act independently of central dictates, and in response to local needs, has contributed to China's resilience and dynamism."*. However, the fierce fiscal competition between local governments due to this ability may also waste a lot of potential fiscal revenue in the industrial land market as I discussed before.

The autonomy of local officials in economic affairs declines significantly after 2013 due to the re-centralization of the governance structure in China and the anti-corruption movement, and I call this change "fiscal centralization" reflected by the decrease of β.

I examine the potential impacts of the decrease of β by counterfactual experiments. Specifically, I decrease the estimated $\hat{\beta}$ by 10%, 25%, 50% respectively, and do thirty simulations to observe the change in the average land price, total land selling revenue and total fiscal revenue across all firms in the data set.

The change of land price distribution is shown in Figure 6, which draws the land price distribution across all firms for the original estimates $\beta = 0.45$, $\beta = 0.405$ (10% decrease), $\beta = 0.338$ (25% decrease), and $\beta = 0.225$ (50% decrease).

Figure 6: Change of land price distribution due to decrease of β



Notes: The estimates $\hat{\beta} = 0.45$. $\beta = 0.414,\ 0.345,\ 0.230$ are equivalent to 10%, 25%, 50% decrease of $\hat{\beta}$ respectively. For each β, I do 30 simulations for each firm to draw the empirical CDF plot.

The figure shows that as β decreases, the empirical CDF curve of land price shifts to the right, which means more lands are sold at higher prices as β decreases. However, for $\beta = 0.405$ and $\beta = 0.338$, the empirical CDFs are close to the baseline ($\beta = 0.45$). Thus, the impact of fiscal centralization on land prices is limited.

The reason behind this change of price distribution is reflected by the trade-off between land-selling revenue and the fiscal revenue generated by landing the firm, which is characterized by (7). Intuitively, as the governments' revenue share of firms' output β decreases, the expected fiscal revenue of attracting firms using low land prices also decrease. Thus, the local officials care more about land-selling revenue even though higher land prices will decrease their probability of getting the firms. However, even if β decreases by 25%, the spillover effects of attracting firms are still sufficiently large, as I discussed in Section 6.2. Thus, the rise of land prices is limited.

I also calculate the percent changes in the average land price, total land selling revenue, and total fiscal revenue at different values of β, and the results are shown in Table 6. It shows the average land price and total land selling revenue increase as β decreases, but the increase is not large compared to the decrease of β. However, I also notice that the total fiscal revenue, which is the summation of total land selling revenue and total output share decreases sharply as β decreases. This is because the land-selling revenue constitutes only a small proportion of the total fiscal revenue compared to the huge revenue generated by landing firms.

Table 6: The impacts of decrease in β

β	change of β	average land price	total land selling revenue	total fiscal revenue
0.405	-10%	2.52%	1.99%	-8.99%
0.338	-25%	7.16%	5.65%	-22.43%
0.225	-50%	17.38%	14.04%	-44.62%

To conclude my discussion in this subsection, the trend of fiscal centralization after 2013 might restrict the fiscal competition between local governments and increase the industrial land price if the total output level across the nation is stable. However, the increase in land price is limited, and the loss of fiscal revenue due to the decrease in local governments' output share is difficult to be compensated by the increasing industrial land-selling revenue. Thus, fiscal centralization poses challenges for the central government in how to compensate the local governments' potential fiscal revenue loss.

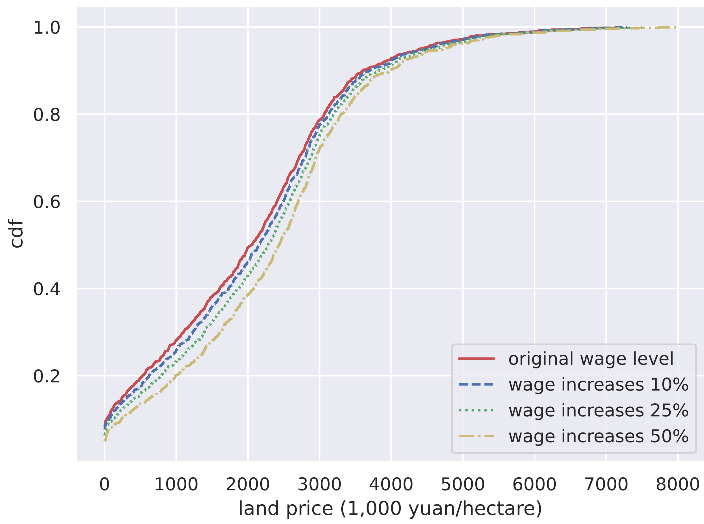
7.3 Potential impacts of the rising wage level

In recent years, the impacts of rising urban wages on China's economy are intensely discussed in both economic literature and public policy debates. The rising urban wage in China after the late 2000s is caused by the rural surplus labor crossing the Lewisian turning point (Cai and Wang, 2010) and the age structure change of population (Cai, 2016). I am particularly interested in the potential impacts of the rising urban wage on the fiscal competition between city governments since labor and land are both inputs in the firm's production function.

To implement the counterfactual analysis, I increase the wage level in all cities by 10%, 25%, and 50% respectively, and examine the impacts on average land price, total land selling revenue, total fiscal revenue by simulations.

I draw the empirical distributions of industrial land price under different rises of wage in Figure 7. As Figure 7 shows, the new empirical CDF curves of land price are close to the baseline if wage increases by 10% or 25% in all cities. If the wage level rises by 50% in all cities, the distribution shifts to the right more clearly. Nevertheless, all three cases show that the average land price is higher if the wage level increases in all cities.

Figure 7: Change of land price distribution due to increase of wage



Notes: The preferred estimates are used in all four cases. I do thirty simulations for each firm to draw the empirical CDF plots.

The intuition behind the rise of average land price as wage level increases in the whole nation is that labor and land are both inputs in the firm's production function, and the national-wide wage increase will make the relative price of labor to land more expensive than before. In other words, firms will be more sensitive to labor costs in their location choice problem than before, and city governments will raise the land price since firms are not as sensitive to land prices as before.

I also show the impacts of the increase in wage on average land price, total land selling revenue, and total fiscal revenue in Table 7. The average land price increases as the wage level increases, but the increase is not sharp.

Table 7: The impacts of rising wage level

wage increase	average land price	total land selling revenue	total fiscal revenue
10.00%	3.96%	5.38%	0.45%
25.00%	9.21%	12.63%	1.06%
50.00%	16.49%	23.00%	1.93%

To conclude my discussion in this subsection, I find that the rising urban wage would increase the average industrial land price in China, i.e. restrict the fiscal competition, but just to a small extent. This implies it is difficult to change the pattern of fiscal competition by the change of production factor prices (pure market power).

8 Conclusion

In this paper, I have developed and estimated a static Bertrand game model between city governments in China as they use land sales discounts to attract industrial firms. The estimates show local governments can get huge benefits from attracting industrial firms. And the ratio of fiscal revenue to firms' output is much higher than the official tax rate.

I show that if the total output level in China doesn't change a lot by fiscal competition, i.e., firms still want to produce in China if the land price is decided by market power rather than special deals between local governments and firms, the major tool for fiscal competition: selling industrial land at low prices cannot improve the allocation efficiency of output but waste potential land-selling revenue. In this sense, my study not only sheds light on the mechanism of fiscal competition with Chinese characteristics, but also supports the tax competition model (Wilson, 1999) in the public economics literature. The counterfactual analysis also shows the decline of the economic autonomy of local governments poses potential challenges to the mode of fiscal competition. But rising wages might not have significant influence on the fiscal competition pattern.

However, my study doesn't imply the fiscal competition between city governments in China is worthless. The real potential benefits of fiscal competition can come from the case that more firms (or higher output levels) are created due to the lower production cost induced by the efforts of local governments to attract firms. But this issue is beyond the scope of my analysis since the firm's output level is exogenous in the model, and I don't consider the entry and exit problems of firms. Analyzing these issues can be a possible extension of the paper.

Incorporating a labor market into my model might be another interesting extension since local governments can use policy instrument (e.g. household registration system reform, enhancing social welfare, etc.) to attract workers to their jurisdictions and decrease the production cost of firms indirectly. And I leave these two possible extensions for further research.

Appendix A: Proofs of the Theorems

Proof of Theorem 1. $f(p_{min}) \leq 0$ implies $\frac{\partial v_{ik}}{\partial p_{ik}} < 0$ for all $p_{ik} \in (p_{min}, p_{max}]$, thus city government k will sell the land at the minimum price. The same logic holds for the case where $f(p_{max}) \geq 0$. If $f(p_{min}) > 0$ and $f(p_{max}) < 0$, there must be exactly one p_{ik} on (p_{min}, p_{max}) s.t. $\frac{\partial v_{ik}}{\partial p_{ik}} = 0$, which is exactly the F.O.C. of k's expected fiscal revenue maximization problem. $f(p_{min}) < 0$ and $f(p_{max}) > 0$ is impossible since $f(p_{ik})$ is decreasing on $\mathbb{R}$. □

Proof of Theorem 2. For any Bertrand Game $\langle \mathbf{C}_i, (S_{ik}), (v_{ik}) \rangle$, I drop the subscript i for convenience. And I define another mapping $F : \mathbb{R}^N \to \mathbb{R}^N$, such that $F(p) = p'$, where p'_k is the root of $f(p_k) = 0$ given $p_{(-k)} = (p_j)_{j \neq k}$. As defined in Section 3.4:

$$f(p_k) = 1 - \frac{1}{\sigma m}\big(1 - P_k(p_k, p_{(-k)})\big)\big(\beta Y + p_k T\big).$$

I fix k and denote the choice probability for city j by $P_j \equiv P_j(p'_k, p_{(-k)})$ for $\forall j$, i.e., the price vector is fixed as $(p'_k, p_{(-k)})$. I define a metric ρ on S^N: $\rho(p, q) = max_k |p_k - q_k|$.

Differentiating both sides of $f(p'_k) = 0$ w.r.t. p_j $(j \neq k)$ and rearrange it, then I get:

$$\frac{\partial p'_k}{\partial p_j} = \frac{\frac{1}{m\sigma} P_k P_j (\beta Y + p'_k T)}{\frac{1}{m\sigma} P_k (1 - P_k)(\beta Y + p'_k T) + (1 - P_k)} < \frac{P_j}{1 - P_k}.$$

Notice that $\frac{\partial p'_k}{\partial p_j}$ is continuous in $p_{(-k)}$ since P_k and P_j are continuous in $p_{(-k)}$. Thus, p'_k is differentiable everywhere w.r.t. $p_{(-k)}$ since all the partial derivatives $\frac{\partial p'_k}{\partial p_j}$ exist and are continuous functions of $p_{(-k)}$. Then according to the mean value theorem, for any $p,\ q \in S^N$:

$$\begin{aligned} |F(p)_k - F(q)_k| &= |\nabla p'_k(\hat{p}_{(-k)}) \cdot (p_{(-k)} - q_{(-k)})| \\ &< \sum_{j\neq k} \frac{\hat{P}_j}{1-\hat{P}_k} \rho(p,q) \\ &= \rho(p,q), \end{aligned} \tag{16}$$

where $\nabla p'_k$ is the gradient of p'_k evaluated at some $\hat{p}_{(-k)}$ between $p_{(-k)}$and $q_{(-k)}$, $\hat{P}_j = P_j(p'_k, \hat{p}_{(-k)})$. The last equality holds since $\sum_{j\neq k} \frac{\hat{P}_j}{1-\hat{P}_k} = 1$.

From (16) and the definition of G in Theorem 2 I have:

$$|G(p)_1 - G(q)_1| = |F(p)_1 - F(q)_1| < \rho(p,q).$$

Then I redefine $p^\star = [G(p)_1, p_2, \ldots, p_N]$, $q^\star = [G(q)_1, q_2, \ldots, q_N]$, using (16) again, I have

$$|G(p)_2 - G(q)_2| = |F(p^\star)_2 - F(q^\star)_2| < \rho(p^\star, q^\star) \leq \rho(p,q).$$

The last inequality holds since $|G(p)_1 - G(q)_1| < \rho(p,q)$ and $|p_j - q_j| \leq \rho(p,q)$ for all $j \geq 2$. Iteratively, It can be proved that $|G(p)_k - G(q)_k| < \rho(p,q)$ for $\forall k$, i.e., $\rho(G(p), G(q)) < \rho(p,q)$.

I have proved that G is contracting on the compact set $S^N = [p_{min},\ p_{max}]^N$ in $\mathbb{R}^N$. Thus, G has a unique fixed point p^*, and $G^n(p) \to p^*$ for any $p \in S^N$ as $n \to \infty$. Since p^* is a pure NE if and only if p^* is a fixed point of G, the Bertrand game has a unique pure NE. □

Appendix B.1: Estimation Under Different Settings of Choice Sets

I present the estimation results under the other two settings of the choice sets $\mathbf{C}_i$ here.

First, instead of (12), I set $|\mathbf{C}_i| = 3$ if $Y_i \leq Median(Y)$, $|\mathbf{C}_i| = 4$ if $Y_i > Median(Y)$. The estimation results are shown in Table 8 below.

Table 8: Estimates of Parameters ($3 \leq |\mathbf{C}_i| \leq 4$)

Calibrated α	Parameters	Estimates	Standard Error
0.33	β	0.537	0.046
	σ	180.438	23.666
0.5	β	0.581	0.047
	σ	190.297	25.614
0.67	β	0.580	0.048
	σ	196.159	25.878

As Table 8 shows, the estimates $\hat{\beta} \approx 0.54,\ 0.58,\ 0.58$ for $\alpha = 0.33,\ 0.5,\ 0.67$ respectively. Thus, $\hat{\beta}$ under this setting are higher than my main results due to the decrease of the size of choice sets, i.e., the fiscal competition is less fierce than the settings in the main text, and $\hat{\beta} = 0.45$ is a conservative estimate.

Second, I set $|\mathbf{C}_i| = 4$ if $Y_i \leq Median(Y)$, $|\mathbf{C}_i| = 5$ if $Y_i > Median(Y)$. The estimation results are shown in Table 9 below.

Table 9: Estimates of Parameters ($4 \leq |\mathbf{C}_i| \leq 5$)

Calibrated α	Parameters	Estimates	Standard Error
0.33	β	0.391	0.027
	σ	177.253	22.551
0.5	β	0.374	0.026
	σ	196.125	24.497
0.67	β	0.387	0.028
	σ	203.421	26.099

As Table 9 shows, the estimates $\hat{\beta} \approx 0.4$ in both cases of $\alpha = 0.33,\ 0.5,\ 0.67$, which is slightly lower than the main results in Table 3. However, the setting that half of the firms have 4 candidate cities and the other half have 5 candidate cities makes the competition for firms very fierce even for a smaller β, thus, the estimates in Table 9 indeed verifies the robustness of my main results in Table 3 since $\hat{\beta}$ is not far away from 0.45 even if the fiscal competition is extremely fierce.

Appendix B.2: Estimation Under Different Participation Constraints

I present the estimation results under the other two settings of participation constraints (action space) here. The specification of choice set $\mathbf{C}_i$ here is the same as which in the main text, i.e., $3 \leq |\mathbf{C}_i| \leq 5$.

First, I try to shrink the action space and set $p^i_{min} = \max\{0,\ p^i_{obs} - 1000\}$, $p^i_{max} = p^i_{obs} + 1000$. The estimation results are shown in Table 10.

Table 10: Estimates of Parameters ($3 \leq |\mathbf{C}_i| \leq 5$, smaller action space)

Calibrated α	Parameters	Estimates	Standard Error
0.33	β	1.122	0.801
	σ	103.807	34.941
0.5	β	0.483	0.033
	σ	181.004	24.844
0.67	β	0.491	0.037
	σ	184.535	29.319

As Table 10 shows, the estimates of β are slightly higher than $\hat{\beta} = 0.45$ in the main text. But the standard error of $\hat{\beta}$ for $\alpha = 0.33$ is extremely large, which implies there might be a weak identification problem when the action space is small. Thus, the specification of action space is less appropriate than which in the main text.

Second, I try to expand the action space and set $p^i_{min} = \max\{0, p^i_{obs} - 3000\}$, $p^i_{max} = p^i_{obs} + 3000$. The estimation results are shown in Table 11 below.

Table 11: Estimates of Parameters ($3 \leq |\mathbf{C}_i| \leq 5$, larger action space)

Calibrated α	Parameters	Estimates	Standard Error
0.33	β	0.572	0.037
	σ	221.916	26.406
0.5	β	0.613	0.039
	σ	245.214	26.756
0.67	β	0.581	0.037
	σ	227.853	27.511

As Table 11 shows, the estimates of β are higher than $\hat{\beta} = 0.45$ in the main text. But $\hat{\beta} = 0.61$ for $\alpha = 0.5$ is still not far away from $\hat{\beta} = 0.45$.

Notice that in both Table 10 and Table 11, the estimates of β with small standard errors vary from 0.48 to 0.61, thus, $\hat{\beta} = 0.45$ is a conservative estimate. To conclude, the specification of action space in the main text is appropriate for my analysis.

References

ADDA, J., AND R. W. COOPER (2003): *Dynamic Economics: Quantitative Methods and Applications*: MIT press.

BAI, C.-E., C.-T. HSIEH, AND Z. SONG (2020): "Special deals with Chinese characteristics," *NBER Macroeconomics Annual*, 34 (1), 341–379.

Brandt, L., C.-T. Hsieh, and X. Zhu (2008): "Growth and structural transformation in China," in *China's great economic transformation* ed. by Brandt, L., and Rawski, T. G.: Cambridge University Press, 683–728.

Cai, F. (2016): *China's economic growth prospects: From demographic dividend to reform dividend*: Edward Elgar Publishing.

Cai, F., and M. Wang (2010): "Growth and structural changes in employment in transition China," *Journal of Comparative Economics*, 38 (1), 71–81.

Cai, H., J. V. Henderson, and Q. Zhang (2013): "China's land market auctions: evidence of corruption?" *The Rand Journal of Economics*, 44 (3), 488–521.

Chen, T., L. Liu, W. Xiong, and L.-A. Zhou (2017): "Real estate boom and misallocation of capital in China," *Working Paper, Princeton University.*

Cheung, S. N. (2014): "The economic system of China," *Man and the Economy*, 1 (1), 1–49.

Gouriéroux, C., and A. Monfort (1996): *Simulation-Based Econometric Methods*: Oxford University Press.

Iskhakov, F., and M. Keane (2021): "Effects of taxes and safety net pensions on life-cycle labor supply, savings and human capital: The case of Australia," *Journal of Econometrics*, 223 (2), 401–432.

Keen, M., and M. Marchand (1997): "Fiscal competition and the pattern of public spending," *Journal of Public Economics*, 66 (1), 33–53.

Kroeber, A. R. (2020): *China's Economy: What Everyone Needs to Know*: Oxford University Press.

Liu, C., and W. Xiong (2020): "China's Real Estate Market," in *The Handbook of China's Financial System* ed. by Amstad, M., Sun, G., and Xiong, W.: Princeton University Press, Chap. 7, 183–207.

Mast, E. (2020): "Race to the bottom? Local tax break competition and business location," *American Economic Journal: Applied Economics*, 12 (1), 288–317.

McFadden, D. (1974): "Conditonal logit analysis of qualitive choice behavior," in *Frontiers in Econometrics* ed. by Zarembka, P. New York: Academic Press, Chap. 4, 105–142.

——— (1989): "A method of simulated moments for estimation of discrete response models without numerical integration," *Econometrica*, 995–1026.

Pakes, A., and D. Pollard (1989): "Simulation and the asymptotics of optimization estimators," *Econometrica*, 1027–1057.

Su, F., and R. Tao (2017): "The China model withering? Institutional roots of China's local developmentalism," *Urban Studies*, 54 (1), 230–250.

Tiebout, C. M. (1956): "A pure theory of local expenditures," *Journal of Political Economy*, 64 (5), 416–424.

Wilson, J. D. (1999): "Theories of tax competition," *National Tax Journal*, 52 (2), 269–304.

Xiong, W. (2018): "The mandarin model of growth," *Working Paper, National Bureau of Economic Research.*

Xu, C. (2011): "The fundamental institutions of China's reforms and development," *Journal of Economic Literature*, 49 (4), 1076–1151.

Zhu, X. (2012): "Understanding China's growth: Past, present, and future," *Journal of Economic Perspectives*, 26 (4), 103–24.