Land-Use Regulation and Economic Development: Evidence from the Farmland Red Line Policy in China*

Yue Yu†

Columbia University

November 24, 2019

Download the latest version here

Abstract

Many countries have land-use regulations to preserve farmland for food security reasons. In this paper, I show that such regulations can distort economic activity across sectors and locations at a substantial cost to aggregate welfare in developing countries during urbanization. I study a major policy restricting farm-to-urban land conversion in China - the Farmland Red Line Policy - to provide causal evidence on the impact of land-use regulation on local development measured by GDP and population growth. The policy imposes a barrier to urban land development, the strength of which depends on exogenous local geographical features. I show that a greater barrier significantly reduces urban land supply, lowers GDP, and decreases population. To understand the aggregate impact of the policy, I develop a quantitative spatial equilibrium model that features endogenous land-use decisions. According to the model, the policy causes an excess supply of farmland and an under-supply of urban land, and the extent of such land misallocation varies across locations due to their local geographical features. In the constrained equilibrium, the spatial and sectoral mobility of workers implies that land misallocation leads to labor misallocation. The calibrated model reveals that the welfare of workers would have been 6% higher in 2010 if the policy had not been implemented. Moreover, a cap-and-trade system that achieved the same aggregate level of farmland would have been far less costly in terms of welfare. The results suggest that fast-growing economies in developing countries need to design land-use policies carefully, as the welfare costs of poorly designed policies can be substantial.

JEL: O1, O5, Q1, R1, R5

Keywords: Land-use regulation, economic development, urbanization, China

*I am very grateful to Donald Davis, Jonas Hjort, Réka Juhász and David Weinstein for their invaluable guidance and continued support. I thank Eric Verhoogen, Michael Best, Jack Willis, Cristian (Kiki) Pop-Eleches, Suresh Naidu, Bernard Salanie, Douglas Almond, Shang-Jin Wei, Dávid Nagy, Cailin Slattery, Christian Hilber, Matthew Turner, Gilles Duranton, Walker Hanlon, Stephen Redding, Ferdinando Monte, David Cuberes, Clare Balboni, Jing Li, Yi Fan, Shuang Zhang, Lin Tian, Yang Jiao, Danyan Zha, Qing Zhang, Haaris Mateen, Iain Bamford, Pablo Ernesto Warnes, Timur Abbiasov, Sang Hoon Kong, Sun Lee, Anurag Singh, Junlong Feng, Ye Zhang, Yining Zhu, and seminar participants at Columbia, Urban Economics Association, AREUEA, ESEM, PWSD, and YES for helpful discussions and comments. All errors are mine.

†Department of Economics, 420 W. 118th St., New York, NY 10027. Email: yy2558@columbia.edu.

1 Introduction

Most countries have land-use regulations that can affect local and national economic development (Duranton and Puga, 2015; Hsieh and Moretti, 2018; Turner et al., 2014). Many of these regulations are designed to ensure food security by protecting farmland from urban sprawl (Liu et al., 2018; Nelson, 1992). Such regulations are set to attract even more attention as the accelerated expansion of urban land and the fast growth of food demand potentially endanger food security in many countries in the next three decades (van Vliet et al., 2017; Seto et al., 2011). However, an often-missing element in these policy discussions is recognition that these regulations can also generate substantial costs. These policies essentially create frictions in land reallocation from the less productive agricultural sectors to the more productive urban sectors. But empirical evidence on how such regulations affect urbanization and economic development is lacking.

In this paper, I study the impact of land-use regulation on economic development using China's Farmland Red Line Policy (1999) - a national land-use policy motivated by food security concerns - as a natural experiment. I examine both the local impacts of the policy on urban land supply, GDP and population and the aggregate impacts of the policy on workers' welfare. The policy aims to preserve at least 1.8 billion acres of farmland nationwide by specifying that each local region must maintain a certain amount of farmland. The policy was implemented in the middle of the largest rural-to-urban migration in human history. From 1980 to 2010, 470 million Chinese people moved from rural to urban areas, and urbanization inevitably involved the conversion of former farmland to urban use. The Chinese government had grown concerned that the conversions of farm to urban land might compromise food security and hence adopted the policy to restrict such conversions (Chen, 2007).

I take three steps to analyze this policy. First, I show that the policy imposes a barrier to rural-to-urban land conversion that depends on exogenous local geographical features. Using a reduced-form analysis, I find that a higher land conversion barrier in a local region significantly reduces its urban land supply, GDP and population. Second, I develop a general equilibrium model to quantify the aggregate welfare cost of the policy. My calibration of the model produces an estimate of the aggregate cost arising from the policy of 6% of workers' welfare. Moreover, distortions from the policy on urbanization manifest mostly in the over-congestion of urban sectors as opposed to less urbanization. Finally, I show that the Farmland Red Line Policy is an inefficient means of protecting farmland by conducting a counterfactual exercise in which the government institutes a cap-and-trade platform that allows local regions to exchange farmland preservation requirements. I show that this system would eliminate 60% of the cost to workers' welfare.

First implemented in 1999, the Farmland Red Line Policy forbids the conversion of farmland into urban land unless an equal amount of unused arable land (within a city jurisdiction) is converted into farmland. The quality of the farmland has to be comparable to the existing farmland in the same city jurisdiction. City jurisdictions in China are administratively categorized into urban and rural land. Rural land is then subdivided into farmland, unused arable land, and unused non-arable land. There was no restriction on the conversion of farmland into urban land before 1999. From 1999, the policy guarantees that the total amount of farmland within each city jurisdiction does not decrease due to urbanization. It is equivalent to a minimum quantity constraint on farmland for each city jurisdiction.

The policy creates an additional cost in urban land development that varies across city jurisdictions. Since the policy was adopted, whenever urban land is converted from farmland, local governments have to bear the additional cost of creating an equivalent amount of new farmland somewhere else in the city jurisdiction. Therefore, the additional cost per unit of new urban land imposed by the policy equals the percentage of new urban land converted from farmland times the unit cost of new farmland development. The cost of farmland development refers to the labor and material expenses involved in the cultivation of unused arable land. Attempts to directly compare the economic outcomes of city jurisdictions with different additional costs of urban land development are problematic because the additional costs are endogenous. For example, a positive urban sector productivity shock after 1999 could boost the local economy and increase the development of urban land. As less unused land is left, the unit cost of new farmland development would increase. As a result, the unobservable productivity shock is an omitted variable that biases the estimates upward.

I isolate exogenous variation in the additional cost of urban land development and define it as the *land conversion barrier*. The exogenous variation consists of two components that are arguably unrelated to historical development patterns or the current development potential of the jurisdiction but do affect the additional cost of urban land development. The first component is the soil-quality predicted farmland percentage of land close to existing urban land prior to 1999. It affects the percentage of urban land converted from farmland because land close to existing urban land is likely the first developed into urban land. The second component is the ruggedness of land close to the administrative boundary. It affects the unit cost of new farmland development because land close to the administrative boundary is most likely used for farmland development, and land ruggedness is the main obstacle to land cultivation. I define the *land conversion barrier* as the interaction of the predicted farmland percentage and land ruggedness. The land conversion barrier is essentially an exogenous determinant of the additional cost of urban land development.

To identify the impact of the land conversion barrier on local economic outcomes, I use a difference-in-difference (DD) estimator with continuous treatment intensity. The empirical strategy is to compare economic outcomes between city jurisdictions with different levels of land conversion barriers before and after 1999. The identification relies on the assumption of parallel trends across city jurisdictions with different land conversion barriers throughout the period in my study had the policy never been adopted. In support of the assumption, I do not find any systematically different trends of growth in terms of the major outcome variables before 1999. A series of alternative regression specifications are provided as robustness checks.

The first empirical finding is that a one-standard-deviation increase in the land conversion barrier significantly reduces a city jurisdiction's urban land by 5.5%, its GDP by 3.9%, and its population by 4.5% after the policy is adopted. This is, to the best of my knowledge, the first causal evidence on the impact of the land-use regulation on local GDP and population. The existing literature exploits variation in land-use regulations at a more local level and focuses on the housing market (Turner et al., 2014; Libecap and Lueck, 2011; Mayer and Somerville, 2000). Next, the decrease in GDP is driven by the decrease in GDP in the secondary sector, which is dominated by the manufacturing sector and hence uses urban land most intensively in production. I also find that the land conversion barrier is positively associated with urban land prices during the 2000s. Finally, I find that the causal relation cannot be explained by alternative channels such as the deterioration of urban compactness or poorer government service provision in more affected jurisdictions. Altogether, the empirical evidence shows that because of the Farmland Red Line Policy, city jurisdictions with a lower land conversion barrier were able to create more urban land after the policy was adopted. Ceteris paribus, urban land is cheaper in these city jurisdictions, which encouraged more workers to move in during China's period of rapid urbanization. Therefore, these city jurisdictions have higher GDP and more population comparing to those in which the land conversion barrier is higher.

The identified local effects of the policy on economic outcomes raises the question of whether the policy generates any significant aggregate impacts on the economy, which matters for policy evaluation. The policy is less of a concern if it does not create any inefficiency at the aggregate level but simply causes a reallocation of economic activities from more treated to less treated locations. Because locations are interlinked through flows of trade and migration, the impact of the policy on a more treated location may generate spillover effects on less treated locations in general equilibrium. Therefore, it is challenging to infer the aggregate impacts solely based on the estimates of the local effects. A general equilibrium model can incorporate the interlinkages between locations and separate the spillover effects from the direct impacts. The model can also simulate counterfactual outcomes under alternative policies and provide informative guidance on a more efficient design of the policy, which the Chinese central government has been considering.

To evaluate the aggregate effects of the Farmland Red Line Policy, I develop a static quantitative spatial equilibrium model that features endogenous land use decisions. In the model, each location has both an urban area and a rural area. Each location-sector produces a variety of a final product. There are two types of agents: workers and landlords. Workers maximize utility by choosing their location and sector, supplying one unit of labor to earn wage income, and spending income on tradable goods and residential land. Each location has a representative and immobile landlord who owns a continuum of land plots. A land plot can be developed into farmland or urban land at a cost or remain unused land. Agricultural and urban sector workers rent farmland and urban land for production and residential use. Landlords choose the amounts of urban land, farmland, and unused land to maximize land development profit.

According to the model, aggregate welfare loss first comes from cross-sector and cross-location land misallocation caused by the Farmland Red Line Policy. Without the Farmland Red Line Policy, the landlord maximizes profit by equalizing the price of urban land, the price of farmland, and the marginal cost of land development. The policy imposes a minimum quantity constraint on farmland. When the constraint is binding, more farmland is created to meet the minimum quantity relative to the no-policy market equilibrium. This increases the marginal cost of land development and hence decreases the urban land supply. Therefore, there is an undersupply (oversupply) of urban land (farmland) compared to the no-policy market equilibrium. Furthermore, the degree of cross-sector land misallocation is smaller in a location if the supply of developed land is more elastic. As the supply elasticity of developed land varies across locations, so does the degree of cross-sector land misallocation.

In general equilibrium, spatial and sectoral labor mobility indicate that land misallocation causes labor misallocation, the second source of aggregate welfare loss. First, an oversupply of farmland and an undersupply of urban land leads to an oversupply of rural workers because farmland is cheaper and an undersupply of urban workers because urban land is more expensive, compared to the no-policy market equilibrium. Second, the degree of land misallocation varies across locations, which leads to the variation in labor misallocation across locations. When an undersupply of urban land occurs in productive yet highly constrained locations, workers have to reside in more affordable yet less productive locations.

I structurally estimate the model to quantify the aggregate welfare loss from the Farmland Red Line Policy. The parameter unique to my setting is the price elasticity of unused land and is identified using the variation from land ruggedness as in the reduced-form analysis. The rest of the parameters are calibrated to match values either from aggregate data or the literature. With the parameters and the observed GDP, employment, land use, and land features, I recover the unobserved productivities, amenities, and prices that rationalize the observed data from 2010 as an equilibrium of the model. I show that the model performs well in simulating the reduced-form results. Amenities and productivities recovered from the model correlate well with observable proxies such as local FDI and presence of theaters and museums. Three counterfactual analyses are conducted based on the calibrated model.

My quantitative model first produces an estimate of a 6% welfare loss for workers arising from the Farmland Red Line Policy. The estimate is derived by comparing the simulated counterfactual equilibrium without the policy and the reality in 2010. Next, the economy would have specialized more in the manufacturing sector in the no-policy counterfactual equilibrium. Specifically, manufacturing output would have been 5.0% higher, while agricultural output would have been 2.8% lower.

One important question is how the policy intervened in the urbanization process between 1999 and 2010, as the policy was adopted when rural-to-urban migration accelerated. The policy-induced undersupply of urban land would both make urban areas more congested and slow urbanization. A quantitative exercise shows that distortions from the policy manifest mostly in overcrowding in urban areas as opposed to less urbanization. Without the policy, the urban population would have been 5.2% higher in 2010. This is not economically large when compared to the real-world increase in the urban population from 1999 to 2010, which was more than 40%. In contrast, without the policy, there would have been 40% more urban land in 2010. This indicates that urban population density would have been dramatically lower, by 25%, decreasing from 12,170 to 9,249 per sq. km.

Next, I show that using a cap-and-trade platform is a more efficient way of protecting farmland and food security than the Farmland Red Line Policy. Through this platform, a local government in one location can pay another to create new farmland if the former location needs to convert farmland into urban land. This cap-and-trade platform guarantees that the amount of farmland nationwide does not decrease, while the amount of farmland in each individual location is allowed to change. I simulate a counterfactual equilibrium with the cap-and-trade platform and find that the platform can eliminate 60% of workers' welfare loss from the Farmland Red Line Policy. Specifically, in the counterfactual equilibrium with a cap-and-trade platform, the welfare of workers would have been 3.5% higher than it was in reality in 2010.

Finally, I find that an alternative cap-and-trade platform that the Chinese central government plans to use is much less effective in reducing workers' welfare loss. Specifically, the Chinese central government announced in 2018 that a cap-and-trade platform for farmland protection is under discussion. Locations in the more developed regions pay 1.6 to 4 times the listed price on the trading platform to have a unit of farmland preserved somewhere else (Notice of the General Office of the State Council [2018] No.16). My estimates show that this alternative platform would only save approximately 40% of workers' welfare loss from the Farmland Red Line Policy because the design relatively restricts urban land expansion and urbanization in more productive regions.

Related Literature

This paper is the first to provide both reduced-form estimates of the local effects of land-use regulations and a quantification of the aggregate welfare implications. This differentiates my paper from the existing literature that relies on structural estimation to estimate the aggregate effects of urban land-use regulation (Hsieh and Moretti, 2018; Bunten, 2017; Parkhomenko, 2018; Allen et al., 2015). The reduced-form evidence helps validate the mechanisms built into the quantitative model.

Second, this paper highlights the role of land-use regulations in shaping the spatial allocation of economic activities, which is often neglected in the literature on spatial economics (Redding and Rossi-Hansberg, 2017; Redding and Sturm, 2008; Desmet and Rossi-Hansberg, 2013; Donaldson and Hornbeck, 2016; Gollin et al., 2016). My paper shows that land-use policies can effectively change the spatial distribution of economic activities, which matters for production efficiency at the aggregate level.

Third, my paper emphasizes land-use regulations as an important source of friction in the resource reallocation during structural economic change (Syrquin, 1988; Williamson, 1988; Banerjee and Duflo, 2005). The ease of resource reallocation matters for the speed of convergence of the economy to the steady state after productivity shocks across sectors (Gollin et al., 2002; Lagakos et al., 2018). This literature has focused on frictions in the labor market. However, as my paper reveals, land market frictions also significantly affect economic transformation, and one major source of frictions is land-use regulations.

Finally, my paper provides a new example of how geographical features affect urban land supply and economic development (Saiz, 2010; Harari, 2018; Nunn and Qian, 2011; Marden et al., 2015; Nunn and Puga, 2012). Different from the existing literature, the link between the geographical features and urban land supply is established because of the land-use policy.

The remainder of this paper proceeds as follows. Section 2 provides institutional information about the Farmland Red Line Policy. Section 3 details the data and unit of analysis. Section 4 conducts regression analysis to study the effect of the Farmland Red Line Policy on local economic development. Section 5 develops a static spatial equilibrium model to demonstrate the impacts of the Farmland Red Line Policy on land allocation and labor allocation. Section 6 quantifies the model and calibrate it to the benchmark year. Section 7 estimates the aggregate effects of the Farmland Red Line Policy. Section 8 evaluates two counterfactual policies and compare the counterfactual outcome with reality. Second 9 concludes the paper.

2 Policy Background

This section provides a detailed discussion of the Farmland Red Line Policy and the relevant institutional background. I first outline the institutional information about land classification in China. Within a city jurisdiction, land is divided into urban land, farmland, unused arable land, and unused non-arable land. Figure 1 provides an example of a stylized city jurisdiction. Urban land is typically located near the geometrical center of a city jurisdiction, surrounded by a mix of farmland, unused arable land, and unused non-arable land. To create new urban land, a local government acquires farmland and unused land from rural residents at the urban fringe and convert it into urban land.[1] After this step, newly developed urban land will be transferred to independent real estate developers through long-term leaseholds. Most land development profit counts as local government revenue (Chen and Kung, 2016).

The Farmland Red Line Policy was implemented at a time when urbanization was accelerating in China and the country switched from being an exporter to being an importer of agricultural products. From 1980 to 2010, 469.4 million Chinese people moved from rural areas to urban areas.[2] On the one hand, the rapid urbanization drew land and labor from the agricultural sector to the manufacturing and service sector.[3] On the other hand, rising household income increased domestic consumption of food. As a result, China's switch from agricultural exporter to importer occurred in 1995 (Brown A., 1995). In September 1995, Lester Brown published the book 'Who Will Feed China?', in which he pessimistically predicted that industrialization in China would soon make China a large importer of food. This would drive up the world food price and cause food shortages in the long run.[4] The book alerted the Chinese central government that rapid urbanization might soon endanger national food security (Wang et al., 2010).

Due to food security concerns, it became national priority to preserve at least 1.8 billion acres of farmland ('National Land Use Plan for 1997-2010') in early 1997. This number is close to the total national amount of farmland around that time. When the goal was first announced, the central government was unclear about how to achieve the goal without fully halting urban land expansion. During 1997 and 1998, members of the central government quickly developed all of the regulatory details of the Farmland Red Line Policy. Meanwhile, to prevent local governments from over-expanding urban land before the new regulation came into force, the central government prohibited rural-to-urban land conversion nationwide.[5] The Farmland Red Line Policy was announced in August 1998. For the

[1]By Chinese law, urban land and most of the unused land is owned by the state; farmland is collectively owned by local rural residents, who are represented by a village committee in each local area. To convert farmland into urban land, the local government first pays compensation to rural residents to transfer ownership of the land to the state. The compensation is based on the output value of the farmland, and the rural residents have little bargaining power. After the land is acquired, the local government cleans up the surface and provides urban infrastructure such as electricity systems for this new piece of land.

[2]Urban population time series data come from the World Bank.

[3]In 1990, there were only 23,640 km^2 of urban land (GeoExplorer II). If no new urban land was created between 1990 and 2010 and the amount of rural-to-urban migration held the same as in reality, by 2010, population density in urban China would have reached 27,941/km^2. This would be even higher than Manhattan's population density, which is 25,846/km^2 (Wikipedia).

[4]The main argument in the book is that the reduction of in farmland combined with an expected increase in food demand as the Chinese population became wealthier would make China a big large importer in the world food market.

[5]The only exception is the national key projects detailed in No.11[1997] of the CPC Central Committee and State

first time, strict regulations were in use to protect farmland, even if it could slow the urbanization process.[6]

The Farmland Red Line Policy prohibits a local government from converting farmland into urban land unless an equal amount of unused land is converted into farmland. Further, the new farmland should be in the same city jurisdiction as the farmland converted into urban land, unless a local government faces extreme difficulty in creating farmland, in which case it might ask another the government of another city jurisdiction within the same province to help create farmland (No.374[2001] of the Ministry of Land and Resources). In practice, asking another city jurisdiction's government for help involves high political costs, and hence local governments typically create new farmland within their own jurisdictional boundaries (Liu et al., 2005).

To guarantee that the local governments would comply with the policy, the central government adopted a sophisticated supervision system and devoted enormous efforts to monitoring urban land development and farmland change across locations. First, urban land development plans have to be approved by the central government on an annual basis; otherwise, the development is illegal. The central government will pay special attention to any urban land development project that uses more than 35 hectares of prime farmland or more than 70 hectares of farmland in total. Second, starting in 2000, the central government began using remote sensing techniques to detect illegal farmland conversion. If according to remote sensing data, urban land increases significantly more than the amount reported by local governors or there is significant loss of farmland in that year, the local government could be investigated. Next, since 2002, all the land legally converted from rural land to urban land has had to be registered with the national land-use registration platform (No.374[2001] of the Ministry of Land and Resources). Each record provides information on the amount of farmland that used to be on a parcel and detailed information about the location and the condition of newly created farmland. The central government randomly selects records from the database and sends officials to the local area to check whether all the information on the record is accurately documented.

As a result of the enormous efforts devoted to regulating and monitoring urban land development across locations, the Farmland Red Line Policy successfully halted the loss of farmland due to urbanization. At the national level, the total amount of farmland has barely changed since the policy's implementation, as shown in Figure 9. Across city jurisdictions, there is no significant negative correlation between the absolute change in urban land quantity and the absolute change in farmland quantity since the policy began, as shown in Appendix Table 14.

The Chinese central government has recognized the obstacles to urban development created by the Farmland Red Line Policy and has been considering alternative policy designs. To reduce the constraint without endangering national food security, the Chinese central government announced in 2018 that there would be a national trading platform through which one city can pay another to

Council.

[6] Although a farmland resource tax existed to protect farmland since 1980s, the stringency, scope and efforts of monitoring were never comparable to those under the Farmland Red Line Policy (Lichtenberg and Ding, 2008). For example, since 1987, local government has had the authority to impose a farmland resource tax on local urban land developers and users. However, it was not carefully implemented by local governors because of unclear instructions for the tax and its non-mandatory nature. Furthermore, local governors had no incentive to add any constraints to urban development because urban development was a very effective local development strategy.

create new farmland (Notice of the General Office of the State Council [2018] No.16). The platform will ensure that the total amount of farmland at the national level does not decrease, while in each location, the amount of farmland can be reduced as long as farmland increases by the same amount somewhere else. A more detailed discussion of the trading platform is provided in Section 8.

3 Data

To understand the estimation, I first need to explain how the data on city jurisdictions are assembled. I construct panel data on city jurisdictions in China by assembling data from a series of statistical yearbooks and the population census. This panel covers the years from 1990 to 2015 and includes most of the city jurisdictions in China. City jurisdictions cover most urban areas and nearby rural areas, but rural areas farther away are not included. Figure 10 shows the geographical coverage of city jurisdictions. The city jurisdictions represent 80.1% of China's total GDP in 2010. They specialize in non-agricultural sectors. The agricultural GDP from city jurisdictions only accounts for 41.5% of national agricultural GDP, while the secondary sector (manufacturing and construction) and tertiary sector GDP from city jurisdictions account for 87.7% and 80.4% of the respective national totals.

The main outcome variables include the amount of urban land, GDP by sector, and population. Urban land and GDP data come from City Statistical Yearbooks, City Development Yearbooks, and ChinaDataOnline. Urban land refers to land that has been developed and used for non-agricultural activities. Next, GDP by sector at the level of city jurisdictions is only available in the yearbooks since 1994. Therefore, the analysis of GDP outcomes is based on panel data that covers the period 1994 to 2015. Finally, the population data are constructed using the 1982, 1990, 2000, and 2010 waves of the population census.[7]

Several other variables are used in the empirical analysis to investigate the channels of the effects. First, government expenditures per capita and the number of hospital beds from 1990 to 2015 are used to measure government service provision. Second, the average price and the floor-to-area ratio (FAR) of newly developed urban land sorted by land-use purpose, such as residential and commercial, are available from 2007 to 2015. Third, the amount of urban land by land-use purpose is available in the City Development Yearbooks from 2002 to 2015. These variables are used to provide suggestive evidence about how urban land price and urban land use respond to the land conversion barrier. Fourth, the compactness index of the urban area in a city jurisdiction for 1995 and 2010 is constructed using 30-m resolution raster data on urban land cover (Liu et al., 2018). This variable is used to test whether the shape of the urban area deteriorates due to the policy.

Summary statistics of the main variables are provided in Table 1. Appendix B.3 provides a more detailed discussion of all the datasets used in the empirical analysis.

[7]The data are from the China Data Center, University of Michigan. The main advantage of using the population census is that it offers a much more accurate accounting of the number of residents in city jurisdictions. See B.3 for a detailed discussion.

4 Empirical Analysis

This section examines the local effect of the Farmland Red Line Policy on economic development. I find that a one-standard-deviation increase in the land conversion barrier significantly reduces a city jurisdiction's urban land by 5.5%, its GDP by 3.9%, and its population by 4.5% after the policy is adopted. Next, the decrease in GDP is driven by the decrease in GDP in the secondary sector, which is dominated by the manufacturing sector and hence uses urban land most intensively in production. I also find that the land conversion barrier is positively associated with urban land prices during the 2000s. Finally, I find that the causal relation cannot be explained by alternative channels such as the deterioration of urban compactness or poorer government service provision in more affected jurisdictions. Altogether, the empirical evidence shows that because of the Farmland Red Line Policy, city jurisdictions with a lower land conversion barrier were able to create more urban land after the policy was adopted. Ceteris paribus, urban land is cheaper in these city jurisdictions, which encouraged more workers to move in during China's period of rapid urbanization. Therefore, these city jurisdictions have higher GDP and more population comparing to those in which the land conversion barrier is higher.

4.1 Identification Strategy

To identify the impact of the Farmland Red Line Policy on local economic development, I use a DD estimator with continuous treatment intensity, and the cross-sectional variation is from the exogenous component of the additional cost of urban land development imposed by the policy. Since the policy was adopted, whenever urban land is converted from farmland, local governments have to bear the additional cost of creating an equivalent amount of new farmland somewhere else in the city jurisdiction. Therefore, the additional cost per unit of new urban land imposed by the policy equals the percentage of new urban land converted from farmland (*farmland density*) times the unit cost of new farmland development.[8] The cost of farmland development refers to the expenses of labor and materials involved in the cultivation of unused arable land to convert it into farmland of similar quality to the existing farmland within the city jurisdiction.

Attempts to directly compare the economic outcomes of city jurisdictions with different additional costs of urban land development are problematic because both components of the additional cost – farmland density and the unit cost of new farmland development – are endogenous. Farmland density is endogenous because it depends on the farmland percentage of the rural land surrounding existing urban land, where new urban land development typically occurs (van Vliet et al., 2013). The latter can be correlated with local population density and agricultural techniques, which might be associated with unobserved local productivity and affect the growth path of the local economy. Next, the unit cost of farmland development is endogenous because a positive urban sector productivity shock after 1999 could boost the local economy and increase the development of urban land. As less unused land is left, the unit cost of new farmland development would increase. As a result, the unobserved productivity shock can bias the estimation upward.

[8] Figure 11 provides two examples to illustrate each aspect.

I isolate exogenous variation in the additional cost of urban land development and define it as the *land conversion barrier*. The exogenous variation consists of two components that are arguably unrelated to historical development patterns or the current development potential of the jurisdiction but do affect the additional cost of urban land development.

The first component of the land conversion barrier is the soil-quality predicted farmland percentage of land close to existing urban land by 1999. It affects the percentage of urban land converted from farmland because land close to existing urban land is likely the first to be developed into urban land. If the soil qualities there make it more suitable for agricultural production, the farmland density is expected to be higher.

I construct the first component in two steps. I first draw a boundary representing the existing urban area in each city jurisdiction before 1999 and create an outward buffer on the urban boundary.[9] Next, I regress farmland density on the mean soil qualities of land within the buffer and obtain the predicted value of the dependent variable.[10] The *predicted farmland density* is essentially a weighted average of the soil quality of land in the outward buffer.

One question that remains is how to specify the width of the outward buffer for each city jurisdiction. In the main analysis, I set the buffer width to be identical for all city jurisdictions, regardless of the difference in the pace of urban land expansion across city jurisdictions during the 2000s. The specification avoids spurious correlations between the soil qualities of land inside the buffer and unobserved local productivity shocks.[11] I further specify the buffer width to be 2 km to equalize the total amount of land in the buffers and the total new urban land from 1999 to 2015 and use alternative values in robustness checks.

The second component is land ruggedness of the local region that is likely to be used for farmland development, which I use to approximate the unit cost of new farmland development.[12] In the agricultural engineering literature, the ruggedness of land is a crucial determinant for the low-cost cultivation of land (Nunn and Puga, 2012). The soil quality of land is not a constraint here because when the original soil does not provide sufficient nutrition, an easy solution and hence a common practice is to move the topsoil of the farmland that is converted into urban land onto to the newly developed farmland.[13]

Which region within a city jurisdiction is most likely to be used for new farmland development?

[9]This specification relies on the assumption that urban area expands by extending all points along the current boundary outwards by an equal distance. The results are robust to alternative assumptions on the urban area expansion path (Harari, 2018).

[10]The dependent variable is constructed using data from the City Development Yearbooks. In the baseline specification, I include the soil features that are the most critical in determining soil suitability for cultivation (Zhang et al., 2010), including organic carbon, pH, salinity, gypsum, gravel fraction, water storage capacity and soil drainage. Appendix B.4 provides a detailed discussion of the regression results and robustness checks.

[11]Suppose instead that we were to consider a wider buffer for jurisdictions with faster urban land expansion during the 2000s. If the average soil suitability for cultivation is lower for land farther away from the current urban area, there would be a positive mechanical correlation between soil quality and the increase in urban land and vice versa.

[12]In the main analysis, I choose the percentage of land grids with a slope of 15 degrees or more because the central government explicitly discourages farmland from being created on surfaces with a slope of 15 degrees or more. The results are robust to alternative definitions of land ruggedness.

[13]Since there is underground construction beneath new urban land (to build, e.g., electrical lines, water systems), the topsoil would have been removed in any case. Hence, the only additional cost here is to ship the topsoil to the new farmland. See National Land Resource Department Annual Report, 2006.

Land close to the administrative boundary of a city jurisdiction is farther away from existing or future urban area. Placing new farmland close to the administrative boundary hence reduces the likelihood that the new farmland will be converted into urban land and unused land from somewhere else will be converted into farmland in the near future. Moreover, land close to the administrative boundary is less developed and more likely to have open space for new farmland development in general. This conjecture is confirmed by the data. Between 2000 and 2010, land grids closer to the administrative boundary were more likely to shift from non-farmland to farmland.[14] Furthermore, moving 1 km farther away from the administrative boundary reduces the probability of a grid changing from non-farmland into farmland by approximately 20%, as shown in Table 13 Column 3. This indicates that land grids more than 5 km away from the administrative boundary of a city jurisdiction are very unlikely to be used for new farmland development. Therefore, I create a 5-km inward buffer from the administrative boundary and define it as the projected farmland development region.[15]

I define the land conversion barrier as the interaction of the predicted farmland density based on the soil qualities of land close to the existing urban land in 1999 and the ruggedness of land close to the administrative boundary. Appendix B.4 provides a detailed description of the steps involved in constructing the land conversion barrier. The granular data on soil quality and land ruggedness were collected during the national soil survey conducted between 1989 and 1993 (Fischer et al., 2008a). Therefore, the land conversion barrier solely depends on predetermined geographical features. The land conversion barrier is essentially an exogenous determinant of the additional cost per unit of new urban land. Ceteris paribus, the higher the barrier is, the less new urban land should be developed after policy implementation.

A series of diagnostics of the land conversion barrier show that it has no clear spatial patterns on the map (Figure 2), that the measure has rich cross-sectional variation (Figure 15), and that city jurisdictions with different land conversion barriers are balanced along various dimensions of economic and demographic characteristics in 1990 (Table 14).[16]

To identify the impact of the land conversion barrier on local economic outcomes, I use a DD estimator with continuous treatment intensity. The regression specification is the following:

$$\ln(y_{it}) = \beta C_{u,i} \times \text{Post1999} + \alpha_i + \gamma_t + \sum_{\tau \in 1991 \text{ to } 2015} X_i' \theta_\tau + \epsilon_{it}. \tag{1}$$

The dependent variable $\ln(y_{it})$ is (log of) the outcome variable of interest in city jurisdiction i in year

[14] To conduct this analysis, I use land cover and land-use database at a 500-m resolution to create a raster file on whether, at each grid, the land cover changes from non-farmland into farmland from 2001 to 2010. The raster data file comes from NASA MODIS MCD12Q2. It classifies every 500-m grid on the map as, among other types, unused land, farmland, urban land.

[15] I exclude land that is inside within the existing urban area before 1999 or inside the outward buffer of the urban boundary. The results are robust to alternative ways of drawing the projected farmland development region.

[16] Locations with different land conversion barriers are overall very similar in terms of growth of population and employment, changes in economic structure (growth of employment in the non-agricultural sector), and human capital accumulation (changes in the illiterate population and college graduates) from 1982 to 1990. They are also quite similar along a broad measure of local economic characteristics in 1990, including the employment structure, education, and in-migration. Locations with a higher land conversion barrier had slightly lower populations and employment rates in 1990. Therefore, I control for the time-varying impacts of population and the employment rate in 1990 in the main analysis.

t. $C_{u,i}$ represents the land conversion barrier, and Post1999 takes value 1 after 1998. β represents the causal impact of the land conversion barrier on the outcome since policy implementation. I control for city jurisdiction fixed effects (α_i) and year fixed effects (γ_t). X_i is a set of state variables for city jurisdiction i and includes region dummies across all regression specifications to guarantee that any cross-region differences in economic development will not bias the estimation.[17] In the central regression specification, I include the population and employment rate in 1990 in the state vector X_i because the land conversion barrier is slightly correlated with these two variables. Finally, the error term is clustered at the city jurisdiction level.

The identification relies on the assumption of parallel trends across city jurisdictions with different land conversion barriers throughout the period of study had the policy never been implemented. I have multiple years of data before the policy was adopted, which allow me to directly test the parallel trends assumption for the years up to 1998. As shown in Section 4.2, I do not find any systematically different trends in growth in terms of the major outcome variables before 1999. Next, given that the land conversion barrier is constructed using predetermined geographical features only, the identification is not contaminated by reverse causality issues.

4.2 Results: Main Outcomes

This subsection estimates the effects of the Farmland Red Line Policy on local economic outcomes.

4.2.1 Urban Land Supply

I first show that a higher land conversion barrier significantly reduces urban land supply after the policy was adopted. In the main analysis (2, Column 2), a one-standard-deviation increase in the land conversion barrier reduces urban land supply by 5.5%. The estimated effects remain stable as additional controls are added, which alleviates concerns that the results are driven by different growth paths across city jurisdictions.[18]

To provide evidence to support the parallel trends assumption, I estimate the effect year-by-year to allow the impact of the land conversion barrier on urban land supply to change over time.[19] Figure 3 confirms that city jurisdictions with different land conversion barriers have parallel trends in urban land supply before policy implementation: β_τs are not significantly different from 0 before 1998. After 1999, the negative impact of the land conversion barrier on urban land supply gradually increases until 2007.

[17] China is divided into 4 economic regions: East, Middle, West and Northeast.

[18] Table 2 Column 1 displays the regression outcome with baseline controls. Column 2 presents the central regression specification, which includes the time-varying effects of the population and employment rate in 1990, additionally. Column 3 further includes the time-varying effects of the percentage of employment in the agricultural sector, the percentage of employment in the construction sector, illiteracy rate, the percentage of college graduates, and the percentage of in-migration in 1990. In Column 4, both the controls used in Column 3 and the province time-varying effects are included.

[19] The regression specification is the following: $\ln(y_{it}) = \sum_{\tau\in 1991 \text{ to } 2015} C_{u,i}\beta_\tau + \alpha_i + \gamma_t + \sum_{\tau\in 1991 \text{ to } 2015} X_i'\theta_\tau + \epsilon_{it}$, where $\{1_{\tau=t}\}_{\tau\in 1991 \text{ to } 2015}$ is a set of dummy variables that take value 1 if the observation is from year τ. β_τ represents whether in year τ city jurisdictions with greater land conversion barriers have less urban land, conditional on any initial difference in 1990.

4.2.2 GDP

Next, a higher land conversion barrier significantly reduces the GDP of a city jurisdiction. The main estimates (Table 3 Panel B) suggest that after policy implementation, a one-standard-deviation increase in the land conversion barrier reduces GDP by 3.9% (Column 1). Furthermore, the negative impact on GDP is due to a 6.0% decrease in GDP in the secondary sector, which is dominated by manufacturing. GDP from the agricultural sector (Column 2) or the service sector (Column 5) does not change significantly. This finding is consistent with the intuition that the secondary sector is more urban land-intensive than the service sector and therefore is more negatively affected by the policy. The patterns are robust to adding alternative control variables, as shown in Table 3 Panels A, C, and D.

I also estimate the effect year-by-year to show that the parallel trends assumption holds well for GDP and GDP by sector. I plot the coefficients in Figure 4. Moreover, the estimates suggest that the Farmland Red Line Policy has substantial impacts on local GDP. By 2010, a one-standard-deviation increase in the land conversion barrier reduces the GDP of a city jurisdiction by 6.0% to 6.9%, depending on the specification of control variables. The impact is equivalent to approximately 15% of one standard deviation of the GDP growth rate across city jurisdictions from 1999 to 2010 and is hence economically significant.

4.2.3 Population and Employment

This subsection shows that a higher land conversion barrier significantly reduces the population and employment of a city jurisdiction. As reported in Table 4, a one-standard-deviation increase in the land conversion barrier reduces both population and employment by 4.5% in the main analysis. Next, I present the year-by-year estimates and plot the coefficients in Figure 5. After 1999, the negative impact of the land conversion barrier grows over time, which is consistent with the dynamics of urban land and GDP during the same period. In the main analysis, a one-standard-deviation increase in the land conversion barrier reduces the population by 5.9% by 2010, which is one-quarter of a standard deviation of population growth across city jurisdictions from 2000 to 2010.

I present additional evidence that the impact of the land conversion barrier to local employment growth is driven by a boom in non-agricultural sector employment growth. I conduct DD analysis on employment by sector and display the results in Appendix Table 36. A comparison of Column 2 and Column 3 suggests that the increase in employment is driven by an increase in the non-agricultural sector. Moreover, the employment growth in the non-agricultural sector is not due to the growth of the construction sector (Column 4). The results confirm that the observed increase in secondary sector GDP is unlikely to have been driven by a boom in the construction sector.

4.2.4 Robustness Checks

This subsection presents a series of robustness checks to address several endogeneity concerns. First, the ruggedness of land close to the existing urban area might be correlated with the soil quality there or the ruggedness of land close to the administrative boundary. The correlation would bias the estimation because the ruggedness of land close to the existing urban area can directly affect the cost of urban land development after 1999 (Saiz, 2010). To address this concern, I control for the time-varying effects of the ruggedness of the land in the outward buffer of the existing urban area. The results for urban land supply (Appendix Table 21 Column 4), GDP (Appendix Table 28 Panel C) and population (Appendix Table 34 Column 4) are quite similar to the main estimates.

Second, the ruggedness of land within a city jurisdiction might affect the cost of transportation between the city jurisdiction and other locations. Transportation costs might have time-varying effects on local production and population growth, which would bias the estimation. To address this concern, I control for the number of railway lines that pass through a city jurisdiction in 2000 – an approximation of the connectivity of the city jurisdiction with other locations – interacted with year dummies. The results for urban land supply (Appendix Table 22 Column 3), GDP (Appendix Table 29 Panel A) and population (Appendix Table 35 Column 2) are close to the main estimates.

Third, there might be concerns that the land conversion barrier is systematically correlated with local economic fundamentals in the 1990s. Locations with different economic fundamentals in the 1990s have different growth paths, and such a relation would bias the estimation. The balance test shows that city jurisdictions with different land conversion barriers are similar along various dimensions except the population and employment rate in 1990. Therefore, the main specification included the time-varying effects of the two variables. As a robustness check, I further include the time-varying effects of state variables on sector structure, human capital, and migration and find that the results are close to the main estimates.[20] This specification approximates the non-linear specific time trends.

Fourth, city jurisdictions from different provinces might have different trends in economic growth and hence be less comparable to one another. As a robustness check, I control for province time-varying effects in addition to the time-varying effects of the full set of economic characteristics. The estimated impacts on urban land supply (Table 2 Column 4), GDP (Table 3 Panel D) and population (Table 4 Panel A Column 4) are similar to those in the main analysis.

Appendix B.5 provides additional robustness checks to show that the results are not affected by potential data measurement errors, sample selection concerns, alternative specifications of the land conversion barriers, and the clustering of error terms.

4.3 Mechanism Investigation

This subsection tests several channels that might explain the causal relation between the land conversion barrier and local economic outcomes. I first find that the land conversion barrier is positively associated with urban land prices during the 2000s. This finding confirms the idea that a lower land conversion barrier benefits the local economy by reducing urban land prices. Moreover, the FAR of buildings cannot freely adjust to the land conversion barrier. Therefore, if urban land supply decreases, the effective urban space is reduced. Finally, I find that the causal relation cannot be explained by alternative channels such as the deterioration of urban compactness or poorer government service

[20] I add region dummies, population, employment rate, the share of agricultural sector employment, the illiteracy rate, the percentage of college graduates in the population, and the percentage of in-migrants in the population in the state vector X_i. Please refer to Table 2 Column 3 for the results on urban land supply, Table 3 Panel C for results on GDP, and Table 4 Panel A Column 3 for results on population.

provision in more affected jurisdictions.

First, I provide further evidence in support of the argument that a lower land conversion barrier benefits the local economy by reducing urban land prices. Urban land price data at the city jurisdiction level are available only after 2006, which makes a standard DD analysis infeasible. Therefore, I first use the economic conditions of a city jurisdiction in 1996 to predict urban land prices in 1996 ($\widehat{\ln P}_{i,1996}$).[21] I then conduct a pooled long-difference regression using specification (2),

$$\ln P_{it} - \widehat{\ln P}_{i,1996} = \beta C_{u,i} + \tau_t + \sum_{\tau=2007\, to\, 2015} X_i \theta_\tau + \epsilon_{it}, \tag{2}$$

where X_i includes both the region dummies and the full set of 1990 characteristics and τ_t represents the year fixed effects. Despite the regression being written in long-difference form, this regression specification essentially compares urban land prices between 2007 and 2015 across city jurisdictions. The identification assumption is that for city jurisdictions within the same economic region that have similar economic conditions in 1990, if they have similar levels of GDP by sector, population, and urban land supply in 1996, they would have had similar urban land prices at the time, as captured by $\widehat{\ln P}_{i,1996}$. Therefore, the difference in urban land prices between such city jurisdictions after 2006 would be explained by the land conversion barrier.[22]

I find that a one-standard-deviation decrease in the land conversion barrier is associated with a 5.8% lower urban land price overall (Table 6 Column 1). In China, urban land is sold through either auction or agreement (a private negotiation with a local government). The price of urban land sold at auction is expected to be closer to the market price. As confirmed by Column 2, a one-standard-deviation decrease in the land conversion barrier is associated with a 12.5% reduction in the price of the urban land sold at auction.

Second, I find that the FAR cannot freely respond to the land conversion barrier, and hence, floor space declines whenever urban land supply decreases. This explains why urban land supply matters even though urban workers ultimately use floor space, which depends on both the urban land supply and the FAR. As shown in Table 8, I find no association between the land conversion barrier and the FAR of newly developed commercial land or industrial land during the 2000s (Column 1 and Column 2). For residential land, the FAR of new residential land is positively correlated with the land conversion barrier (Column 3). However, after accounting for the reduction in urban land supply, a one-standard-deviation decrease in the land conversion barrier is still associated with a 4.3% increase in total residential space.[23] The increasing marginal cost of adding another floor to a building can

[21] I first regress lnP_{it} from 2007 to 2015 on GDP, GDP from the secondary sector and the tertiary sector, population, and urban land amount in the corresponding year. Next, I use the coefficients estimated in the regression to predict the urban land price in 1996.

[22] An alternative regression specification is to employ $\widehat{\ln P}_{i,1996}$ as a control variable or directly control for the time-varying effects of variables used to predict $\widehat{\ln P}_{i,1996}$. Regression results using these two alternative regression specifications are similar to the baseline results and reported in Appendix Table 37 and 38.

[23] A one-standard-deviation decrease in the land conversion barrier leads to 7.9% more urban land; 7.9% is the estimated response of urban land supply to a one-standard-deviation decrease in the land conversion barrier in 2010, averaged across regression specifications. Given that the allocation of urban land to different purposes does not significantly correlate with the land conversion barrier, the change in urban residential land supply is the same as the overall change. Next, the FAR of new residential buildings decreases by 3.4% with respect to a one-standard-deviation decrease in the land conversion barrier, according to Table 8 Column 3. Therefore, the total change in urban residential space is jointly

explain why the supply of urban space is not perfectly elastic. Furthermore, for many industrial and commercial buildings, the FAR is subject to industry-specific standards and hence cannot easily change even when there is a lack of urban land.

Third, I do not find evidence in support of the government service provision channel. This channel states that because a large fraction of the revenue from urban land sales in China goes to the local government, a lower land conversion barrier leads to higher local government revenue and better public services. Such jurisdictions attract more workers and have higher GDP and more population.

I test this channel by using a DD regression specification (1) to estimate the impact of the land conversion barrier on government expenditures per capita and the number of hospital beds. Both outcome variables approximate government public service provision and are available both before and after the policy was adopted. The majority of local government expenditure is spent on providing social security and public education. These services are rivalrous, and hence, I use government expenditures per capita instead of total government expenditures to represent government services. As shown in Table 5 Column 1, a decrease in the land conversion barrier does not increase the number of hospital beds after 1999, suggesting that health care provision does not improve after urban land expansion. Next, from Table 5 Column 2, there is no clear evidence that government spending per capita decreases with the land conversion barrier after 1999.

Fourth, I find no evidence that a higher land conversion barrier deteriorates the shape of the urban area in a city jurisdiction. Harari (2018) notes that the compactness of a city can directly improve commuting efficiency within the city and is hence valued by workers. On the one hand, the expensive farmland next to existing urban land may incentivize the local government to develop unused land farther away from existing urban land, which reduces the compactness of the urban area. On the other hand, the local government is responsible for providing public transportation infrastructure and subsidizing public transportation. This makes the local government internalize the benefit of the compactness of the urban area. If the saved cost on transportation from a more compact urban area is greater than the additional cost of using farmland, the local government would use the farmland for urban land development and bear the additional farmland development cost.

To test whether compactness deteriorates with a higher land conversion barrier, I conduct a DD analysis with the normalized remoteness measure as the dependent variable. The remoteness measure is constructed by using a 30-m resolution urban land cover database (Liu et al., 2018), and it exists for both 1995 and 2010. As displayed in Table 9 Column 1, there is no significant change in remoteness with respect to the land conversion barrier after policy implementation. This indicates that the urban compactness channel cannot explain the causal impact of the land conversion barrier on local economic development.

Finally, I find no evidence that the land conversion barrier affects how local government allocates urban land between residential use and business use. I calculate the share of urban land as residential land, business land, land for public facilities, and land for public transport and green area from 2002 to 2015. Then, I test whether the share of urban land in each category is associated with the land conversion barrier. As shown in Table 7, none of them is significantly associated with the land

determined by the change in urban residential land and the FAR of residential buildings: (1+7.9%)(1-3.4%)-1=4.3%.

conversion barrier. This analysis shows that when the amount of urban land increases, urban land for each purpose increases proportionally.

Summary This section shows that because of the Farmland Red Line Policy, city jurisdictions with a lower land conversion barrier were able to create more urban land after the policy was adopted. Ceteris paribus, urban land is cheaper in these city jurisdictions, which attracts more workers to move in during the rapid urbanization period in China. Therefore, these city jurisdictions have higher GDP and more population than the others. The findings show that the Farmland Red Line Policy has a significant impact on the spatial distribution of economic activities.

5 Model

The identified local effects of the policy on economic outcomes leave open the question of whether the policy generates any significant aggregate impacts on the economy, which matters for policy evaluation. The policy is less of a concern if it does not create any inefficiency at the aggregate level but simply causes a reallocation of economic activities from more treated to less treated locations. Because locations are interlinked through flows of trade and migration, the impact of the policy on a more treated location may generate spillover effects on less treated locations in general equilibrium. Therefore, it is challenging to infer the aggregate impacts solely based on the estimates of the local effects. A general equilibrium model can incorporate the interlinkages between locations and separate the spillover effects from the direct impacts. The model can also simulate counterfactual outcomes under alternative policies and provide informative guidance on a more efficient design of the policy, which the Chinese central government has been considering.

I develop a static quantitative spatial equilibrium model with endogenous land-use decisions to evaluate the aggregate impact of the Farmland Red Line Policy. When examined through the lens of the model, the Farmland Red Line Policy causes an under-supply of urban land and an excess supply of farmland. Moreover, the degree of cross-sector land misallocation depends on the predetermined geographical features. In general equilibrium, land misallocation leads to labor misallocation due to the labor mobility between the agriculture and the manufacturing sector and across space. Therefore, abolishing the policy would generate gains in terms of workers' welfare.

5.1 Model Setup

In the model, each location has both an urban sector and a rural sector. Each location-sector produces a variety of a final product. Consumers' love of variety means that products produced in one location are shipped to the rest of the locations, and transportation of final products is subject to an iceberg trade cost. Next, workers maximize utility by choosing the location and sector, supplying one unit of labor to earn wage income, and spending income on tradable goods and housing. Finally, in each location, there is an immobile representative landlord who owns a continuum of land plots. A land plot can be developed into farmland or urban land at a cost or remain unused land. Agricultural sector workers rent farmland for production and housing, while manufacturing sector workers rent urban land for housing. The landlord chooses the amount of urban land, farmland and unused land to maximize the land development profit. Finally, the landlord spends the land profit on tradable goods consumption.

5.1.1 Locations

I model China as consisting of N locations. Each location n has both an urban area where manufacturing production takes place and a rural area where agricultural production takes place. In each location and sector, a differentiated final good is produced. All the final goods can be traded between any two locations subject to an iceberg trade cost $T_{nn'}$.

A location n is distinguished by its productivity in the manufacturing sector and the agricultural sector $\{A_{Mn}, A_{Fn}\}$ and the amenity in the urban and the rural area $\{B_{Mn}, B_{Fn}\}$. Agglomeration effects exist in each urban sector: $A_{Mn} = \bar{A}_{Mn}(L_{Mn})^{\alpha}$, where α represents the degree of the agglomeration in production. Local amenities can respond endogenously to how populated the location is: $B_{sn} = \bar{B}_{sn}(L_{sn})^{\beta}$, $s \in \{M, F\}$.[24]

5.1.2 A Worker's Problem

The economy is inhabited by a measure $\bar{L}$ of workers. Workers are homogeneous except for their idiosyncratic preferences over locations and sectors, and they can freely choose their location and sector.[25] A worker in the manufacturing sector lives in the urban area and a worker in the agricultural sector lives in the rural area. Workers in the economy derive utility from manufacturing and agricultural goods consumption, housing consumption, and local amenities. Workers are price takers in all markets.

The utility function of worker i who lives in location n and works in sector s is:

$$U(i,n,s) = \frac{\left(C_F(i,n,s)^{\mu} C_M(i,n,s)^{1-\mu}\right)^{\theta} h(i,n,s)^{1-\theta} B_{sn} z_{i,n,s}}{(\mu\theta)^{\mu\theta}((1-\mu)\theta)^{(1-\mu)\theta}(1-\theta)^{1-\theta}}.$$

where $C_M(i,n,s)$ is a CES bundle of manufacturing goods, $C_F(i,n,s)$ is a CES bundle of agricultural goods and $h(i,n,s)$ is housing consumption (the amount of residential land). Specifically, $C_M(i,n,s) = \left(\sum_{n' \in N_c} c_{Mn'}(i,n,s)^{\frac{\sigma_M - 1}{\sigma_M}}\right)^{\frac{\sigma_M}{\sigma_M - 1}}$, where $c_{Mn'}(i,n,s)$ is the manufacturing product from location n' and σ_M is the elasticity of substitution of the manufacturing products from alternative locations; $C_F(i,n,s) = \left(\sum_{n' \in N} c_{Fn'}(i,n,s)^{\frac{\sigma_F - 1}{\sigma_F}}\right)^{\frac{\sigma_F}{\sigma_F - 1}}$, where $c_{Fn'}(i,n,s)$ is the agricultural product from location n' and σ_F is the elasticity of substitution of the manufacturing products from alternative locations. θ represents the share of spending on tradable goods. Within the tradable goods category, μ is the share of spending on agricultural products and $1-\mu$ is the share of spending on manufacturing

[24] This assumption allows local amenities to endogenously respond to the size of local population. Under this assumption, more populous regions can be either more polluted or congested (which indicates a negative β) or more attractive due to public goods sharing (which indicates a positive β).

[25] In reality, changing sectors and locations is subject to non-negligible switching costs (Tombe and Zhu, 2018). This makes the labor supply to a specific location and sector an upward-sloping curve. Although switching costs are not explicitly modeled here, workers' preference heterogeneity guarantees that the labor supply acts similarly for any real wage shocks. A high switching cost is approximated by a high preference dispersion.

products. Finally, worker i derives utility from location n and sector s. B_{sn} is the utility common to all workers, while $z_{i,n,s}$ is the idiosyncratic utility of worker i. For model tractability, I assume that each $z_{i,n,s}$ is an independent draw from a Fréchet distribution: $z_{i,n,s} \sim F_z(x) = e^{-x^{-\tilde{\nu}}}$.

Next, I introduce the budget constraints of workers in the economy. Workers have a fixed unit of working time, normalized to 1. I denote w_{sn} as the wage rate for one unit of working time in sector s and location n. A worker spends labor income on tradable goods and housing consumption. The unit price of a sector s' product from location n' is $p_{s'n'}T_{n'n}$, where $p_{s'n'}$ is the product price at origin and $T_{n'n}$ is the iceberg trade cost from n' to n. A worker in an urban area rents urban land for housing and pays the unit urban land price p_{Hn} to the landlord; a worker in the rural area rents farmland for housing and pays the unit farmland rent p_{Rn} to the landlord. Therefore, her budget constraint is

$$\sum_{s' \in \{F,M\}} \sum_{n' \in N} p_{s'n'} T_{n'n} c_{s'n'}(i,n,s) + p_{H,sn} h(i,n,s) \leq w_{sn},$$

where $p_{H,Mn} = p_{Hn}$ and $p_{H,Fn} = p_{Rn}$.

The utility maximization problem for a worker is the following. She first draws a vector of the idiosyncratic utilities she derives from different locations and sectors. Based on her realized preferences, she chooses the location and sector and earns wage income. She then allocates her income to tradable goods and housing. The worker solves the utility maximization problem through backward induction. In Stage 2, conditional on the choice of location and sector, she optimally allocates income across tradable goods and housing. In Stage 1, she chooses the location and sector combination that offers her the highest utility.

In Stage 2, a worker i who has already chosen to live at location n and work in sector s receives her maximum utility $\frac{w_{sn}}{\tilde{p}_n^{\theta} p_{H,sn}^{1-\theta}} B_{sn} z_{i,n,s}$, where $\tilde{p}_n$ is the price index of tradable goods,

$$\tilde{p}_n = \tilde{p}_{Fn}^{\mu} \tilde{p}_{Mn}^{1-\mu}. \tag{3}$$

$\tilde{p}_{sn}$ is the price index of the bundle of consumption goods from sector s chosen by consumers in location n,

$$\tilde{p}_{sn} = \left(\sum_{n' \in N} \left(T_{nn'} p_{sn'} \right)^{1-\sigma_s} \right)^{\frac{1}{1-\sigma_s}}. \tag{4}$$

In Stage 1, a worker chooses the location and sector combination that offers her the highest utility. Given that $z_{i,n,s}$ follows the Frechét distribution, the share of workers who choose location n and sector s is $\pi_{sn}^L = \frac{\left(w_{sn} \tilde{p}_n^{-\theta} p_{H,sn}^{\theta-1} B_{sn}\right)^{\tilde{\nu}}}{\sum_{s' \in \{F,M\}} \sum_{n' \in N} \left(w_{s'n'} \tilde{p}_{n'}^{-\theta} p_{H,s'n'}^{\theta-1} B_{sn}\right)^{\tilde{\nu}}}$.[26] We define $\nu = \frac{\tilde{\nu}}{1-\beta\tilde{\nu}}$ to simplify calculation. The expected utility of a worker is defined as the expected utility before her idiosyncratic preferences

[26] This equals the probability that a worker will choose location and n sector s before her idiosyncratic preferences are realized.

are realized. It can be expressed as $\tilde{V} = \bar{V}\bar{L}^{\beta}\Gamma\left(1-\tilde{\nu}^{-1}\right)^{\frac{1}{\nu}}$, and $\bar{V}$ is expressed as follows:

$$\bar{V} = \left(\sum_{s'\in\{F,M\}}\sum_{n'\in N}\left(w_{s'n'}\tilde{p}_{n'}^{-\theta}p_{H,s'n'}^{\theta-1}\bar{B}_{sn}\right)^{\nu}\right)^{\frac{1}{\nu}}. \tag{5}$$

The labor supply for the manufacturing sector and the agricultural sector in location n can be expressed as

$$L_{sn} = \bar{L}\bar{V}^{-\nu}\left(w_{sn}\tilde{p}_{n}^{-\theta}p_{H,sn}^{\theta-1}\bar{B}_{sn}\right)^{\nu}. \tag{6}$$

Next, the share of spending on the product variety from location n' and sector s' is $\pi^{C}_{n'n,s'} \equiv \frac{(p_{s'n'}T_{n'n})^{1-\sigma_{s'}}}{\tilde{p}_{s'n}^{1-\sigma_{s'}}}$. Finally, the aggregate demand for urban land in location n is:

$$(1-\theta)w_{Mn}L_{Mn} = p_{Hn}H_{n}. \tag{7}$$

The aggregate demand for farmland used for housing in location n is:

$$(1-\theta)w_{Fn}L_{Fn} = p_{Rn}R_{Hn}. \tag{8}$$

A detailed derivation of workers' utility maximization problem is provided in Appendix A.1.

5.1.3 Production

The production function of the manufacturing good variety produced in location n, Y_{Mn}, is

$$Y_{Mn} = A_{Mn}L_{Mn}. \tag{9}$$

Next, the production function of the agricultural good variety in location n, Y_{Fn}, is

$$Y_{Fn} = \bar{A}_{Fn}L_{Fn}^{\gamma}R_{Fn}^{1-\gamma}. \tag{10}$$

All product markets and labor markets are assumed to be competitive. At a given wage rate in the manufacturing sector w_{Mn} and the agricultural sector w_{Fn}, farmland price p_{Fn}, and product prices p_{Mn} and p_{Fn}, the demand for labor from the manufacturing sector in location n is

$$w_{Mn} = p_{Mn}\bar{A}_{Mn}L_{Mn}^{\alpha}, \tag{11}$$

the demand for labor from the agricultural sector in location n is

$$w_{Fn} = \gamma p_{Fn}\bar{A}_{Fn}L_{Fn}^{\gamma-1}R_{Fn}^{1-\gamma}, \tag{12}$$

and the demand for farmland is

$$p_{Rn}R_{Fn} = \frac{1-\gamma}{\gamma}w_{Fn}L_{Fn}. \tag{13}$$

5.1.4 The Landlord's Problem

At each location n, an immobile representative landlord owns land with measure $\tilde{R}_n$. A fraction, $1-\phi_n$, of the land is developable. The developable land can be divided into a continuum of land plots indexed by $l \in [0,1]$. A land plot l can be developed into either farmland or urban land at a cost of $\tilde{p}_n f(\psi_n) x_{nl}$.[27] $f(\psi_n)$ represents the constant development cost due to the overall geographical suitability for land development in location n, and x_{nl} represents the heterogeneity of land development suitability across land plots within a location.[28] For model tractability, x_{nl} is an independent draw from the Pareto distribution: $x_{nl} \sim F(X) = 1 - X^{-\frac{1}{\zeta}}$.[29]

A landlord obtains payoff p_{Rn} by providing one unit of farmland in the farmland rental market and p_{Hn} by providing one unit of urban land in the urban land rental market. Farmland is rented by agricultural sector workers for agricultural goods production and housing. Urban land is rented by manufacturing sector workers for housing. The profit from land development is obtained by the immobile landlord and spent on tradable goods.[30]

The farmland supply function and the urban land supply function can be derived by solving a landlord's profit maximization problem. First, without any policy intervention, the urban land rent equals the farmland rent.[31] Second, a landlord will develop a plot into farmland or urban land iff:

$$p_{Rn} = p_{Hn} \geq \tilde{p}_n f(\psi_n) x_{nl}.$$

Therefore, at given land rents p_{Rn}, the proportion of developable land being urban land or farmland is:

$$\frac{R_n + H_n}{(1-\phi_n)\tilde{R}_n} = 1 - \left(\frac{p_{Rn}}{\tilde{p}_n f(\psi_n)}\right)^{-\frac{1}{\zeta}}. \tag{14}$$

[27]By assuming that the costs of converting an unused land plot into farmland and into urban land are the same, the farmland density in urban land development does not affect the urban land supply function. Farmland density determines the proportion of farmland and unused land in urban land development. In the baseline model, converting unused land into urban land is equivalent to converting the same unused land into farmland and then converting it into urban land at no additional cost. Therefore, whether urban land is converted from farmland, unused land, or a combination of the two leads to the same urban land supply function. In Appendix A.8, I extend the model to make the farmland density affect the urban land development cost. The results of the quantitative analysis based on the extended model are close to those simulated using the baseline model.

[28]Heterogeneous suitability of land for development guarantees that the farmland supply function is upward sloping.

[29]Any distribution that is not bounded from above would guarantee that as the amount of developed land in a location is infinitely close to the total amount of developable land $(1-\phi_n)\tilde{R}_n$, the marginal cost of land development goes to infinity. This feature helps avoids the corner solution in which the farmland price is greater than the marginal farmland cost in equilibrium. Therefore, all the qualitative results discussed in this section are robust to any distribution that satisfies this feature. I choose the Pareto distribution for the quantitative exercise because it delivers a simple functional form of the farmland supply.

[30]This assumption makes the model incorporate general equilibrium effects from changes in the value of urban land and farmland, without introducing any mechanical externality into workers' location decisions from the local redistribution of land development profits (Monte et al., 2018). In a robustness check, I assume that a national portfolio aggregates the land development profits of the whole economy and equally redistributes them across all the workers in the economy. Please refer to A.7 for a more detailed discussion.

[31]This is because conditional on developing a land plot, a landlord always turns a land plot into the type of land that offers her the highest payoff. Therefore, she only supplies farmland if $p_{Hn} < p_{Rn}$ or urban land if $p_{Hn} > p_{Rn}$. However, the Cobb-Douglas functional form of the utility function and the production function guarantees that as the quantity of either type of land goes to zero, its price goes to infinity. As a result, the only possible outcome is that $p_{Hn} = p_{Rn}$, in which case the landlord is indifferent between developing farmland and urban land.

I denote by R_n the amount of farmland rented by agricultural sector workers and by H_n the amount of urban land rented by urban workers. Rearranging Equation (14) makes it possible to derive the farmland supply function as follows:

$$p_{Rn} = \tilde{p}_n f(\psi_n) \left(1 - \frac{R_n + H_n}{(1 - \phi_n)\tilde{R}_n}\right)^{-\zeta}. \tag{15}$$

Urban land supply is

$$p_{Hn} = p_{Rn}. \tag{16}$$

5.1.5 Land Supply Decision under the Farmland Red Line Policy

The Farmland Red Line Policy stipulates that farmland cannot be converted into urban land unless an equal amount of unused land is converted into farmland. This is equivalent to imposing a minimum farmland quantity $\bar{R}_n$ that the landlord in location n has to supply. The minimum quantity is determined by the amount of farmland that existed in location n by the time the policy was adopted. When there is a quantity requirement, the farmland supply falls into one of the following two scenarios: either the landlord's pure economic incentive makes her provide farmland that is weakly greater than the requirement $\bar{R}_n$, in which case the marginal land development cost still equals the marginal farmland rent or the landlord supplies excess farmland to meet the quantity requirement $R_n = \bar{R}_n$ and the marginal land development cost is greater than the marginal farmland rent. This complementary slackness condition is expressed as follows:

$$\tilde{p}_n f(\psi_n) \left(1 - \frac{R_n + H_n}{(1 - \phi_n)\tilde{R}_n}\right)^{-\zeta} \geq p_{Rn},\ R_n \geq \bar{R}_n,$$

$$\left(\tilde{p}_n f(\psi_n) \left(1 - \frac{R_n + H_n}{(1 - \phi_n)\tilde{R}_n}\right)^{-\zeta} - p_{Rn}\right)\left(R_n - \bar{R}_n\right) = 0. \tag{17}$$

The urban land price equals the marginal cost of land development regardless of the constraint binding condition,

$$p_{Hn} = \tilde{p}_n f(\psi_n) \left(1 - \frac{R_n + H_n}{(1 - \phi_n)\tilde{R}_n}\right)^{-\zeta}. \tag{18}$$

When the constraint is not binding, profit maximization indicates an equalization of the price of urban land, the price of farmland, and the marginal cost of land development. When the constraint is binding, more farmland is created to meet the minimum quantity requirement relative to the no-policy market equilibrium. This increases the marginal land development cost and hence decreases the urban land supply. Therefore, there is an undersupply of urban land and an oversupply of farmland compared to the no-policy market equilibrium. The constraint is binding if after policy implementation, there are demand shifts for urban land and farmland such that it is profitable to reduce farmland.

Finally, when the constraint is binding, an excess supply of farmland causes a less severe undersupply of urban land if the supply of developed land is more elastic in a location. The intuition is that when the supply of developed land is more elastic, supplying an additional unit of farmland crowds out more unused land and less urban land. As the supply elasticity of developed land depends on suitability of land for cultivation, ϕ_n, and hence varies across locations, so does the degree of cross-sector land misallocation. Appendix A.2 provides a more detailed discussion of the features of the land markets.

5.2 General Equilibrium

At the general equilibrium, all the markets clear. Therefore, the aggregate demand for the final good produced in location n and sector s from all the locations equals the total sales of that final good:

$$p_{Fn}Y_{Fn} = \mu \sum_{s'} \sum_{n'} p_{s'n'}Y_{s'n'} \frac{(p_{sn}T_{nn'})^{1-\sigma_F}}{\tilde{p}_{sn'}^{1-\sigma_F}}, \tag{19}$$

$$p_{Mn}Y_{Mn} = (1-\mu) \sum_{s'} \sum_{n'} p_{s'n'}Y_{s'n'} \frac{(p_{sn}T_{nn'})^{1-\sigma_M}}{\tilde{p}_{sn'}^{1-\sigma_M}}. \tag{20}$$

Next, labor markets for the manufacturing sector and the agricultural sector in each location also clear. Finally, the urban land market and the farmland market in each location clear. The farmland is used for both rural residential purpose and agricultural production. Therefore, at equilibrium,

$$R_{Fn} + R_{Hn} = R_n. \tag{21}$$

Formally, an equilibrium of the economy is defined as follows. Given the parameters of the model $\{\alpha, \theta, \mu, \nu, \sigma_F, \sigma_M, \gamma, \zeta\}$, total working population $\bar{L}$, vectors of exogenous location characteristics $\{\bar{A}_{Mn}, \bar{A}_{Fn}, \bar{B}_{Mn}, \bar{B}_{Fn}, \bar{R}_n, \psi_n, \phi_n, \tilde{R}_n\}_{n\in N}$ and trade costs $\{T_{nn'}\}_{n,n'\in N}$, the general equilibrium of the model when there is no Farmland Red Line Policy is referenced by 17 vectors $\{w_{Mn}, L_{Mn}, w_{Fn}, L_{Fn}, H_n, R_n, R_{Fn}, R_{Hn}, p_{Hn}, p_{Rn}, Y_{Fn}, p_{Fn}, \tilde{p}_{Fn}, Y_{Mn}, p_{Mn}, \tilde{p}_{Mn}, \tilde{p}_n\}_{n\in N}$ and one scalar $\{\bar{V}\}$. These 18 components of the equilibrium are determined by labor supply (6), labor market demand (11) and (12), manufacturing products supply (9), agricultural products supply (10), product market clearing conditions (19) and (20), tradable goods price index (4) and (3), urban land supply (16) and demand (7), farmland supply (15) and demand (10), (13) and (21), and expected utility of a representative worker (5).

When the Farmland Red Line Policy exists, the general equilibrium is referenced by the same set of vectors and scalars. The equilibrium conditions it has to satisfy are the same except for the farmland and urban land supply functions, which change to (17) and (18) respectively.

In general equilibrium, sectoral and spatial labor mobility implies that land misallocation causes labor misallocation. First, an oversupply of farmland and an undersupply of urban land leads to an oversupply of rural workers because farmland is cheaper and an undersupply of urban workers because urban land is more expensive, compared to the no-policy market equilibrium. Second, the degree of land misallocation varies across locations, which leads to the variation in labor misallocation across locations. When an undersupply of urban land occurs in productive yet highly constrained locations, workers have to reside in more affordable yet less productive locations. Finally, inefficiency due to the misallocation of land and labor is amplified through the regional trade network: trading with a less efficient location makes oneself worse off.

5.3 Model Extension

Two extensions are made to the baseline model before the model is taken to the data. First, I incorporate rural regions in China that do not belong to any city jurisdiction to relax the assumption that there is no migration or trade between the city jurisdictions and the other rural regions. After the extension, migration and trade flows between the city jurisdictions and the rural regions can also respond to the policy change. Second, I introduce a markup between the urban land price and the marginal land development cost to capture the price distortion caused by other land-use controls (Fu et al., 2019). I separately calibrate the land price distortions caused by the Farmland Red Line Policy and the other land-use restrictions. Therefore, in the counterfactual equilibrium, only the price distortion caused by the Farmland Red Line Policy is removed.

Incorporating the rest of the rural locations. Rural regions in China are regions that do not belong to any city jurisdictions. They account for 46% of the population in China and specialize in the agricultural sector: 70% of employment is in the agricultural sector, and these rural regions produce approximately half of the national agricultural GDP.[32] To allow for migration and trade flows between the city jurisdictions and the other, rural regions, I add N_r rural regions to the model. Each rural region is similar to the rural sector in a city jurisdiction: it produces one differentiated agricultural product that is traded across locations.[33] Landlords in these rural regions only produce farmland. The extension has no qualitative effect on the change in the variables of interest in the counterfactual analysis. Please refer to Section 7.2 for a detailed comparison of the counterfactual results with and without incorporating the rural regions.

Incorporating other land-use controls. The Farmland Red Line Policy is not the only land-use regulation that applies to rural-to-urban land conversion in reality. Rural-to-urban land conversion is subject to several taxes on urban land development, such as the urban maintenance and construction tax and the property tax. The central government also regulates the amount of new urban land that can be developed in city jurisdictions every year (Fu et al., 2019). These other land-use controls also create distortions in land prices, which need to be isolated from the price distortions caused by the Farmland Red Line Policy.

These additional taxes and regulations on urban land development are modeled as a location-specific markup, λ_{Hn}, on the marginal land development cost. λ_{Hn} is assumed not to change in the counterfactual equilibria. Therefore, when removing the Farmland Red Line Policy from the economy, the only urban land price distortion corrected is the part caused by the Farmland Red Line Policy. Any price distortions caused by other land-use taxes are assumed not to change. The price of urban

[32] The number is calculated using the 2010 population census. The majority of these people work in the agricultural sector. These rural regions are farther away from urban areas than the rural areas that belong to city jurisdictions.

[33] A rural region is defined as any rural area within a prefecture that does not belong to any city jurisdiction. The rural regions differ from one another in terms of productivity, amenities, and land development cost. The Farmland Red Line Policy imposes a minimum farmland quantity constraint on rural regions. In the consumers' utility function, there are N_r more varieties of agricultural products.

land now becomes

$$p_{Hn} = \lambda_{Hn}\tilde{p}_n f(\psi_n)\left(1 - \frac{R_n + H_n}{(1-\phi_n)\tilde{R}_n}\right)^{-\zeta}. \tag{22}$$

Denote the marginal farmland development cost as c_{Rn}, and the farmland supply function is

$$c_{Rn} = \tilde{p}_n f(\psi_n)\left(1 - \frac{R_n + H_n}{(1-\phi_n)\tilde{R}_n}\right)^{-\zeta}. \tag{23}$$

Any profit caused by the markup is obtained by the representative landlord. Note that when $\lambda_{Hn} = \frac{1+(1-\theta)\nu}{(1-\theta)\nu}$ for all the city jurisdictions, the model is isomorphic to one in which each representative landlord is a monopolist in the local urban land market and she takes into account that the number of urban workers is endogenous to the urban land price, as described by the manufacturing labor supply function (6). In this alternative setup, the landlords from different locations are essentially in monopolistic competition. In this case, the markup comes from the monopolistic competition of urban land market. A more detailed discussion is provided in Appendix A.6.

6 Structural Estimation

This section structurally estimates the model from Section 5. I first outline the procedure to calibrate the model's parameters related to the workers and tradable goods production. With the parameters and the observed variables, I recover the unobservable productivities, amenities and prices that rationalize the observed data as an equilibrium of the model. Next, geographical features that affect land development are constructed using geo-databases. Finally, with the imputed land prices and geographical features across locations, I calibrate the rest of parameters related to the land markets. Model diagnostics results are provided at the end of this section.

6.1 Parameter Calibration

I estimate the spending share on agricultural products, μ, to be 0.28 to match the national share of GDP from all rural areas (the rural sector in city jurisdictions and the rural regions). I set the farmland share in agriculture production, $1-\gamma$, to be 0.23 to match the farmland share in agricultural production at the national level in 2010.[34]

I allow the income share on residential land to differ between the rural and urban sectors. This is because although urban land is modeled as residential land for simplification, we interpret it as urban workers using a fraction of urban land for production purposes and the rest for residential housing. The income share on urban land is therefore specified as the sum of the structure share in manufacturing production and the labor share in manufacturing production times the urban workers' spending share on residential land. The structure share and labor share in the manufacturing production are specified as 0.156 and 0.844, following Tsivanidis (2018). I set the urban residential housing spending share in

[34] The data come from China Rural Statistical Yearbook, 2010. The labor share, farmland share and intermediate input share are 0.51,0.15 and 0.34, respectively. The farmland share is derived after a re-normalization to exclude intermediate inputs that are absent from my model. This value is close to the cost share of farmland in the US circa the 1980s (Caselli and Coleman, 2001).

China to be 0.22, as estimated in Reinbold et al. (2018). This indicates that $1-\theta_M$ equals 0.344.[35] I consider the robustness of the results to alternative values of the residential housing spending share ranging from 0.1 to 0.3 that have been used in the literature (Tombe et al., 2018). Finally, I set $1-\theta_F$ to be 0.1 to match the rural housing expenditure share in 2010 reported in the China Rural Statistical Yearbook (2010).

The values for the rest of parameters are taken from the literature.[36] First of all, I specify the degree of agglomeration effect in the manufacturing sector, α, as 0.05 in the baseline according to the review in Combes and Gobillon (2015). I provide robustness checks to the value of α in the range $\alpha \in [0.02, 0.08]$.

Next, I set the labor supply elasticity to real income, ν, to be 3 and provide robustness of the results to alternative values within the range of 2 to 4 (Tombe et al., 2018; Morten and Oliveira, 2014; Bryan and Morten, 2018; Balboni, 2019).

The last parameter is the by-sector elasticity of substitution between products from different locations. The estimates about the elasticity of substitution between products produced from different locations within the same country ranges from 4 to 9 (Ossa, 2015; Allen and Arkolakis, 2014). But a distinction between agricultural products and non-agricultural products is rarely considered. I set the elasticity of substitution between manufacturing products, σ_M, to be 7. The elasticity of substitution between agricultural products, σ_F, is chosen as 8.3, following Donaldson and Hornbeck (2016). By making σ_F greater than σ_M, the model allows agricultural products produced across locations more substitutable relative to manufacturing products. I consider the sensitivity of results to allowing the two parameters to have the same value as well as specifying alternative values in the range of 4 to 9 to σ_M and σ_F.

6.2 Recover Unobservables

To conduct counterfactual analysis, a calibration of the model to the benchmark year 2010 is needed. This requires a recovery of the values of unobserved variables that rationalize the observed data from 2010 as an equilibrium. Adapting Ahlfeldt et al. (2015) to my setup, it can be shown that given parameters $\{\alpha, \sigma_M, \sigma_F, \gamma, \nu, \theta_M, \theta_F\}$, bilateral trade costs $\{T_{nn'}\}$, and location level data about land use and economic outcome by sector $\{H_n, R_n, L_{Mn}, L_{Fn}, E_{Mn}, E_{Fn}\}$ in 2010, there exist unique values of productivity in urban and rural areas ($\bar{A}_{Mn}$ and $\bar{A}_{Fn}$) that are consistent with the data up to a normalization, which corresponds to a choice of price level.[37] Besides, there exist unique values of residential amenities ($\bar{B}_{Mn}$ and $\bar{B}_{Fn}$), which are consistent with the data up to a normalization, which corresponds to a choice of the unit in which to measure amenities. Correspondingly, the rest of unobserved prices $\{p_{Rn}, p_{Hn}, \tilde{p}_n, \tilde{p}_{Mn}, \tilde{p}_{Fn}, p_{Mn}, p_{Fn}, w_{Mn}, w_{Fn}\}$ and quantities $\{R_{Hn}, R_{Fn}, Y_{Mn}, Y_{Fn}\}$ can be uniquely determined as well.

The calibration of the model proceeds in four steps. First, I solve the set of wages $\{w_{Mn}, w_{Fn}\}$, and land prices $\{p_{Hn}, p_{Rn}\}$ that are consistent with the data according to the equilibrium condi-

[35] The income share on urban land, $1-\theta_M$, is 0.156+0.844x0.22=0.344.

[36] These parameters are best estimated using firm level data or bilateral trade/migration flow data.

[37] The bilateral trade costs are calibrated using the method adopted in Redding (2016). Details are explained in Appendix A.4.

tions. Next, I use tradable goods market clearing condition (19) and tradable goods price index (4) to recover $\{\tilde{p}_{Mn}, \tilde{p}_{Fn}, \tilde{p}_n, p_{Mn}, p_{Fn}, Y_{Mn}, Y_{Fn}\}$ that are consistent with observed GDP by sector, trade costs and elasticity of substitution between varieties of final goods. In Step 3, given $\{\tilde{p}_{Mn}, \tilde{p}_{Fn}, w_{Mn}, w_{Fn}, L_{Mn}, L_{Fn}, H_n, R_n\}$, $\{\bar{A}_{Mn}, \bar{A}_{Fn}\}$ are uniquely determined by the manufacturing and the agricultural sector labor demand functions (11) and (12). In Step 4, $\{\bar{B}_{Mn}, \bar{B}_{Fn}\}$ and $\bar{V}$ can be solved up to a constant using labor supply functions (6) and labor market clearing conditions (5). Finally, the ratio of urban land price to the marginal land development cost, λ_{Hn}, is calibrated after the farmland cost-to-rent ratio in 2010 is imputed, as introduced in Section 6.4. A more detailed explanation about the above five steps is provided in Appendix A.4.

Geographical features of a location, including the soil quality, ψ_n, the percentage of undevelopable land, ϕ_n, and the total area of land in a location, $\tilde{R}_n$, are calculated using the World Harmonized Soil Database and the shapefile of administrative boundaries in China. Second, the minimum farmland quantity constraint, $\bar{R}_n$, equals the amount of farmland in location n immediately before the Farmland Red Line Policy was implemented.[38] Finally, the minimum farmland quantity constraint is binding in a city jurisdiction if and only if the quantity of farmland in 2010 equals the minimum farmland quantity and the amount of urban land increase from 1999 to 2010. Appendix B.6 provides a more detailed description of the data construction process.

6.3 Estimation of the Price Elasticity of Unused Land

The price elasticity of unused land, ζ, governs the average supply elasticity of developed land. The higher its value, the greater the increase in the marginal cost of developed land there is for a marginal unit of land development. By log-linearizing (23), we have the following regression specification:

$$\ln c_{Rnt} = -\zeta \ln \left(1 - \frac{H_{nt} + R_{nt}}{(1-\phi_n)\tilde{R}_n}\right) + \psi_n \beta + X_n \gamma_t + \epsilon_{nt}.$$

I introduce the time dimension because data from multiple years are used for the estimation. I control for province-level time-varying effects and time-varying effects of the population, employment rate, and agricultural employment in 1990 to control for local labor costs and material costs, which matter for farmland development.

An OLS regression could generate a biased estimation of ζ because the percentage of unused developable land is affected by the endogenous amount of urban land and farmland. For example, more advanced local agricultural techniques that reduce the cost of land cultivation would make the residents develop more farmland. It would bias the estimation of ζ towards zero. Moreover, the actual ϕ_n could also be endogenous because, in historically more densely populated areas, the land surface might be flattened already to make it more suitable for living and production.

To address these endogeneity concerns, I use the percentage of undevelopable land in the area close to the administrative boundary, $\hat{\phi}_n$, to predict the actual percentage of undevelopable land. As discussed in Section 4, land close to the administrative boundary is more likely to be undeveloped and

[38] The nearest year for which the farmland quantity is available for all the locations is 2000. Therefore, I use the amount of farmland in 2000 to approximate the corresponding value in 1999. A more detailed discussion of concerns related to measurement error and how to address them are provided in Appendix B.6.

hence preserve its natural features. Next, I use $\left(1-\hat{\phi}_n\right)\tilde{R}_n$, the predicted amount of developable land in location n, to instrument $\ln\left(1-\frac{H_{nt}+R_{nt}}{(1-\phi_n)\tilde{R}_n}\right)$.[39]

The relevance condition is likely to be satisfied because for two locations from the same province that have similar levels of population, agricultural sector employment and soil quality, they tend to have similar amounts of developed land in use. Therefore, a higher stock of developable land within the administrative boundary leads to a greater amount of unused developable land, hence a higher percentage of unused developable land. The exclusion restriction assumption is that conditional on population, agricultural sector employment, soil quality, and being in the same province, the instrument is uncorrelated with unobserved factors that affect the marginal farmland development cost.

Data on the unit cost of farmland development are very difficult to obtain in general.[40] Fortunately, between 1999 and 2004, the China Land Resource Yearbooks published the local government compensation per unit of new farmland to meet the policy requirement. The compensation goes to the rural households who develop new farmland.[41] Therefore, the unit cost of farmland development equals the local government compensation per unit of farmland development plus the farmland rent.[42]

Table 10 Panel A reports baseline regression results, and Appendix Table 40 shows the rest of the robustness checks discussed in this subsection. Across all regression specifications, the estimated value of ζ is between 1.25 and 2.52. The baseline specification suggests that the price elasticity of farmland supply is 1.78, which indicates a 5.6% increase in farmland supply with respect to a 10% increase in farmland price (Column 1). The corresponding OLS estimate (reported in Column 5) is biased toward zero and not significant. This is explained by the endogeneity issue discussed earlier and attenuation bias due to measurement errors in the independent variable.[43] A more detailed discussion of the alternative regression specifications and results is provided in the footnote of Table 10.

For the counterfactual analysis in Sections 7 and 8, I set the baseline value of ζ to be 1.78 and show that the counterfactual results are not sensitive to ζ within the range of 1.25 to 2.52.

[39] Note that I rely on cross-sectional variation to identify ζ because government spending on farmland development, the variable used to impute farmland development costs at the local level, is only available from 1999 to 2004. During this period, the farmland market was already distorted, and demand shocks to farmland cannot be used to estimate supply-side parameters.

[40] Given that the farmland market has been distorted since 1999, one cannot infer the marginal cost from farmland rents in locations that are binding on the minimum farmland quantity constraint.

[41] The local government surveyed villages in rural areas to learn how much it costs to develop a unit of new farmland.

[42] The farmland rent is imputed using the data on agricultural output value and farmland amount in the corresponding year. Both the marginal farmland development cost and government compensation per unit of farmland need to be amortized to make them comparable with farmland rent in a static model. Suppose that the depreciation rate of farmland is zero and the actual marginal development cost is C_{Rn}; it must be that $C_{Rn}=C_{Fn}+\sum_{t=0}^{\infty}\frac{p_{Rn}}{(1+r)^t}$, where C_{Fn} is the actual government payment per unit of farmland and r is the interest rate. I denote $\delta=\sum_{t=0}^{\infty}\frac{1}{(1+r)^t}$, and the above equation is expressed as $\frac{C_{Rn}}{\delta}=\frac{C_{Fn}}{\delta}+p_{Rn}$. The amortized government payment per unit of farmland, c_{Fn}, is simply $\frac{C_{Fn}}{\delta}$, while the amortized farmland development cost, c_{Rn}, is $\frac{C_{Rn}}{\delta}$. In the baseline, I specify the interest rate to be 0.058, which was the average lending interest rate in China from 2000 to 2015. As a robustness check, I set the interest rate to be 0.070, which is the median interest rate faced by rural households calculated from China Household Finance Survey. This result is reported in Appendix Table 40.

[43] The farmland data from 2001 to 2004 and the urban land for smaller cities from 1999 to 2001 are missing and were hence imputed using data from other years. A detailed explanation is provided in the footnote of Table 10.

6.4 Impute the Cost-to-Rent Ratio of Farmland in 2010

The cost-to-rent ratio of farmland, denoted as τ_n, is defined as the ratio of the marginal cost of farmland to the farmland rent. A higher τ_n indicates more severe cross-sector land misallocation in a location. It is greater than 1 if the minimum quantity constraint of farmland is binding and equals 1 if it is not binding. Moreover, conditional on the constraint binding, a larger τ_n means greater compensation is needed to induce additional farmland development to meet the farmland quantity requirement and hence more severe cross-sector land misallocation. In general equilibrium, a higher τ_ns indicates greater distortion in the land market, which leads to more labor misallocation and a greater loss of workers' welfare from the policy. Therefore, the value of τ_ns is critical in determining the aggregate welfare loss of workers due to the Farmland Red Line Policy.

The value of τ_n for 2010 ($\tau_{n,2010}$ hereafter) is not observable because the cost of farmland development can only be constructed using the China Land Resource Yearbooks for years up to 2004. In the baseline specification, I assume that $\tau_{n,2010}$ is the same as $\tau_{n,2004}$.[44] It is likely that $\tau_{n,2010}$ is greater than the corresponding value in 2004, given that the urbanization rate continued to rise between 2005 and 2010. Therefore, imputation based on this assumption would provide a lower bound on the aggregate effect of the policy.

As a robustness check, an alternative way of imputing $\tau_{n,2010}$ is to infer the change in the farmland development cost from 2004 to 2010 based on the change in unused developable land during this period. According to the farmland supply function (23), $c_{Rn,2010}$ can be recovered through the following relation: $\frac{c_{Rn,2010}}{c_{Rn,2004}} = \frac{\tilde{p}_{n,2010}}{\tilde{p}_{n,2004}}\left(\frac{(1-\phi_n)\tilde{R}_n - R_{n,2010} - H_{n,2010}}{(1-\phi_n)\tilde{R}_n - R_{n,2004} - H_{n,2004}}\right)^{-\zeta}$. I use the national average change in the farmland development cost from 2004 to 2011 to approximate $\frac{\tilde{p}_{n,2010}}{\tilde{p}_{n,2004}}$.[45] After $c_{Rn,2010}$ is imputed, $\tau_{n,2010}$ can also be imputed.[46] The correlation between $\tau_{n,2010}$ imputed using the baseline version and this alternative version is 0.61. In Section 7.2, I show that this alternative imputation method generates slightly larger GDP and welfare gains in the counterfactual equilibrium without the policy.

6.5 Model Fit

Before proceeding to the counterfactual analysis, I show that the model is a good approximation of the real economy in two steps. In Step 1, I use the calibrated model structure to simulate the reduced-form effects of the Farmland Red Line Policy and compare these predictions to the realized reduced-form results from Section 4. The simulated results are quantitatively similar to the reduced-form results,

[44] The value of $\tau_{n,2004}$ is calculated using the marginal farmland development cost and the farmland rent constructed in Section 6.3. For the main analysis, I use the baseline specification of farmland development cost in 2004 explained in Section 6.3 to calculate $\tau_{n,2004}$ and impute $\tau_{n,2010}$. This choice returns the lowest mean of $\tau_{n,2010}$, hence generating the most conservative estimate of the aggregate cost of the policy. Please refer to Table 47 for a detailed discussion of the baseline specification. I also provide alternative ways to calculate $\tau_{n,2004}$ as robustness checks.

[45] The national average government compensation per unit of farmland is available for 2011. After adjusting for inflation and amortization, the national average government compensation per unit of farmland is approximately the difference between the national farmland development cost $c_R^{k,2010}$ and the national average farmland rent $p_R^{k,2010}$. Therefore, $c_R^{k,2010} = c_F^{k,2010} + p_R^{k,2010}$. $c_R^{k,2004}$ can be calculated in the same way. The national growth rate of the farmland development cost is $\frac{c_R^{k,2010}}{c_R^{k,2004}}$.

[46] τ_n is adjusted to 1 if it is less than 1. This is caused by the measurement errors in the imputed farmland development cost in 2010.

which suggests that the model captures the cross-sectional variation well. In Step 2, I show that the recovered unobservables such as productivity, amenities, and land prices are closely correlated with proxy variables not used in the calibration.

6.5.1 Structural Simulation of the Reduced Form Results

This subsection simulates the impacts of the Farmland Red Line Policy on each location and compares these predictions to the realized reduced-form results from Section 4. None of the realized reduced-form results from Section 4 are explicitly targeted when calibrating the model. Therefore, if the simulated results are close to the realized reduced-form results, this suggests that the model is a good approximation of the real economy both qualitatively and quantitatively.

To conduct the exercise, I simulate the counterfactual equilibrium following the removal of the Farmland Red Line Policy. $d \ln y_n$ is defined as the difference between the realized outcome and the counterfactual outcome for city jurisdiction n, where y is the variable of interest, including urban land supply, GDP by sector and population. $d \ln y_n$ hence represents the simulated impact of the Farmland Red Line Policy on outcome y in location n.

Next, I regress $d \ln y_n$ against the land conversion barrier constructed in Section 4 and compare the simulated results with the realized results. As the baseline model only accounts for the variation in farmland development cost, I adjust the measure of the land conversion barrier by replacing the location-specific farmland share with the sample mean. This mutes the variation in the farmland share in the land conversion barrier. I re-estimate the DD regression by using the adjusted land conversion barrier, $\widetilde{c_{un}}$, as the independent variable. The results are shown in Table 11 Panel A. I also estimate the long-difference regression only using observations from years 1996 and 2010 and report the results in Panel B.[47] The negative impact of land conversion barrier is smaller in the short run than in the long run. Therefore, the coefficients estimated only using data from 1996 and 2010 are greater than using the full set of data. Finally, in Panel C, I regress the simulated change in the variables of interest against the adjusted land conversion barrier. All the regressions include the full set of controls.[48]

Comparing Table 11 Panel B with Panel A reveals that the simulated impacts of the land conversion barrier on urban land supply, GDP, and population are close to the estimates based on the realized data. From Panel B Column 1, a one-standard-deviation decrease in the land conversion barrier increases urban land supply by 4.9%, which is quite close to the estimates using the realized data (Panel A Column 1). Next, based on the simulation, a one-standard-deviation decrease in the land conversion barrier leads to a 1.4% decrease in GDP, which is driven by a 2.0% decrease in GDP from the non-agricultural sector and a 0.8% decrease in population. These estimates are of the same order as the estimates based on the realized data.[49]

[47] Urban land data for smaller cities are not available for 1997 and 1998. Hence I specify the benchmark year as 1996, the last year for which data across all city jurisdictions are available before the policy began.

[48] The results are robust to only controlling for region time-varying effects.

[49] The fact that the coefficients from estimation using the simulated data are slightly smaller than those from the realized data suggests that the baseline quantitative model is, if anything, too conservative in predicting the impacts of the policy on economic development. As shown in Appendix Table 41, if the labor supply elasticity is specified to be 6, which is towards the high end of the estimates of this parameter in the literature, the simulated impacts on GDP and population would be very close to the estimates using the realized data. The aggregate gain in real GDP and welfare from removing the policy is larger here than in the baseline case.

6.5.2 Out-of-Sample Test

This subsection evaluates the fit of the model by testing the correlation between recovered unobservables with proxy variables not used in the calibration. I find that the productivity of the manufacturing sector is strongly correlated with foreign direct investment (FDI thereafter) per worker, the percentage of college graduates and the average years of education in the population. Next, the amenities in the urban sector are positively correlated with the number of theaters and the number of books collected by the public library. Third, the farmland price calibrated using the model is close to its counterpart calculated using rural household survey data. Finally, the urban land price recovered from the model is positively correlated with its counterpart calculated based on urban land transaction data. These results suggest that the quantitative model, albeit simplified from reality, provides a good approximation of the real world. These findings lend additional confidence to the simulated counterfactual results. A detailed discussion of the analysis and findings is provided in Appendix B.7.

7 The Aggregate Effects of the Farmland Red Line Policy

This section evaluates the aggregate effects of the Farmland Red Line Policy by simulating its removal from the economy and comparing the simulated counterfactual equilibrium with reality in 2010.[50] Without the Farmland Red Line Policy, the welfare of workers would have increased by 6%. Moreover, distortions from the policy on urbanization manifest mostly in the over-congestion of urban sectors as opposed to less urbanization. The counterfactual results are robust to various parameter values in the range discussed in the literature and alternative model specifications.

7.1 Baseline Results

My quantitative model first produces an estimate of worker welfare loss of 5.78% from the Farmland Red Line Policy (Figure 6). The estimate is derived by comparing the simulated counterfactual equilibrium without the policy and the reality in 2010. The welfare loss comes from the misallocation of both land and labor. The estimates are of the same order as the estimated welfare gain if US were to adopt optimal zoning regulation (Bunten, 2017) or use federal policies to weaken incentives to regulate the housing supply (Parkhomenko, 2018). Next, the economy would have specialized more in the manufacturing sector in the no-policy counterfactual equilibrium. Specifically, manufacturing output would have been 4.95% higher, while agricultural output would have been 2.79% lower.

One important question is how the policy intervenes in the urbanization process between 1999 and 2010, given that the policy was adopted when rural-to-urban migration accelerated. The policy-induced undersupply of urban land would both make urban areas more congested and slow urbanization. A quantitative exercise shows that distortions from the policy manifest mostly in overcrowding in the urban areas as opposed to less urbanization. Without the policy, the urban population would have been 5.2% higher in 2010. This is not economically large when compared to the increase in the urban population in reality from 1999 to 2010, which was more than 40%. In contrast, without the

[50] A detailed explanation of how the counterfactual equilibrium is solved is provided in Model Appendix A.5.1.

policy, there would have been 40% more urban land in 2010. This indicates that urban population density would have dropped dramatically, by 25%, decreasing from 12,170 to 9,249 per sq. km.[51]

To see why urban population growth is small relative to the urban land increase, I decompose the change in the urban population into a weighted average of the urban population increase across locations in the counterfactual equilibrium:[52]

$$d\ln L_M = \underbrace{-1.5 * d\ln \bar{V}}_{-8.3\%} + \underbrace{0.5 * \sum_n \pi_n d\ln H_n}_{13.8\%} + \underbrace{1.0 * \sum_n \pi_n d\ln \tilde{w}_{Mn}}_{-1.9\%} + \underbrace{\varepsilon_{Jensen}}_{1.6\%}. \tag{24}$$

$d\ln \bar{V}$ is the change in welfare, $d\ln H_n$ is the change in urban land supply in location n, and $\tilde{w}_{Mn}$ is the real wage in terms of consumption goods. The main negative driver $(-1.5 * d\ln \bar{V})$ is that in the counterfactual scenario, an overall improvement in other locations makes a location with no change in urban land supply or real wage relatively less attractive. Urban population in such a location would move to other locations with more urban land supply, causing a decline in urban population density in this location. Note that the increase in urban land supply is heterogeneous across locations and therefore leads to a smaller urban population increase than if there were a uniform increase in urban land by 39.58% (which would increase the second term in (24) to 19.79%).

Another important question is to what extent the policy protects farmland and food security. I find that without the policy, the agricultural output would have been 2.79% lower and the farmland would have declined by 6.66% compared to the reality in 2010. I also find that the main reason for the reduction in agricultural output is because of agricultural sector workers switching to the manufacturing sector. To see this, I simulate an alternative counterfactual equilibrium in which the Farmland Red Line Policy did not exist, but workers cannot switch locations or sectors.[53] In this alternative counterfactual senario, farmland would have decreased by 6.81%, which is similar to the finding in the baseline. However, GDP in the agricultural sector decreases by only 0.97%.[54]

Finally, there is substantial spatial relocation of the urban population across city jurisdictions; 40.0% of the city jurisdictions would have lost more than 5% of their urban population, while 36.5% of the city jurisdictions would have experienced an increase in urban population of at least 5%.

7.2 Robustness and Model Extensions

I first demonstrate the robustness of the quantitative results to alternative parameter values and different ways to impute the cost-to-rent ratio. Appendix Table 46 shows the robustness of the quantitative results to alternative parameter values. The increase in national output in the counterfactual equilibrium is always between 1.96% and 3.15% and the increase in workers' welfare is between 3.85% and 7.09% across specifications. Next, as shown in Appendix Table 47, alternative ways to impute the

[51]For comparison, the population density in New York City in 2017 was 10,947 per sq. km.

[52]The equation is derived by plugging the function of manufacturing labor supply change, $d\ln L_{Mn} = (1-\theta_M)\kappa d\ln H_n + \theta_M \kappa \ln \tilde{w}_{Mn} - \kappa d\ln \bar{V}$, into the decomposition of total manufacturing labor change: $d\ln L_M = \sum_n \pi_n dln L_{Mn} + \varepsilon_{Jensen}$. $\kappa \equiv \frac{\nu}{1+\nu(1-\theta_M)}$.

[53]This is modeled as specifying the wage elasticity of labor supply to be very close to 0 in the counterfactual equilibrium.

[54]Heterogeneous reductions in farmland across locations mean that the decrease in agricultural GDP is smaller than when the reduction in farmland is uniform, which would be 6.66%×0.23=1.53%.

cost-to-rent ratio generate similar quantitative results. Moreover, the baseline specification yields the most conservative estimate of the increase in workers' welfare in the counterfactual.

Next, I extend the model by making the increase in urban land in the counterfactual equilibrium increase with the percentage of urban land converted from farmland. This is done by revising two assumptions in the baseline specification. First, the cost to convert unused land into urban land is set to be smaller than that of converting unused land into farmland. Second, in location n, the landlord has to use λ_{Fn} units of farmland and $1-\lambda_{Fn}$ units of unused land to produce a unit of urban land. As shown in Appendix Table 48 Columns 3 to 5, in the extended model, the estimated welfare gain for workers in the counterfactual equilibrium is in the range of 5.43% to 6.02%. The baseline estimates fall in this range. A detailed description of the land markets in this extended model is provided in Appendix A.8.

Third, I compare the difference in the results when not including the rural regions that do not belong to any city jurisdictions. As shown in Appendix Table 48 Column 2, the change in variables yields qualitatively similar results to the baseline. By dropping the rural regions, the trade and migration flow between the city jurisdictions and the rural regions cannot adjust to the policy change. The first consequence is that more urban land would increase the utility of workers who already reside in the city jurisdictions but not the total population across city jurisdictions. Therefore, the welfare increase is higher, but the urban population increase is smaller. The second consequence is that city jurisdictions cannot shift agricultural production to the rural regions and therefore cannot specialize in manufacturing production as much as in the baseline. This can be seen in the smaller increase in manufacturing output relative to the baseline.

Finally, instead of assuming that the immobile landlords spend the land development profit, I assume that the land development profits are collected in a national portfolio and equally redistributed across workers. As shown in Appendix Table 48 Column 6, the changes in the main variables are quantitatively similar to those in the baseline outcome.

8 Policy Counterfactuals

In 2018, the Chinese central government announced that a cap-and-trade platform for farmland creation is under development (Notice of the General Office of the State Council [2018] No.16). On this trading platform, one location can pay another to create new farmland if the former location needs to convert farmland into urban land. This cap-and-trade platform guarantees that the nationwide amount of farmland does not decrease, while in each individual location, the quantity of farmland can change freely. This section evaluates the aggregate production and welfare change if a cap-and-trade platform had been launched instead.

Next, the central government plans to charge Beijing and Shanghai 4 times the listed price on the trading platform; the other locations in the more developed regions have to pay 1.6 to 3 times the price. This design essentially restricts urban land expansion and hence slows urbanization in more developed regions. In Subsection 7.2, I evaluate the welfare change if a national trading platform with price differentiation had been in use instead.

8.1 Introduce a Cap-and-Trade Platform

This subsection evaluates the change in GDP and workers' welfare if a cap-and-trade platform had been used to prevent decrease in farmland at the national level. I first discuss how the cap-and-trade platform changes the farmland supply decision and the rest of the equilibrium conditions. I then compare the simulated counterfactual outcome with reality with the counterfactual outcome when there is no farmland quantity target.

When there is a cap-and-trade platform, the landlords receive additional payoff c_F from the trading platform for each unit of farmland above the minimum quantity $\bar{R}_n$. In contrast, if a landlord develops less farmland than the minimum quantity $\bar{R}_n$, she pays the platform c_F for each unit of shortage. With this cap-and-trade platform, the supply function of farmland becomes:

$$c_{Rn} = p_{Rn} + c_F, \tag{25}$$

where c_F is the additional payoff from (payment to) the trading platform for each unit of farmland above (below) the minimum quantity. As long as c_F is positive, there is an over-investment in farmland in all locations. However, the cap-and-trade platform reduces the aggregate cost compared to requiring each location to meet the minimum farmland quantity requirement. This is because the degree of distortion, represented by the gap between the marginal farmland development cost and farmland rent, is equalized across locations.

c_F is the new endogenous variable in this equilibrium. In equilibrium, c_F guarantees that

$$\sum_n R_n \geq \sum_n \bar{R}_n, c_F \geq 0,$$

$$\left(\sum_n R_n - \sum_n \bar{R}_n \right) c_F = 0. \tag{26}$$

Next, there are payments among landlords across locations and trade balance no longer holds. If a location has more farmland than the minimum requirement, the landlord receives an additional payment of $(R_n - \bar{R}_n)c_F$ in total. Therefore, the product market clearing condition becomes

$$p_{sn}Y_{sn} = \sum_{n'} \frac{(p_{sn}T_{nn'})^{1-\sigma_s}}{\tilde{p}_{sn'}^{1-\sigma_s}} p_{Mn'}Y_{Mn'} + \sum_{n'} \frac{(p_{sn}T_{nn'})^{1-\sigma_s}}{\tilde{p}_{sn'}^{1-\sigma_s}} \left(p_{Fn'}Y_{Fn'} + (R_{n'} - \bar{R}_{n'})c_F \right). \tag{27}$$

This new equilibrium is solved through iteration. Appendix A.5.2 explains the algorithm in detail.

The simulation indicates that approximately 60% of the aggregate cost incurred when implementing the Farmland Red Line Policy could have been saved if this cap-and-trade platform had been in use. As shown in Figure 7, in the counterfactual equilibrium with a cap-and-trade platform, the welfare of workers would have been 3.54% higher than in reality. Moreover, the manufacturing output would have grown by 3.00% and the agricultural output would have been 1.65% lower. Finally, the urban population would have increased by 2.95%, and the urban land would have increased by 21.74%. The total amount of farmland would have barely changed because the minimum farmland quantity constraint still holds at national level.

8.2 Introduce a Cap-and-Trade Platform with Price Differentiation

The central government is considering adopting a cap-and-trade platform and charging more developed locations 1.6 to 4 times the price. When more developed locations such as Beijing and Shanghai pay a higher price for each unit of farmland below the quantity requirement, the urban land expansion will be more tilted towards less developed locations. This creates land mis-allocation and labor mis-allocation since less productive places become more affordable and even more workers choose to reside in the less productive places. The modeling of this alternative cap-and-trade platform is similar to the procedure described in the previous section. A detailed explanation of the farmland development decisions in this setting is provided in Model Appendix A.5.3.

The simulation indicates that approximately 35% of the aggregate cost incurred when implementing the Farmland Red Line Policy could have been saved if this alternative cap-and-trade platform had been in use. As shown in Figure 8, compared to reality, the welfare of workers would have been 2.53% higher. Moreover, manufacturing output would have grown by 1.75%, and agricultural output would have been 1.47% lower. Finally, the urban population would have increased by 2.23%, and urban land would have increased by 17.05%.

9 Conclusion

This paper used China's Farmland Red Line Policy as a natural experiment to study the impact of land-use regulation on local economic development and the aggregate welfare of workers. At the local level, city jurisdictions with a lower barrier to rural-to-urban land conversion due to the policy had significantly more urban land supply, higher GDP, and more population after the policy was adopted. At the aggregate level, the policy reduced worker welfare by 6% and generated a sub-optimal spatial distribution of economic activities. Moreover, distortions from the policy on urbanization manifest mostly in the over-congestion of urban sectors as opposed to less urbanization. Finally, a cap-and-trade platform that allows local regions to exchange farmland preservation requirements can achieve the same aggregate level of farmland, while costs 60% less of workers' welfare compared to the Farmland Red Line Policy.

Global food security is threatened by the increasing demand for food, the slow growth of crop yields, climate change, and land erosion and salinization in the 21st century (Ray et al., 2013; FAO, 2015). The international policy agenda has recognized farmland preservation policies as a potential solution to maintaining food security (Skog and Steinnes, 2016). But such policies do not come for free. My study demonstrates that land-use regulations motivated by food security concerns can generate a substantial cost to aggregate welfare in developing countries during urbanization.

More generally, my results matter in understanding the process of urbanization in China, which was one of the most important economic phenomena in the world at the turn of the 21st century. Urbanization and urban land development have been subject to strong regulations. Given that another 350 million Chinese are expected to move to urban sectors in the next two decades, the welfare gains from urbanization in China will depend substantially on the efficiency of these regulations (Woetzel et al., 2009). My paper highlights the large costs arising from the Farmland Red Line Policy. Therefore, the social value of preserving a sufficient amount of farmland to maintain food security should be higher than its substantial cost to make the policy desirable.

My results also shed light on land-use regulation in other countries and motivated by alternative reasons that constrain urban growth. Research and policy discussions typically focus on the benefit sides of the land-use regulations but neglect the cost sides. As my paper demonstrates, the cost of the land-use regulations arises from inefficient land allocation and can be economically significant. Two features of the context make the welfare loss from the Farmland Red Line Policy exceptionally high. The first is that China adopted the policy at a time when there was a high demand for rural-to-urban land conversion. The second is that it is unlikely for workers to reside in a city jurisdiction with sufficient land while enjoying the productivity of another city jurisdiction. To the extent that similar features apply to a land-use regulation, the efficiency cost is likely to be high and warrant a thorough examination.

Finally, my results emphasize the crucial role of governments in designing efficient policies in developing countries. Externalities and frictions in the markets create room for policy interventions to improve the efficacy and efficiency of the public and private sectors. This paper reveals that a cap-and-trade platform to preserve farmland can save more than half of the costs of workers' welfare from the Farmland Red Line Policy. It highlights that a market mechanism can provide a more cost-efficient solution to achieve a policy goal.

References

Gabriel M Ahlfeldt, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf. The Economics of Density: Evidence from the Berlin Wall. *Econometrica*, 83(6):2127–2189, 2015.

Treb Allen and Costas Arkolakis. Trade and the Topography of the Spatial Economy. *The Quarterly Journal of Economics*, 129(3):1085–1140, 2014.

Treb Allen, Costas Arkolakis, and Xiangliang Li. Optimal City Structure. *Yale University, mimeograph*, 2015.

Clare Alexandra Balboni. *In harm's way? Infrastructure investments and the persistence of coastal cities*. PhD thesis, The London School of Economics and Political Science (LSE), 2019.

Abhijit V Banerjee and Esther Duflo. Growth theory through the lens of development economics. *Handbook of economic growth*, 1:473–552, 2005.

Elizabeth Brown A. Who will feed china? a book review by elizabeth a. brown. *THE SOCIAL CONTRACT*, V(4):310, 1995.

Gharad Bryan and Melanie Morten. The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia. *Journal of Political Economy*, 2018.

Devin Bunten. Is the Rent Too High? Aggregate Implications of Local Land-Use Regulation. 2017.

Francesco Caselli and Wilbur John Coleman. Cross-country technology diffusion: The case of computers. *American Economic Review*, 91(2):328–335, 2001.

Jie Chen. Rapid urbanization in china: A real challenge to soil protection and food security. *Catena*, 69(1):1–15, 2007.

Ting Chen and JK-S Kung. Do Land Revenue Windfalls Create a Political Resource Curse? Evidence from China. *Journal of Development Economics*, 123:86–106, 2016.

Ying Chen, J Vernon Henderson, and Wei Cai. Political Favoritism in China's Capital Markets and its Effect on City Sizes. *Journal of Urban Economics*, 98:69–87, 2017.

Pierre-Philippe Combes and Laurent Gobillon. The Empirics of Agglomeration Economies. In *Handbook of regional and urban economics*, volume 5, pages 247–348. Elsevier, 2015.

Klaus Desmet and Esteban Rossi-Hansberg. Urban accounting and welfare. *American Economic Review*, 103(6):2296–2327, 2013.

Dave Donaldson and Richard Hornbeck. Railroads and american economic growth: A "market access" approach. *The Quarterly Journal of Economics*, 131(2):799–858, 2016.

Gilles Duranton and Diego Puga. Urban Land Use. In *Handbook of regional and urban economics*, volume 5, pages 467–560. Elsevier, 2015.

Benjamin Faber and Cecile Gaubert. Tourism and economic development: evidence from mexico's coastline. *American Economic Review*, 109(6):2245–93, 2019.

ITPS FAO. Status of the world's soil resources (swsr)–main report. *Food and agriculture organization of the United Nations and intergovernmental technical panel on soils, Rome, Italy*, 650, 2015.

G. Fischer, F. Nachtergaele, S. Prieler, H.T. van Velthuizen, L. Verelst, and D. Wiberg. Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008). Technical report, 2008a.

Guenther Fischer, Freddy Nachtergaele, Sylvia Prieler, HT Van Velthuizen, Luc Verelst, and David Wiberg. Global agro-ecological zones assessment for agriculture (gaez 2008). *IIASA, Laxenburg, Austria and FAO, Rome, Italy*, 10, 2008b.

Shihe Fu, Xiaocong Xu, and Junfu Zhang. Land conversion and misallocation across cities in china. *Available at SSRN 3342551*, 2019.

Douglas Gollin, Stephen Parente, and Richard Rogerson. The role of agriculture in development. *American economic review*, 92(2):160–164, 2002.

Douglas Gollin, Remi Jedwab, and Dietrich Vollrath. Urbanization with and without industrialization. *Journal of Economic Growth*, 21(1):35–70, 2016.

Mariaflavia Harari. Cities in Bad Shape: Urban Geometry in India. *February. http://real. wharton. upenn. edu/˜ harari/Harari_Papers/CityShapeHarariMarch2016_updated. pdf*, 2018.

Jonathan E Haskel, Sonia C Pereira, and Matthew J Slaughter. Does inward foreign direct investment boost the productivity of domestic firms? *The review of economics and statistics*, 89(3):482–496, 2007.

Chang-Tai Hsieh and Enrico Moretti. Housing Constraints and Spatial Misallocation. *American Economic Journal: Macroeconomics*, 2018.

David Lagakos, Ahmed Mushfiq Mobarak, and Michael E Waugh. The welfare effects of encouraging rural-urban migration. 2018.

Gary D Libecap and Dean Lueck. The Demarcation of Land and the Role of Coordinating Property Institutions. *Journal of Political Economy*, 119(3):426–467, 2011.

Erik Lichtenberg and Chengri Ding. Assessing Farmland Protection Policy in China. *Land use policy*, 25(1):59–68, 2008.

Xiaoping Liu, Guohua Hu, Yimin Chen, Xia Li, Xiaocong Xu, Shaoying Li, Fengsong Pei, and Shaojian Wang. High-resolution multi-temporal mapping of global urban land using landsat images based on the google earth engine platform. *Remote sensing of environment*, 209:227–239, 2018.

Zhaoshun Liu, Deming Yang, Dongyan Wang, and Xikui Wang. Running mechanism for balance of cultivated land's occupation and compensation in different region in china. *Research of Agricultural Modernization*, 26(5):386–389, 2005.

Samuel Marden et al. The Agricultural Roots of Industrial Development:'Forward Linkages' in Reform Era China. 2015.

Christopher J Mayer and C Tsuriel Somerville. Land Use Regulation and New Construction. *Regional Science and Urban Economics*, 30(6):639–662, 2000.

Ferdinando Monte, Stephen J Redding, and Esteban Rossi-Hansberg. Commuting, migration, and local employment elasticities. *American Economic Review*, 108(12):3855–90, 2018.

Melanie Morten and Jaqueline Oliveira. Migration, Roads and Labor Market Integration: Evidence from a Planned Capital City. *Unpublished Manuscript*, 2014.

Freddy Nachtergaele, Harrij van Velthuizen, Luc Verelst, NH Batjes, Koos Dijkshoorn, VWP van Engelen, Guenther Fischer, Arwyn Jones, and L Montanarela. The harmonized world soil database. In *Proceedings of the 19th World Congress of Soil Science, Soil Solutions for a Changing World, Brisbane, Australia, 1-6 August 2010*, pages 34–37, 2010.

Arthur C Nelson. Preserving prime farmland in the face of urbanization: lessons from oregon. *Journal of the American Planning Association*, 58(4):467–488, 1992.

Nathan Nunn and Diego Puga. Ruggedness: The Blessing of Bad Geography in Africa. *Review of Economics and Statistics*, 94(1):20–36, 2012.

Nathan Nunn and Nancy Qian. The Potato's Contribution to Population and Urbanization: Evidence from a Historical Experiment. *The Quarterly Journal of Economics*, 126(2):593–650, 2011.

Ralph Ossa. A Quantitative Analysis of Subsidy Competition in the U.S. Working Paper 20975, National Bureau of Economic Research, February 2015.

Andrii Parkhomenko. The Rise of Housing Supply Regulation in the US: Local Causes and Aggregate Implications. *Working Paper*, 2018.

Deepak K Ray, Nathaniel D Mueller, Paul C West, and Jonathan A Foley. Yield trends are insufficient to double global crop production by 2050. *PloS one*, 8(6):e66428, 2013.

Stephen J Redding. Goods trade, factor mobility and welfare. *Journal of International Economics*, 101:148–167, 2016.

Stephen J Redding and Esteban Rossi-Hansberg. Quantitative spatial economics. *Annual Review of Economics*, 9:21–58, 2017.

Stephen J Redding and Daniel M Sturm. The costs of remoteness: Evidence from german division and reunification. *American Economic Review*, 98(5):1766–97, 2008.

Brian Reinbold, Yi Wen, et al. Housing costs and regional income inequality in china and the us. *The Regional Economist*, 26(2), 2018.

Albert Saiz. The Geographic Determinants of Housing Supply. *The Quarterly Journal of Economics*, 125(3):1253–1296, 2010.

Karen C Seto, Michail Fragkias, Burak Güneralp, and Michael K Reilly. A meta-analysis of global urban land expansion. *PloS one*, 6(8):e23777, 2011.

Curtis J Simon and Clark Nardinelli. Human capital and the rise of american cities, 1900–1990. *Regional Science and Urban Economics*, 32(1):59–96, 2002.

Kristine Lien Skog and Margrete Steinnes. How do centrality, population growth and urban sprawl impact farmland conversion in norway? *Land use policy*, 59:185–196, 2016.

Moshe Syrquin. Patterns of structural change. *Handbook of development economics*, 1:203–273, 1988.

Trevor Tombe, Xiaodong Zhu, et al. Trade, Migration and Productivity: A Quantitative Analysis of China. *American Economic Review*, 2018.

Nick Tsivanidis. The aggregate and distributional effects of urban transit infrastructure: Evidence from bogotá's transmilenio. *Unpublished manuscript*, 2018.

Matthew A Turner, Andrew Haughwout, and Wilbert Van Der Klaauw. Land Use Regulation and Welfare. *Econometrica*, 82(4):1341–1403, 2014.

Jasper van Vliet, Nick Naus, Ron JA van Lammeren, Arnold K Bregt, Jelle Hurkens, and Hedwig van Delden. Measuring the neighbourhood effect to calibrate land use models. *Computers, Environment and Urban Systems*, 41:55–64, 2013.

Jasper van Vliet, David A Eitelberg, and Peter H Verburg. A global analysis of land take in cropland areas and production displacement from urbanization. *Global environmental change*, 43:107–115, 2017.

Meinong Wang, Xu Liu, and Bo Wang. Changing and developing trend of the farmland requisition-compensation balance policy in china. *Journal of Anhui Agricultural Sciences*, 38(33):19034–19037+, 2010.

Jeffrey G Williamson. Migration and urbanization. *Handbook of development economics*, 1:425–465, 1988.

Jonathan Woetzel, Lenny Mendonca, Janamitra Devan, Stefano Negri, Yangmel Hu, Luke Jordan, Xiujun Li, Alexander Maasry, Geoff Tsen, Flora Yu, et al. Preparing for China's Urban Billion. *McKinsey Global Institute*, 2009.

Zan Yang, Rongrong Ren, Hongyu Liu, and Huan Zhang. Land leasing and local government behaviour in china: Evidence from beijing. *Urban Studies*, 52(5):841–856, 2015.

Ganlin Zhang, Yunjin Wu, and Yuguo Zhao. Physical Suitability Evaluation of Reserve Resources of Cultivated Land in China based on SOTER. *Transactions of the Chinese Society of Agricultural Engineering*, 26(4):1–8, 2010.

Tables

Table 1: Summary Statistics

Variables	(1) N of observations	(2) Mean	(3) SD
Land conversion barrier (C_u)	631	0.128	0.117
Predicted farmland percentage	631	0.436	0.071
Land ruggedness	631	0.303	0.276
ln(urban land)	13,186	3.428	0.914
ln(GDP)	13,552	4.700	1.344
ln(GDP$_{agriculture}$)	13,552	2.368	1.140
ln(GDP$_{non-agriculture}$)	13,552	4.525	1.436
ln(GDP$_{secondary}$)	13,552	3.914	1.511
ln(GDP$_{tertiary}$)	13,552	3.663	1.421
ln(N of hospital bed)	14,553	7.720	0.947
ln(government expenditure per capita)	14,361	7.103	1.308
ln(population)	2,524	13.379	0.786
ln(employment)	2,524	12.735	0.818
ln(employment$_{agriculture}$)	2,524	11.904	1.073
ln(employment$_{non-agriculture}$)	2,524	11.814	1.012
ln(employment$_{construction}$)	1,893	9.179	1.245
ln(FAR of commercial land)	4,235	1.371	0.421
ln(FAR of industrial land)	4,648	0.639	0.264
ln(FAR of residential land)	4,650	1.556	0.279
ln(price of new urban land)	4,790	5.515	1.234
ln(price of existing urban land)	4,769	6.194	1.153
ln(price of industrial land)	4,758	5.090	0.799
ln(price of commercial land)	4,583	6.579	1.073
ln(price of residential land)	4,846	6.516	1.222
% of urban land for residency	8,701	0.332	0.097
% of urban land for public facilities	8,701	0.156	0.062
% of urban land for transport and green area	8,701	0.251	0.088
% of urban land for industrial and commercial use	8,701	0.260	0.088
Remoteness	1,082	1.094	0.210

Notes. This table provides the summary statistics of the main variables used in the empirical analysis. Predicted farmland percentage refers to the predicted percentage of urban land converted from farmland since 1999. Land ruggedness is the percentage of land close to the administrative boundary that has a local slope above 15 degrees. The land conversion barrier is the interaction of the predicted farmland percentage and the land ruggedness. The unit for urban land, GDP, urban land price, and government expenditure per capita are km^2, 10^8 yuan in current prices, and 10^4 yuan per hectare in current prices, and yuan in current prices correspondingly. Employment of the construction sector is only available since 1990, while other variables related to employment were available since 1982. Data about the Floor-to-Area ratio is missing in some cases, and hence the number of observations with urban land price data is more than that with the data of Floor-to-Area ratios.

Table 2: Impact of the Land Conversion Barrier on Urban Land Supply

Depvar: ln(urban land)	(1)	(2)	(3)	(4)
CuxPost1999	-0.069***	-0.055***	-0.049***	-0.049**
	(0.019)	(0.018)	(0.018)	(0.020)
Region time-varying effects	Yes	Yes	Yes	No
Province time-varying effects	No	No	No	Yes
Time trends of population and employment rate in 1990	No	Yes	Yes	Yes
Additional controls	No	No	Yes	Yes
R-squared	0.911	0.994	0.995	0.995
Observations	13,186	13,186	13,186	13,172
N of jurisdictions	631	631	631	631

Notes. This table shows that after the Farmland Red Line Policy was adopted, a one standard deviation decrease of the land conversion barrier raises the urban land supply between 4.9% and 6.9%, depending on the specification of control variables. Column 1 displays the regression outcome with baseline controls, including city jurisdiction fixed effects, year fixed effects, and the region time-varying effects. Column 2 includes the time-varying effects of population and employment rate in 1990 additionally. Column 3 further includes the time-varying effects of illiteracy rate, percent of college graduates, percent of employment from the agricultural sector, percent of employment from the construction sector, and percent of immigrants from outside the city jurisdictions in 1990. Column 4 includes both the full set of controls and the province time-varying effects. Note that in Column 4, 14 observations are dropped because they are the only city jurisdiction from the corresponding province in a specific year. This is due to the missing data issue during 1997 and 2001 discussed in Section 3. Beijing, Tianjin, Shanghai, and Chongqing are provincial-level cities and do not have other city jurisdictions from the same province. To avoid dropping these four city jurisdictions whenever controlling province time-varying effects, I treat them as from one province instead; hence, the province time-varying effects can be estimated properly. The results are not sensitive to dropping these four city jurisdictions. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Impact of the Land Conversion Barrier on GDP

Depvar ln of GDP	(1) All	(2) Agriculture	(3) Non-agriculture	(4) Secondary	(5) Service
	Panel A. Baseline				
CuxPost1999	-0.033**	-0.021	-0.037**	-0.055***	-0.016
	(0.014)	(0.017)	(0.015)	(0.017)	(0.015)
R-squared	0.972	0.930	0.971	0.959	0.969
	Panel B. Control for the time trends of population and employment rate in 1990				
CuxPost1999	-0.039***	-0.027	-0.041***	-0.060***	-0.021
	(0.014)	(0.016)	(0.015)	(0.018)	(0.016)
R-squared	0.998	0.987	0.997	0.995	0.996
	Panel C. Additional controls				
CuxPost1999	-0.031**	-0.023	-0.042***	-0.061***	-0.018
	(0.014)	(0.017)	(0.016)	(0.018)	(0.016)
R-squared	0.998	0.987	0.997	0.995	0.996
	Panel D. Additional controls plus the province time-varying effects				
CuxPost1999	-0.014	0.003	-0.022	-0.042**	-0.005
	(0.015)	(0.020)	(0.017)	(0.020)	(0.017)
R-squared	0.998	0.989	0.998	0.996	0.997
Observations	13,552	13,552	13,552	13,552	13,552
N of jurisdictions	631	631	631	631	631

Notes. This table shows the impact of the land conversion barrier on GDP by sector. The estimates with baseline controls (Panel B) suggest that after the policy was implemented, a one standard deviation decrease of land conversion barrier raises GDP by 3.9% (Panel B Column 1). This is driven by a 6.0% increase in GDP in the secondary sector (Panel B Column 4), which is dominated by the manufacturing sector. GDP from the agricultural sector or service sector does not change significantly (Panel B Column 2 and 5). The pattern is robust to alternative control variable specifications, as shown in Panel A, C, and D. Panel A only controls for the region time-varying effects. Panel B controls for both region time-varying effects and the time-varying effects of population and employment rate in 1990. In Panel C, the additional controls include the time-varying effects of illiteracy rate, percent of college graduates, percent of employment from the agricultural sector, percent of employment from the construction sector, and percent of immigrants from other city jurisdictions in 1990. In Panel D, I include the full set of controls and further control for the province time-varying effects. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Impacts of the Land Conversion Barrier on Population and Employment

	(1)	(2)	(3)	(4)
	Panel A. Population			
CuxPost1999	-0.038***	-0.045***	-0.022**	-0.039***
	(0.012)	(0.012)	(0.009)	(0.012)
R-squared	0.932	1.000	1.000	1.000
	Panel B. Employment			
CuxPost1999	-0.038***	-0.045***	-0.035***	-0.049***
	(0.012)	(0.012)	(0.011)	(0.014)
R-squared	0.933	1.000	1.000	1.000
Region time-varying effects	Yes	Yes	Yes	No
Province time-varying effects	No	No	No	Yes
Time trends of population and employment rate in 1990	No	Yes	Yes	Yes
Additional controls	No	No	Yes	Yes
Observations	2,524	2,524	2,524	2,524
N of jurisdictions	631	631	631	631

Notes. This table shows that a lower land conversion barrier significantly raises the local population and employment after policy implementation. The results are robust to alternative control specifications. In Column 1, only region time-varying effects are controlled. Column 2 controls both region time-varying effects, and the time-varying effects of population and employment rate in 1990 are controlled. In Column 3, the additional controls include the time-varying effects of illiteracy rate, percent of college graduates, percent of employment from the agricultural sector, percent of employment from the construction sector, and percent of immigrants from other city jurisdictions in 1990. In Column 4, I include the full set of controls and further control for the province time-varying effects. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Impact of the Land Conversion Barrier on Government Service Provision

	(1)	(2)	(3)	(4)
Depvar: ln of	N of hospital bed	Government expense per capita	N of hospital bed	Government expenditure per capita
CuxPost1999	-0.000	-0.017	-0.001	-0.018
	(0.011)	(0.014)	(0.011)	(0.014)
Observations	14,553	14,361	14,553	14,361
R-squared	0.981	0.999	0.982	0.999
N of jurisdictions	631	631	631	631

Notes. This table shows that there is no change in government expenditure per capita or public facility provision in response to the land conversion barrier after 1999. Across all the columns, region time-varying effects and the time-varying effects of population and employment rate are controlled. In Column 3 and 4, I further control for the time-varying effects of illiteracy rate, percent of college graduates, percent of employment from the agricultural sector, percent of employment from the construction sector, and percent of immigrants from other city jurisdictions in 1990. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Correlation of the Land Conversion Barrier and Urban Land Prices

Depvar:	(1)	(2)	(3)	(4)
$\ln P_t - \widehat{\ln P_{1996}}$	All transactions	Sold through auction	Industrial land	Commercial land
Cu	0.058**	0.125***	0.077***	0.098***
	(0.024)	(0.022)	(0.020)	(0.025)
Observations	4,978	4,906	4,758	4,583
R-squared	0.976	0.984	0.983	0.983
N of jurisdictions	608	606	601	605

Notes. This table suggests that urban land price is positively associated with the land conversion barrier, controlling for region fixed effects and the full set of economic characteristics in 1990. The dependent variable is the difference in urban land price (ln) between year t during 2006 and 2015 and 1996. The urban land price in 1996 is not available and hence imputed using GDP by sector, urban land supply, and population in 1996. In Column 1, the price is calculated based on the transaction of urban land plots for all purposes. In Column 2, I only use land sold through auction to calculate urban land prices. In Column 3 and 4, I use land used for industrial purposes and commercial purposes to calculate urban land prices. All the regressions control for the region time-varying effects and the time-varying effects of population, employment rate, illiteracy rate, percent of college graduates, percent of employment from the agricultural sector, percent of employment from construction sector, and percent of immigrants from other city jurisdictions in 1990. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Correlation of the Land Conversion Barrier and the Urban Land Use

Depvar	(1)	(2)	(3)	(4)
% of urban land for	Residency	Public facilities	Industrial and commercial	Transport and green area
Cu	0.033	-0.166	0.106	0.003
	(0.297)	(0.187)	(0.277)	(0.275)
Observations	8,701	8,701	8,701	8,701
R-squared	0.934	0.884	0.898	0.914
N of jurisdictions	630	630	630	630

Notes. This table suggests that there is no correlation between the land conversion barrier and the proportion of urban land used for residency, business, public facilities, or transportation plus green areas. I regress the percentage of urban land used for each purpose from 2002 to 2015 against the land conversion barrier. All the regressions include the following controls: (a) region time-varying effects; (b) the time-varying effects of population, employment rate, illiteracy rate, percent of college graduates, percent of employment from the agricultural sector, percent of employment from construction sector, and percent of immigrants from other city jurisdictions in 1990; (c) the time-varying effects of GDP, GDP from non-agricultural sector, and population in 1996. Robust standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 8: Correlation of the Land Conversion Barrier and the Floor-to-Area Ratio

Depvar:	(1)	(2)	(3)
ln of the FAR of	Commercial land	Industrial land	Residential land
Cu	-0.008	0.009	0.034***
	(0.010)	(0.007)	(0.008)
Observations	4,493	4,681	4,825
R-squared	0.910	0.883	0.970
N of jurisdictions	603	601	605

Notes. This table shows that there is no correlation between the land conversion barrier and the Floor-to-Area Ratio (FAR) of newly developed commercial land or industrial land during the 2000s (Column 1 and Column 2). For residential land, the FAR of new residential land is positively correlated with the land conversion barrier (Column 3). The average FAR of new buildings in each land use category is calculated as a land area-weighted average FAR across all land transactions in that land use category. LandChina.com publishes the maximum and the minimum FAR of a land plot approved by the local government. For residential and commercial buildings, the maximum FAR is close to the actual FAR chosen by the real estate firms. For industrial buildings, the minimum FAR is close to the actual FAR chosen by the manufacturing firms. Therefore, I use the maximum FAR to approximate the FAR of the commercial and residential buildings and the minimum FAR to approximate the FAR of the industrial buildings. For robustness check, I use the average of the maximum and the minimum of the FAR to approximate the FAR chosen by individual urban land user and report the corresponding regression results in Appendix Table 39. The only difference of results based on the alternative specification is that the FAR of industrial land is also positively associated with the land conversion barrier. All the regressions include the following controls: (a) region time-varying effects; (b) the time-varying effects of population, employment rate, illiteracy rate, percent of college graduates, percent of employment from the agricultural sector, percent of employment from construction sector, and percent of immigrants from other city jurisdictions in 1990; (c) the time-varying effects of GDP, urban land supply, and population in 1996. Robust standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 9: Impact of the Land Conversion Barrier on the Compactness of an Urban Area

Depvar: Normalized Remoteness Index	(1)	(2)	(3)
CuxPost1999	-0.011	-0.013	-0.015
	(0.010)	(0.011)	(0.011)
Observations	1,082	1,082	1,082
R-squared	0.693	0.989	0.989
N of jurisdictions	541	541	541

Notes. This table shows that urban areas do not become less compact in city jurisdictions with a higher land conversion barrier. The compactness of an urban area is measured by the Normalized Remoteness Index, which approximates the average commuting cost within the urban area. The remoteness index is defined as the average distance between all interior points and the centroid. Only 541 observations are included because the boundaries of the urban areas in 1995 and 2015 are constructed using urban land raster data, which is derived from the satellite image. Weather conditions might make some urban land in a local region undetectable, and city jurisdictions in such situations in either 1995 or 2015 are dropped as the index cannot be calculated. In Column 1, only region time-varying effects are controlled. In Column 2, both region time-varying effects and the time-varying effects of population and employment rate are controlled. In Column 3, I further control for the time-varying effects of illiteracy rate, percent of college graduates, percent of employment from the agricultural sector, percent of employment from the construction sector, and percent of immigrants from other city jurisdictions in 1990. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 10: Estimation of the Supply Elasticity of Unused Land

Depvar: $\ln c_{Rnt}$	(1)	(2)	(3)	(4)	(5)
	IV Regression				OLS
$\ln\left(1 - \frac{H_{nt}+R_{nt}}{(1-\phi_n)\tilde{R}_n}\right)$	-1.779***	-2.521***	-1.390***	-1.248***	0.011
	(0.323)	(0.447)	(0.270)	(0.227)	(0.070)
Wald-F	71.45	68.07	92.70	78.30	
Observations	4,728	4,734	5,070	5,135	4,728
	First Stage				
	0.187***	0.162***	0.205***	0.212***	
	(0.022)	(0.020)	(0.021)	(0.024)	
R-squared	0.929	0.889	0.915	0.908	

Notes. This table reports the IV estimation of the price elasticity of unused land. The baseline specification (Column 1) suggests that a 10% increase in farmland price leads to a 5.6% decrease of unused land. The dependent variable is the unit cost of farmland development in city jurisdiction n in year t. The primary independent variable is the percentage of unused land. The instrument is the predicted percentage of unused land based on the total amount of unused arable land conditional on the level of local economic development. In the baseline regression specification, I trim the bottom 5% of observations and the top 5% of observations to avoid the influence of outliers. As shown in Appendix Figure 20, quite a few observations concentrate close to 0, and these locations have a low degree of land development. The result is robust to trimming the bottom 1% and top 1% observations and using the full sample, as shown in Column 3 and 4 correspondingly. The results are robust to controlling the region time-varying effect (Column 2) instead of the province time-varying effect. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

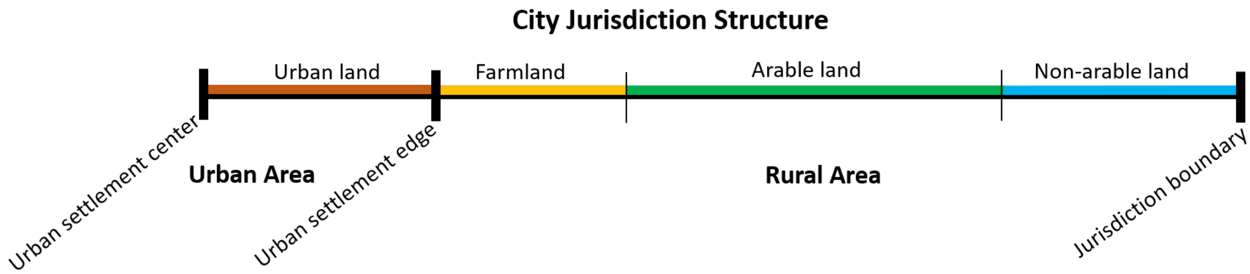
Table 11: Structural Simulation of the Reduced Form Results

	(1)	(2)	(3)	(4)
	urban land	GDP	$\text{GDP}_{nonagri}$	population
Depvar: Δln of	Panel A. Long-Difference Outcome			
$\widetilde{c_{un}}$	-0.061***	-0.044**	-0.043**	-0.037***
	(0.023)	(0.019)	(0.020)	(0.011)
Observations	631	631	631	631
R-squared	0.203	0.109	0.050	0.452
Depvar: Δln of	Panel B. Simulated Outcome			
$\widetilde{c_{un}}$	-0.049***	-0.014**	-0.020***	-0.008*
	(0.017)	(0.006)	(0.007)	(0.005)
Observations	631	631	631	631
R-squared	0.090	0.088	0.094	0.069
Depvar: ln of	Panel C. Difference-in-Difference Outcome			
$\widetilde{c_{un}}$xPost1999	-0.048***	-0.033**	-0.040**	-0.025**
	(0.017)	(0.015)	(0.016)	(0.010)
Observations	13,186	13,552	13,552	2,524
R-squared	0.995	0.998	0.997	1.000

Notes. This table suggests that the quantitative model can simulate the reduced-form results about the impact of the land conversion barrier on urban land supply, GDP, and population. From Panel B Column 1, a one standard deviation decrease of the land conversion barrier increases urban land supply by 4.9%, which is quite close to the estimates using the realized data (Panel A Column 1). Next, based on the simulation, a one standard deviation decrease of the land conversion barrier leads to a 1.4% decrease of GDP, which is driven by a 2.0% decrease of GDP from the non-agricultural sector and 0.8% decrease of population. These estimates are at the same order as the estimates based on the realized data. Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

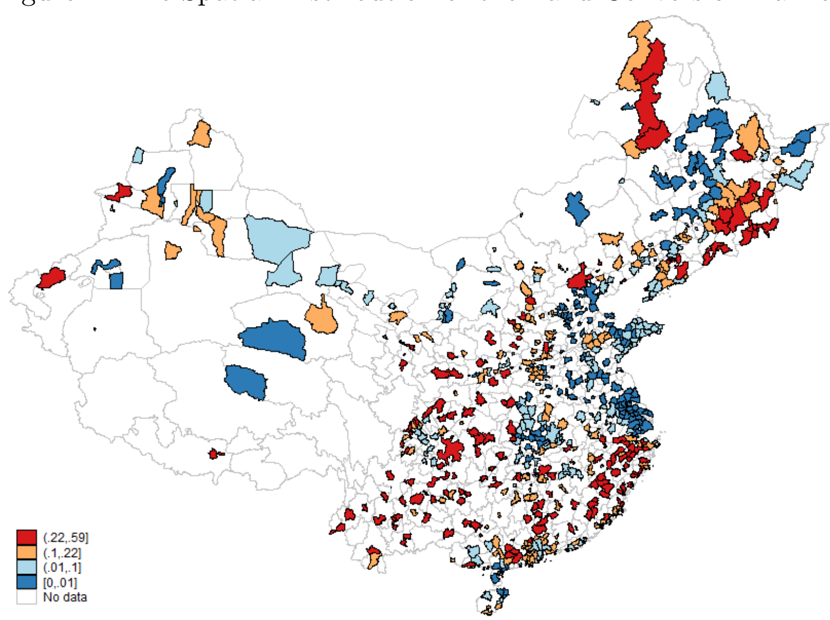
Figures

Figure 1: A Stylized City Jurisdiction



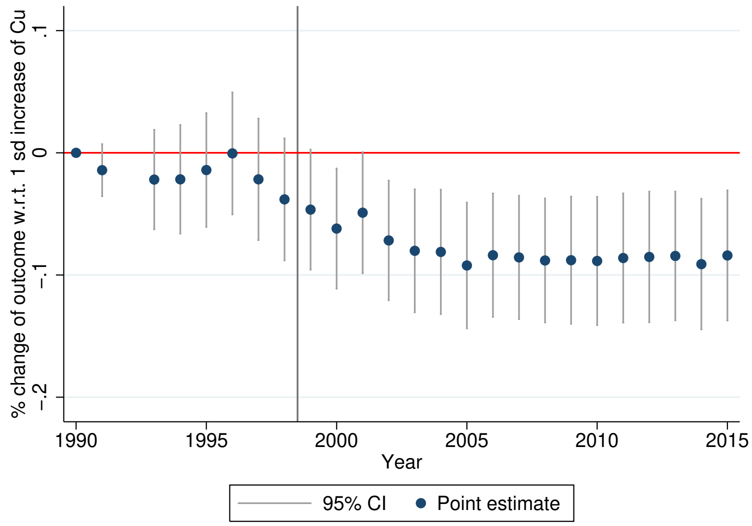
Notes. A city jurisdiction in China is administratively categorized into urban and rural land. Rural land is then subdivided into farmland, unused arable land, and unused non-arable land. Unused arable land refers to land that is neither farmland nor urban land, and can be easily cultivated into farmland. Non-arable land is unused land that is very difficult to be cultivated into farmland. The left end in the figure represents the center of the urban area. In the middle of the figure is the boundary of the urban area, which can move to the right as urbanization keeps going. The right end represents the administrative boundary of the city jurisdiction, which does not change in almost all city jurisdictions during the period of my study. Finally, in the rural area, the three types of land are mixed, and it is not necessarily that farmland plots are always next to each other, or all the arable land plots are always next to each other.

Figure 2: The Spatial Distribution of the Land Conversion Barrier



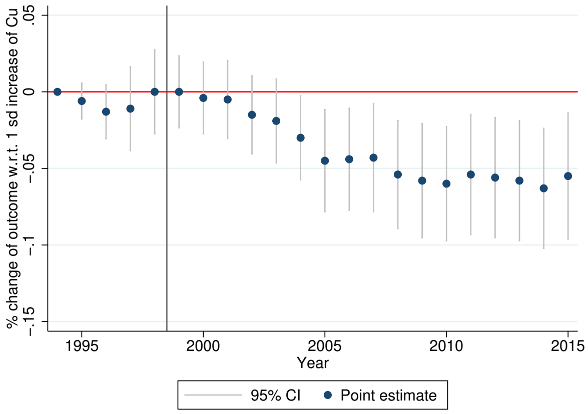
Notes. This figure shows that there is no clear regional pattern of the spatial distribution of the land conversion barrier. The value is higher towards the red end and lower towards the blue end. Appendix Figure 14 provides the spatial distribution of the two components of the land conversion barrier separately.

Figure 3: Impacts of the Land Conversion Barrier on Urban Land

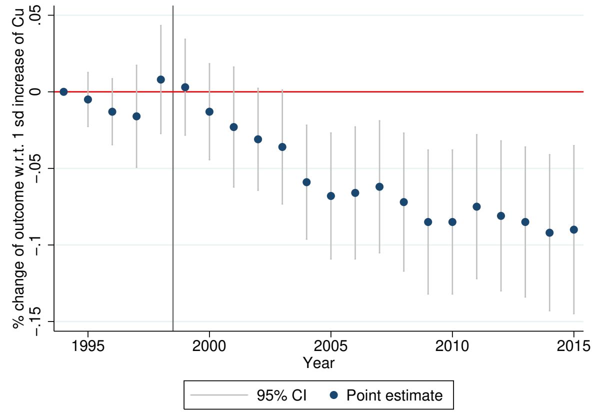


Notes. This figure suggests that city jurisdictions with different land conversion barriers had parallel trends of urban land supply before the policy implementation; after 1999, a lower land conversion barrier significantly increased urban land supply. The fact that the impact gradually showed up suggests it takes time for local government to expand urban land. The slight negative (not significant though) coefficient from 1998 is likely due to the fact that some city jurisdictions start to follow the new rule right after the regulation passed in August 1998, even though it is only officially effective from January the 1st 1999 (Order of the President of the People's Republic of China, No 8). Data on urban land supply is not available for 1992, hence the coefficient for that year is dropped. Each coefficient β_τ represents that in year τ, comparing to the year 1990, whether city jurisdictions with different levels of land conversion barrier have significantly different urban land supply. The coefficients related to the land conversion barrier are standardized such that each coefficient represents the change in outcome variables of interest in response to a one standard deviation increase in land conversion barrier. The results are robust to using a subgroup of 208 city jurisdictions that always report urban land area data since 1990, as shown in Appendix Figure 16b. Next, the results are robust to controlling time-varying impacts of population and employment rate in 1990 (Appendix Figure 16c) and controlling the time-varying impacts of the full set of economic characteristics of a city jurisdiction in 1990 (Appendix Figure 16d).

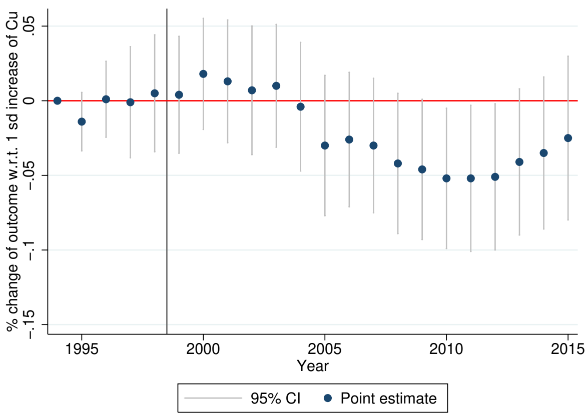
Figure 4: Impacts of the Land Conversion Barrier on GDP by Sector



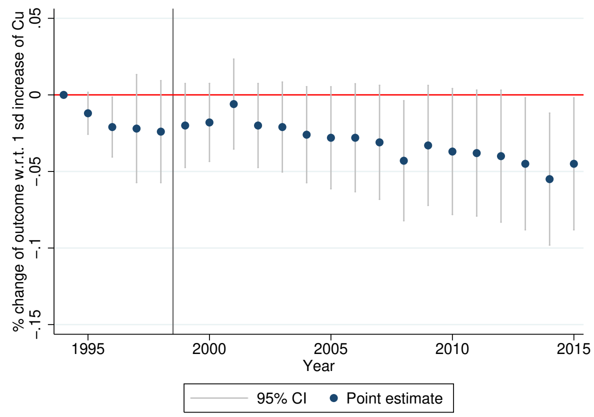
(a) Impacts on GDP



(b) Impacts on GDP from the secondary sector



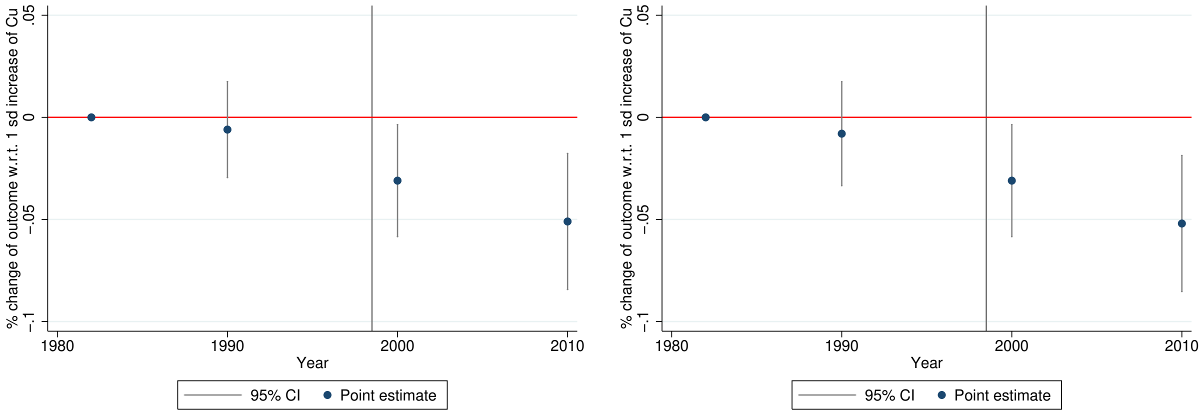
(c) Impacts on GDP from the agricultural sector



(d) Impacts on GDP from the service sector

Notes. This figure suggests that city jurisdictions with different land conversion barriers had parallel trends of GDP by sectors before the policy implementation; after 1999, a lower land conversion barrier significantly increased GDP, which is driven by its impact on GDP from the secondary sector. The results are robust to controlling time-varying impacts of population and employment rate in 1990 (Appendix Figure 17) and controlling the time-varying impacts of the full set of economic characteristics of a city jurisdiction in 1990 (Appendix Figure 18).

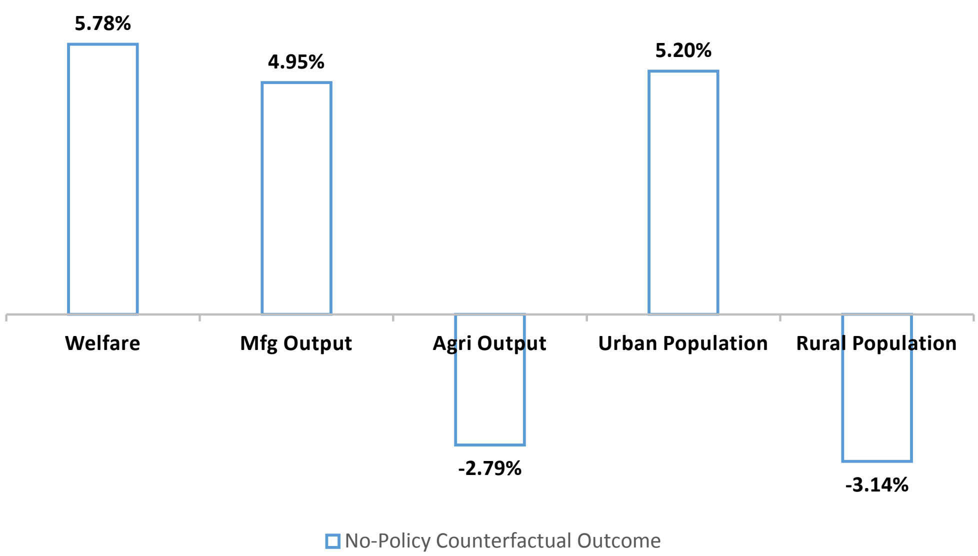
Figure 5: Impacts of the Land Conversion Barrier on Population and Employment



(a) Impacts on population (b) Impacts on employment

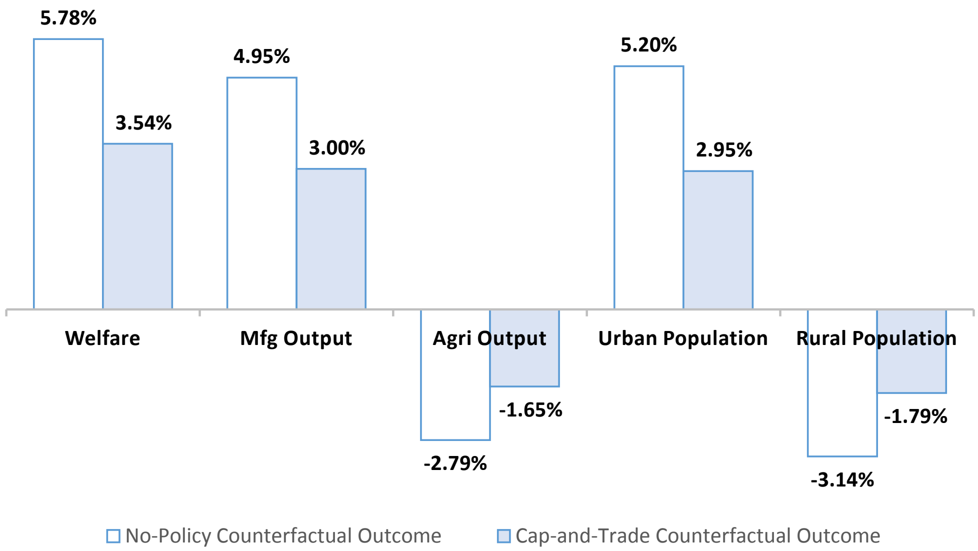
Notes. This figure suggests that city jurisdictions with different land conversion barriers had parallel trends of population and employment before the policy implementation; after 1999, a lower land conversion barrier significantly increased the local population and employment. The results are robust to controlling time-varying impacts of population and employment rate in 1990 (Appendix Figure 19 a and b) and controlling the time-varying impacts of the full set of economic characteristics of a city jurisdiction in 1990 (Appendix Figure 19 c and d).

Figure 6: No-Policy Counterfactual Outcomes



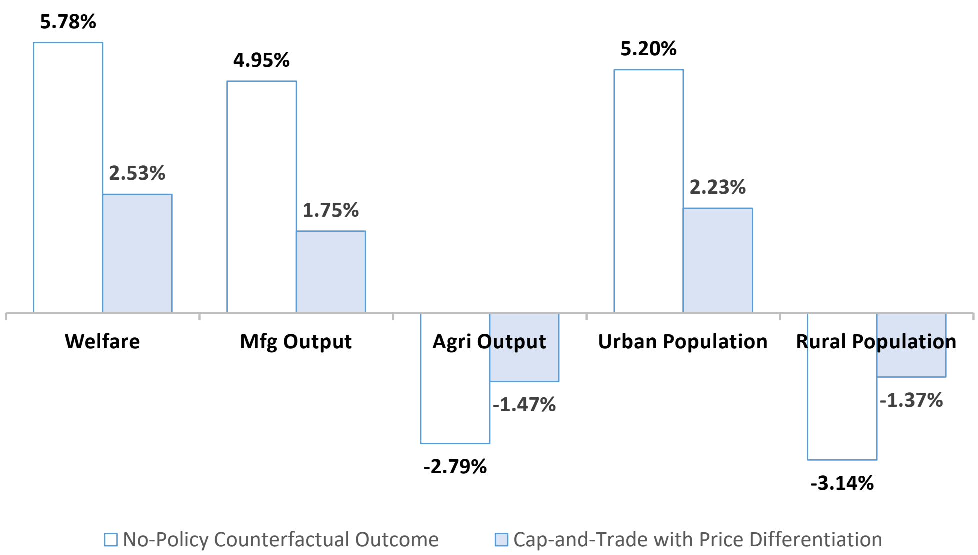
Notes. This figure suggests that the welfare of workers would have increased by 5.78% had Farmland Red Line Policy not been implemented. The economy would have been specialized more in the manufacturing sector relative to reality in 2010. There is a 4.95% increase in the manufacturing sector output and a 2.79% decrease of the agricultural sector output in the counterfactual world. The urban population would have been 5.20% more than in reality in 2010.

Figure 7: Cap-and-Trade Counterfactual Outcomes



Notes. In the counterfactual equilibrium with a cap-and-trade platform, the welfare of workers would have been 3.54% higher relative to reality. Next, the manufacturing output would have grown by 3.00%, and the agricultural output would have been 1.65% lower. Finally, the urban population would have increased by 2.95%.

Figure 8: Cap-and-Trade with Price Differentiation



Notes. In the counterfactual equilibrium with a cap-and-trade platform that has price differentiation, the welfare of workers would have been 2.53% higher relative to reality. Next, the manufacturing output would have grown by 1.75%, and the agricultural output would have been 1.47% lower. Finally, the urban population would have increased by 2.23%.

Appendix

Table of Contents

A Model Appendix

A.1 Derivation of the Optimization Conditions of Workers

This subsection shows the derivation of a worker's optimization conditions of her utility maximization problem. A worker solves the utility maximization problem through backward induction. In Step 2, conditional on choosing location n and sector s, she optimally allocates her labor income to maximize her utility. Given income w_{sn}, she spends $\mu\theta$ on agricultural products, $(1-\mu)\theta$ on manufacturing products, and $1-\theta$ on housing. Therefore, the following conditions hold,

$$h(i,n,s)^* = \frac{(1-\theta)w_{sn}}{p_{H,sn}}, \tag{28}$$

$$T_{nj}p_{Fj}c_{Fj}(i,n,s) = \frac{(T_{nj}p_{Fj})^{1-\sigma_F}}{\tilde{p}_{Fn}^{1-\sigma_F}}\mu\theta w_{sn} \quad \forall j \in N, \tag{29}$$

$$T_{nj}p_{Mj}c_{Mj}(i,n,s) = \frac{(T_{nj}p_{Mj})^{1-\sigma_M}}{\tilde{p}_{Mn}^{1-\sigma_M}}(1-\mu)\theta w_{sn} \quad \forall j \in N \tag{30}$$

where

$$\tilde{p}_{sn} = \left(\sum_{n' \in N}(T_{nn'}p_{sn'})^{1-\sigma_s}\right)^{\frac{1}{1-\sigma_s}}.$$

From (29), the maximum consumption of agricultural goods bundle, $C_F(i,n,s)^*$, is

$$C_F(i,n,s)^* = \frac{\mu\theta w_{sn}}{\tilde{p}_{Fn}}.$$

From (30), the maximum consumption of manufacturing goods bundle, $C_M(i,n,s)^*$, is

$$C_M(i,n,s)^* = \frac{(1-\mu)\theta w_{sn}}{\tilde{p}_{Mn}}.$$

As a result, the maximum utility a worker can get conditional on choosing location n and sector s is

$$V_{sn}^* = \frac{\left(\frac{\mu\theta w_{sn}}{\tilde{p}_{Fn}}\right)^{\mu\theta}\left(\frac{(1-\mu)\theta w_{sn}}{\tilde{p}_{Mn}}\right)^{(1-\mu)\theta}\left(\frac{(1-\theta)w_{sn}}{p_{H,sn}}\right)^{1-\theta}B_{sn}z_{i,n,s}}{(\mu\theta)^{\mu\theta}((1-\mu)\theta)^{(1-\mu)\theta}(1-\theta)^{1-\theta}} = \frac{w_{sn}B_{sn}}{\tilde{p}_n^{\theta}p_{H,sn}^{1-\theta}}z_{i,n,s},$$

where

$$\tilde{p}_n = \tilde{p}_{Fn}^{\mu}\tilde{p}_{Mn}^{1-\mu}.$$

We derive the aggregate demand for residential land and consumption goods. By aggregating individual worker's demand for residential land in the urban sector, we have

$$(1-\theta)w_{Mn}L_{Mn} = p_{Hn}H_n,$$

where p_{Hn} is urban land price. Similarly, by aggregating individual worker's demand for farmland used for residential purpose in the rural sector, we have

$$(1-\theta)w_{Fn}L_{Fn} = p_{Rn}R_{Hn}. \tag{31}$$

where p_{Rn} is farmland price. Finally, the total demand from workers in location n for agricultural good produced in location n' is $\frac{(T_{nn'}p_{Fn'})^{1-\sigma_F}}{\tilde{p}_{Fn}^{1-\sigma_F}}\mu\theta\left(w_{Mn}L_{Mn}+w_{Fn}L_{Fn}\right)$. The total demand from workers in location n for manufacturing good produced in location n' is $\frac{(T_{nn'}p_{Mn'})^{1-\sigma_M}}{\tilde{p}_{Mn}^{1-\sigma_M}}(1-\mu)\theta\left(w_{Mn}L_{Mn}+w_{Fn}L_{Fn}\right)$.

In Step 1, a worker chooses the location and sector that offers her the highest utility. Given that $z_{i,n,s}$ follows Frechét distribution, the probability of one choosing location n and sector s is

$$\pi_{sn}^{L} = \frac{\left(w_{sn}\tilde{p}_n^{-\theta}p_{H,sn}^{\theta-1}B_{sn}\right)^{\tilde{\nu}}}{\sum_{n'}\sum_{s'}\left(w_{s'n'}\tilde{p}_{n'}^{-\theta}p_{H,s'n'}^{\theta-1}B_{s'n'}\right)^{\tilde{\nu}}}. \tag{32}$$

Correspondingly, the expected utility, $\tilde{V}$, is expressed as following,

$$\tilde{V} = \Gamma\left(1-\tilde{\nu}^{-1}\right)\left(\sum_{n'}\sum_{s'}\left(w_{s'n'}\tilde{p}_n^{-\theta}p_{H,s'n'}^{\theta-1}B_{s'n'}\right)^{\tilde{\nu}}\right)^{\frac{1}{\tilde{\nu}}}. \tag{33}$$

Define $\nu = \frac{\tilde{\nu}}{1-\beta\tilde{\nu}}$, (32) is equivalent to

$$\pi_{sn}^{L} = \frac{(w_{sn}\tilde{p}_n^{-\theta}p_{H,sn}^{\theta-1}\bar{B}_{sn})^{\nu}}{\sum_{n'}\sum_{s'}\left(w_{s'n'}\tilde{p}_{n'}^{-\theta}p_{H,s'n'}^{\theta-1}\bar{B}_{s'n'}\right)^{\nu}}. \tag{34}$$

This is because

$$\frac{L_{s'n'}}{L_{sn}} = \frac{(w_{s'n'}\tilde{p}_{n'}^{-\theta}p_{H,s'n'}^{\theta-1}\bar{B}_{s'n'}L_{s'n'}^{\beta})^{\tilde{\nu}}}{(w_{sn}\tilde{p}_n^{-\theta}p_{H,sn}^{\theta-1}\bar{B}_{sn}L_{sn}^{\beta})^{\tilde{\nu}}} = \frac{(w_{s'n'}\tilde{p}_{n'}^{-\theta}p_{H,s'n'}^{\theta-1}\bar{B}_{s'n'})^{v}}{(w_{sn}\tilde{p}_n^{-\theta}p_{H,sn}^{\theta-1}\bar{B}_{sn})^{\nu}}.$$

Therefore, $\frac{1}{\pi_{sn}^{L}}$ can be rewritten as

$$\frac{1}{\pi_{sn}^{L}} = \frac{\sum_{n'}\sum_{s'}(w_{s'n'}\tilde{p}_{n'}^{-\theta}p_{H,s'n'}^{\theta-1}\bar{B}_{s'n'})^{v}}{(w_{sn}\tilde{p}_n^{-\theta}p_{H,sn}^{\theta-1}\bar{B}_{sn})^{\nu}},$$

which is exactly (34).

It can be shown that $\bar{V}$ is a linear transformation of $\tilde{V}$, where $\bar{V} \equiv \left(\sum_{n'}\sum_{s'}(w_{s'n'}\tilde{p}_{n'}^{-\theta}p_{H,s'n'}^{\theta-1}\bar{B}_{s'n'})^{v}\right)^{\frac{1}{\nu}}$. To see this, replace $B_{sn} = \bar{B}_{sn}\left(\bar{L}\pi_{sn}\right)^{\beta}$ in (33), we have

$$\tilde{V}^{\tilde{\nu}} = \bar{L}^{\beta\tilde{\nu}}\Gamma\left(1-\tilde{\nu}^{-1}\right)\left(\sum_{n'}\sum_{s'}\left(w_{Mn'}\tilde{p}_{n'}^{-\theta}p_{H,s'n'}^{\theta-1}\bar{B}_{s'n'}\right)^{\tilde{\nu}}(\pi_{s'n'}^{L})^{\beta\tilde{\nu}}\right). \tag{35}$$

Replace $\pi^L_{s'n'}$ with (34), the following condition holds

$$\tilde{V}^{\tilde{\nu}} = \bar{V}^{-\nu\beta\tilde{\nu}}\bar{L}^{\beta\tilde{\nu}}\Gamma\left(1-\tilde{\nu}^{-1}\right)\sum_{n'}\sum_{s'}\left(w_{Mn'}\tilde{p}_{n'}^{-\theta}p_{H,s'n'}^{\theta-1}\bar{B}_{s'n'}\right)^{\tilde{\nu}+\nu\beta\tilde{\nu}}.$$

Since $\tilde{\nu}+\nu\beta\tilde{\nu} = \tilde{\nu}(1+\frac{\tilde{\nu}\beta}{1-\beta\tilde{\nu}}) = \nu$,

$$\tilde{V}^{\tilde{\nu}} = \bar{V}^{-\nu\beta\tilde{\nu}}\bar{L}^{\beta\tilde{\nu}}\Gamma\left(1-\tilde{\nu}^{-1}\right)\sum_{n'}\sum_{s'}\left(w_{Mn'}\tilde{p}_{n'}^{-\theta}p_{H,s'n'}^{\theta-1}\bar{B}_{s'n'}\right)^{\nu} = \bar{V}^{\tilde{\nu}}\bar{L}^{\beta\tilde{\nu}}\Gamma\left(1-\tilde{\nu}^{-1}\right)$$

Therefore, we get the following equation:

$$\tilde{V} = \bar{V}\bar{L}^{\beta}\Gamma\left(1-\tilde{\nu}^{-1}\right)^{\frac{1}{\nu}}. \tag{36}$$

As $\bar{L}^{\beta}\Gamma\left(1-\tilde{\nu}^{-1}\right)^{\frac{1}{\nu}}$ is a constant, $\bar{V}$ is a linear transformation of $\tilde{V}$.

Note that (34) can be further simplified to

$$\pi^L_{sn} = \bar{V}^{-\nu}(w_{Mn}\tilde{p}_n^{-\theta}p_{H,sn}^{\theta-1}\bar{B}_{sn})^{\nu}. \tag{37}$$

Therefore, the labor supply in sector s, location n is

$$L_{sn} = \bar{L}\bar{V}^{-\nu}(w_{sn}\tilde{p}_n^{-\theta}p_{H,sn}^{\theta-1}\bar{B}_{sn})^{\nu}. \tag{38}$$

A.2 Impacts of the Farmland Red Line Policy on Land Markets

This subsection studies the impacts of the Farmland Red Line Policy on land markets through a partial equilibrium version of the model. In the partial equilibrium, the supply functions of urban land and farmland are derived from the landlord's profit maximization problem, while the demand for urban land and farmland are taken as given and are subject to demand shocks.

I first discuss the condition under which the constraint binds since only then the supply of urban land and farmland deviate from the no-policy market equilibrium outcomes. I define the state of the economy right before policy implementation as the *initial equilibrium.* The minimum farmland quantity $\bar{R}_n$ is the optimal amount of farmland chosen by the landlord in location n in the initial equilibrium. If demands for urban land or farmland do not change after policy implementation, the profit-maximizing quantity of farmland always equals the optimal amount of farmland in the initial equilibrium. As a result, the constraint is not binding. Proposition 1 summarizes the condition for the minimum farmland quantity constraint to bind in a location.

Proposition 1 *The minimum farmland quantity constraint is binding in a location if after policy implementation, the increase in local demand for urban land is large relative to the increase in demand for farmland, such that it is profitable to reduce farmland supply.*

Intuitively, the constraint is binding if and only if without the policy, it is profitable for the landlord to reduce farmland supply. Moreover, the landlord would want to reduce farmland to supply more urban land if the increase in demand for urban land is large relative to the increase in demand for farmland.

To see the second point through the model, we rewrite the demand function of urban land at the initial equilibrium as $p_{Hn}H_n = E_{Hn}$ and the demand function of farmland as $p_{Rn}R_n = E_{Rn}$. E_{Hn} represents the total spending on urban land, while E_{Rn} represents the total spending on farmland. In this subsection, E_{Hn} and E_{Rn} are taken as given and are subject to exogenous shocks.[55] Suppose demand shocks to urban land and farmland occurs to the location. Without the Farmland Red Line Policy, the change in farmland, $d\ln R_n$ can be expressed as a function of the demand shock to urban land $d\ln E_{Hn}$ and the demand shock to farmland $d\ln E_{Rn}$:

$$d\ln R_n = k_{1n} d\ln E_{Rn} - k_{2n} d\ln E_{Hn}. \tag{39}$$

where $k_{1n} \equiv \frac{(1-\phi_n)\tilde{R}_n+(\zeta-1)H_n-R_n}{(1-\phi_n)\tilde{R}_n+(\zeta-1)(H_n+R_n)}$ and $k_{2n} \equiv \frac{\zeta H_n}{(1-\phi_n)\tilde{R}_n+(\zeta-1)(H_n+R_n)}$.[56] The minimum farmland quantity constraint is binding if without the policy, the landlord reduces the farmland supply after the demand shocks: $d\ln R_n < 0$. Therefore, the condition under which the constraint is binding is derived:

$$d\ln E_{Hn} - \frac{k_{1n}}{k_{2n}} d\ln E_{Rn} \geq 0. \tag{40}$$

This relation indicates that the constraint is binding if the urban land demand shock is large relative to the farmland demand shock.

The constraint binding condition is likely to be met in many locations in China for multiple reasons. First, many places had a rapid urbanization process after 1999. It works as a positive demand shock to urban land and negative demand shock for farmland, hence making the constraint bind. Moreover, joining WTO also brings business opportunities that primarily benefit the manufacturing sector. It serves as a positive demand shock to urban land demand while there is no change in demand for farmland, which again makes the constraint bind. Finally, the initial difference in the real wage between the rural sector and urban sector and the further relaxation of Hukou restriction during 2000 and 2005 also make more rural populations move to urban sectors. It is a positive demand shock to urban land demand and negative demand shock for farmland. It again guarantees that the constraint is binding.

Next, I discuss the impact of the Farmland Red Line Policy on the quantity of urban land and farmland when the constraint binds. First of all, the quantity of farmland in the new equilibrium is greater than in the no-policy case. It is because the farmland supply cannot decrease as in no-policy case after demand shocks happen. Second, there is an under-supply of urban land because the policy increases the percent of land development and drives up the marginal land development cost.[57] As a

[55]Note that in general equilibrium, both E_{Hn} and E_{Rn} are determined by the total income in the urban sector and rural sector, hence endogenous.

[56]This expression is derived by log differentiating the demand and supply functions of urban land and farmland around the initial equilibrium. Appendix A.3 provides detailed derivation.

[57]Suppose the percent of land development is weakly lower when there is the policy. It indicates that p_{Hn} is weakly lower and at least the same amount of H_n or even more is consumed than in the no-policy case. But farmland is also greater than in the no-policy case. This suggests that the percent of land development, $\frac{H_n+R_n}{(1-\phi_n)\tilde{R}_n}$, must be higher than

result, the urban land price is higher and less urban land is supplied. Therefore, we get the following proposition.

Proposition 2 *If the constraint is binding, there is an over-supply of farmland and an under-supply of urban land.*

At the end of this subsection, I discuss how the percentage of undevelopable land, ϕ_n, affects the likelihood of the constraint to be binding and the degree of the under-supply of urban land when the constraint binds. Intuitively, a higher ϕ_n indicates a lower price elasticity of developed land. As a result, at a given positive demand shock for urban land, the landlord is more likely to reduce farmland to increase urban land supply without the policy. Therefore, the constraint is more likely to be binding under the policy. Next, given that the constraint is binding, an extra unit of farmland drives up the marginal cost of land development further when the supply elasticity of developed land is smaller. This further increases the urban land price and decreases the quantity of urban land.

Proposition 3 *A higher percent of undevelopable land, ϕ_n, makes the farmland quantity constraint more likely to be binding. Moreover, given that the constraint is binding, a higher ϕ_n causes a more severe under-supply of urban land.*

Proof. Suppose the urban land demand shock and the farmland demand shock are drawn from a joint probability distribution $F_{X_h,X_r}(x_h, x_r)$. The probability that the constraint is binding is

$$Pr\left(d\ln E_{Hn} \geq a_n d\ln E_{Rn}\right) = \int_{-\infty}^{+\infty}\int_{-\infty}^{\frac{x_h}{a_n}} f_{X_h,X_r}(x_h, x_r)dx_r dx_h$$

where $a_n = \frac{(1-\phi_n)\tilde{R}_n+(\zeta-1)H_n-R_n}{\zeta H_n}$ and a_n decreases with ϕ_n. Therefore, at any given x_h, a higher ϕ_n increases $\frac{x_h}{a_n}$ and hence $\int_{-\infty}^{\frac{x_h}{a_n}} f_{X_h,X_r}(x_h, x_r)dx_r$. As a result, a higher ϕ_n leads to a higher probability for the constraint to be binding.

Next, we show that a higher ϕ_n causes a more severe under-supply of urban land given that the constraint is binding. To see this, the difference of urban land supply with and without the policy after the demand shocks is:

$$d\ln H_n' - d\ln H_n = -\frac{R_n}{H_n}k_{2n}\left(\frac{k_{2n}}{k_{1n}}d\ln E_{Hn} - d\ln E_{Rn}\right)$$

Both k_{2n} and $\frac{k_{2n}}{k_{1n}}$ increases with ϕ_n, which indicates that the difference of new urban land supply after demand shocks is more negative. Hence a higher ϕ_n causes a more severe under-supply of urban land. Q.E.D.

Proposition 3 indicates that ceteris paribus, a location with a higher percentage of undevelopable land has a smaller increase in urban land supply after policy implementation. This is consistent with

when there is no policy, which is a contradiction.

the empirical findings from Section 4 that locations with a higher land ruggedness in the area where new farmland is likely to be developed have less urban land supply after the policy began.

A.3 Derivation of the Change in Land Use in Response to Land Demand Shocks

By log differentiating the demand functions of urban land and farmland around the initial equilibrium, we have

$$d\ln p_{Hn} + d\ln H_n = d\ln E_{Hn}, \tag{41}$$

$$d\ln p_{Rn} + d\ln R_n = d\ln E_{Rn}. \tag{42}$$

When there is no Farmland Red Line Policy, we log differentiate the corresponding urban land supply function and farmland supply function and get

$$d\ln p_{Hn} = d\ln p_{Rn}, \tag{43}$$

$$d\ln p_{Rn} = \frac{\zeta H_n}{(1-\phi_n)\tilde{R}_n - R_n - H_n} d\ln H_n + \frac{\zeta R_n}{(1-\phi_n)\tilde{R}_n - R_n - H_n} d\ln R_n. \tag{44}$$

Combining (41), (42), (43), and (44), we have 4 equations with 4 unknowns: $dlnR_n$, $dlnH_n$, $dlnp_{Rn}$, and $dlnp_{Hn}$. Therefore, $dlnR_n$ can be expressed as a function of $dlnE_{Hn}$ and $dlnE_{Rn}$:

$$dlnR_n = \frac{\Big((1-\phi_n)\tilde{R}_n + (\zeta-1)H_n - R_n\Big)\, dlnE_{Rn} - \zeta H_n dlnE_{Hn}}{(1-\phi_n)\tilde{R}_n + (\zeta-1)(H_n+R_n)}. \tag{45}$$

Denote $k_{1n} = \frac{(1-\phi_n)\tilde{R}_n + (\zeta-1)H_n - R_n}{(1-\phi_n)\tilde{R}_n + (\zeta-1)(H_n+R_n)}$ and $k_{2n} = \frac{\zeta H_n}{(1-\phi_n)\tilde{R}_n + (\zeta-1)(H_n+R_n)}$, (45) can be rewritten as

$$dlnR_n = k_{1n} dlnE_{Rn} - k_{2n} dlnE_{Hn}. \tag{46}$$

Similarly, $dlnH_n$ can be expressed as:

$$dlnH_n = \frac{(1-\phi_n)\tilde{R}_n + (\zeta-1)\,R_n - H_n}{(1-\phi_n)\tilde{R}_n + (\zeta-1)\,(R_n+H_n)} dlnE_{Hn} - \frac{\zeta R_n}{(1-\phi_n)\tilde{R}_n + (\zeta-1)\,(R_n+H_n)} dlnE_{Rn}. \tag{47}$$

Next, we derive the change in urban land and farmland in response to the demand shocks for urban land and farmland when the Farmland Red Line Policy intervenes the land development decisions. At the initial equilibrium, the policy is announced first, and the demand shocks $dlnE_{Hn}$ and $dlnE_{Rn}$ happen afterward. Denote the change in farmland and urban land quantities as $dlnR'_n$ and $dlnH'_n$.

If the demand shocks make the farmland weakly increasing without the policy requirement, which means $dlnR \geq 0$, the constraint is not binding, and $dlnR'_n$ and $dlnH'_n$ are the same as $dlnR_n$ and $dlnH_n$ correspondingly. If $dlnR < 0$, without the policy, it is profitable for the landlord to reduce farmland. Therefore, with the policy, the constraint is binding and $dlnR'_n = 0$. Next, $dlnH'_n$ is expressed as

$$dlnH'_n = \frac{(1-\phi_n)\tilde{R}_n - R_n - H_n}{\zeta H_n + (1-\phi_n)\tilde{R}_n - R_n - H_n} dlnE_{Hn}. \tag{48}$$

Finally, the difference of urban land supply after the demand shocks with and without the policy is simply

$$dlnH_n'-dlnH_n=-\frac{\zeta R_n}{\left(1-\phi_n\right)\tilde{R}_n+\left(\zeta-1\right)\left(R_n+H_n\right)}\left(\frac{\zeta H_n}{\left(1-\phi_n\right)\tilde{R}_n+\left(\zeta-1\right)H_n-R_n}dlnE_{Hn}-dlnE_{Rn}\right).$$

Note that $\frac{\zeta R_n}{(1-\phi_n)\tilde{R}_n+(\zeta-1)(R_n+H_n)}=\frac{R_n}{H_n}k_{2n}$ and $\frac{\zeta H_n}{(1-\phi_n)\tilde{R}_n+(\zeta-1)H_n-R_n}=\frac{k_{2n}}{k_{1n}}$. The expression above can be written as

$$dlnH_n'-dlnH_n=-\frac{R_n}{H_n}k_{2n}\left(\frac{k_{2n}}{k_{1n}}dlnE_{Hn}-dlnE_{Rn}\right) \tag{49}$$

A.4 Recover the Unobservables

This subsection introduces the procedure to recover the unobserved variables that rationalize the observed data from 2010 as an equilibrium. The recovery of unobservables takes four steps.

In Step 1, I solve the set of wages, $\{w_{Mn},w_{Fn}\}$, and land prices, $\{p_{Hn},p_{Rn}\}$, that are consistent with the data according to the equilibrium conditions. The observed data include output from the manufacturing and agricultural sector in location n, $\{E_{Mn},E_{Fn}\}$, working population from the manufacturing and agricultural sector in location n, $\{L_{Mn},L_{Fn}\}$, and urban land and farmland, $\{H_n,R_n\}$, correspondingly.[58] Specifically, they are recovered using the following conditions:

$$\begin{aligned}
w_{Mn}&=E_{Mn}/L_{Mn},\\
w_{Fn}&=\gamma E_{Fn}/L_{Fn},\\
p_{Hn}&=(1-\theta_M)E_{Mn}/H_n,\\
p_{Rn}&=(1-\gamma\theta_F)E_{Fn}/R_n.
\end{aligned}$$

In Step 2, I recover the prices of manufacturing and agricultural products from across locations. I first parameterize bilateral trade cost as the bilateral trade data is not available at city jurisdiction level. Therefore, the bilateral trade cost within China is parameterized as a function of bilateral distances between two locations, as in Redding (2016):

$$T_{nn'}^{-\sigma_M}=d_{nn'}^{-D_M},$$

where $d_{nn'}$ is the distance between the centroid of the two locations n and n'. The distance decay elasticity for trade in manufacturing goods is specified as 1, following the literature (Faber and Gaubert, 2019).

I then use the market clearing condition for the manufacturing goods (19) and the price index of the manufacturing goods (4) to recover $\{p_{Mn},Y_{Mn},\tilde{p}_{Mn}\}$ that are consistent with the observed output

[58] For rural regions, all the working population is classified as agricultural workers, and all the final output is classified as agricultural output. This is a realistic simplification because, according to the population census and City Statistical Yearbook in 2010, 73% of employment belongs to the agricultural sector, and 56% of output in rural regions is agricultural output.

by sector and calibrated trade costs $\{T_{nn'}\}$. p_{Mn} and $\tilde{p}_{Mn}$ are solved by iterating over (50) and (51).[59]

$$\tilde{p}_{Mn} = \left(\sum_{n'} \left(T_{nn'} p_{Mn'}\right)^{1-\sigma_M} \right)^{\frac{1}{1-\sigma_M}}, \tag{50}$$

$$p_{Mj}^{\sigma_M - 1} = \frac{1-\mu}{E_{Mj}} \sum_n \frac{\left(T_{jn}\right)^{1-\sigma_M}}{\tilde{p}_{Mn}^{1-\sigma_M}} \left(E_{Mn} + E_{Fn}\right). \tag{51}$$

In Step 3, I calculate $\{\bar{A}_{Mn}, \bar{A}_{Fn}\}$ using the production function of the manufacturing and agricultural sectors:

$$\bar{A}_{Mn} = \frac{w_{Mn}}{p_{Mn} L_{Mn}^{1+\alpha}},$$
$$\bar{A}_{Fn} = \frac{E_{Fn}}{p_{Fn} L_{Fn}^{\gamma} R_{Fn}^{1-\gamma}},$$

where $R_{Fn} = \frac{1-\gamma}{1-\theta_F \gamma} R_n$.

In Step 4, $\{\bar{B}_{Mn}, \bar{B}_{Fn}\}$ and $\bar{V}$ can be solved up to a constant using labor supply functions (6) and labor clearing condition (5). I specify the measure of utility such that at equilibrium, the expected utility of a representative worker is 1. Therefore, $\bar{B}_{Mn}$ is uniquely determined by

$$\bar{B}_{Mn} = \left(\frac{L_{Mn}}{\bar{L}} \right)^{\frac{1}{\nu}} \bar{V} w_{Mn}^{-1} \tilde{p}_{Fn}^{\mu\theta_M} \tilde{p}_{Mn}^{(1-\mu)\theta_M} p_{Hn}^{1-\theta_M},$$

and $\bar{B}_{Fn}$ is uniquely determined by

$$\bar{B}_{Fn} = \left(\frac{L_{Fn}}{\bar{L}} \right)^{\frac{1}{\nu}} \bar{V} w_{Fn}^{-\theta_F} \tilde{p}_{Fn}^{\mu\theta_F} \tilde{p}_{Mn}^{(1-\mu)\theta_F} p_{Rn}^{1-\theta_F}.$$

A.5 Solve the Counterfactual Equilibrium

A.5.1 Remove the Farmland Red Line Policy

I apply the hat algebra to simulate the counterfactual outcome without the Farmland Red Line Policy in 2010. By expressing the equilibrium conditions of the model in changes relative to their baseline values ($\hat{x} = \frac{x'}{x}$) and expressing $\left\{ \hat{Y}_{Fn}, \hat{Y}_{Mn}, \hat{p}_{Fn}, \hat{p}_{Mn}, \hat{\tilde{p}}_n, \hat{p}_{Hn}, \hat{p}_{Rn}, \hat{R}_{Hn}, \hat{R}_{Fn} \right\}$ as functions of the rest of variables, there are 9 equilibrium conditions with 9 unknown vectors $\{\hat{L_{Mn}}, \hat{w_{Mn}}, \hat{R}_n, \hat{L_{Fn}}, \hat{w_{Fn}}, \hat{H}_n, \hat{\tilde{p}}_{Fn}, \hat{\tilde{p}}_{Mn}, \hat{\bar{V}}\}$.

$$\left(\hat{\tilde{p}}_{Fn}\right)^{1-\sigma_F} = \sum_{n' \in N} \pi_{pFnn'} \left(\hat{w_{Fn'}}\right)^{1-\sigma_F} \left(\hat{L_{Fn'}}\right)^{(1-\gamma)(1-\sigma_F)} \left(\hat{R_{n'}}\right)^{(\gamma-1)(1-\sigma_F)}, \tag{52}$$

[59] At interation round t = 0, I specify $p_{Mn}^0 = 1$ for all n. Plug $\{p_{Mn}^0\}$ into (50), we solve $\tilde{p}_{Mn}^1$ that satisfy the set of equations. Next, by pluging $\tilde{p}_{Mn}^1$ into (51), we get $\{p_{Mn}^1\}$, which is an update of product prices. Iterate until $p_{Mn}^t \to p_{Mn}^{t+1}$ and $\tilde{p}_{Mn}^t \to \tilde{p}_{Mn}^{t+1}$. $\{p_{Fn}, Y_{Fn}, \tilde{p}_{Fn}\}$ are recovered in the same way. Finally, $\tilde{p}_n$ can be recovered once $\tilde{p}_{Mn}$ and $\tilde{p}_{Fn}$ are available.

where $\pi_{pFnn'} = \frac{\left(T_{nn'}(\gamma A_{Fn'})^{-1} w_{Fn'} L_{Fn'}^{1-\gamma} R_{n'}^{\gamma-1}\right)^{1-\sigma_F}}{\tilde{p}_{Fn}^{1-\sigma_F}}$.

$$\left(\hat{\tilde{p}}_{Mn}\right)^{1-\sigma_M} = \sum_{n' \in N_c} \pi_{pMnn'} \left(\hat{w_{Mn'}}\right)^{1-\sigma_M} \left(\hat{L_{Mn'}}\right)^{-\alpha(1-\sigma_M)}, \tag{53}$$

where $\pi_{pMnn'} = \frac{\left(T_{nn'}\left(\bar{A}_{Mn'}\right)^{-1} w_{Mn'} L_{Mn'}^{-\alpha}\right)^{1-\sigma_M}}{\tilde{p}_{Mn}^{1-\sigma_M}}$.

$$\left(\hat{L_{Mn}}\right)^{\frac{1}{\nu}+1-\theta_M} = \left(\hat{\tilde{V}}\right)^{-1} \left(\hat{w_{Mn}}\right)^{\theta_M} \left(\hat{\tilde{p}}_{Fn}\right)^{-\mu\theta_M} \left(\hat{\tilde{p}}_{Mn}\right)^{-(1-\mu)\theta_M} \left(\hat{H_n}\right)^{1-\theta_M}. \tag{54}$$

$$\left(\hat{L_{Fn}}\right)^{\frac{1}{\nu}+1-\theta_F} = \left(\hat{\tilde{V}}\right)^{-1} \left(\hat{w_{Fn}}\right)^{\theta_F} \left(\hat{\tilde{p}}_{Fn}\right)^{-\mu\theta_F} \left(\hat{\tilde{p}}_{Mn}\right)^{-(1-\mu)\theta_F} \left(\hat{R_n}\right)^{1-\theta_F}. \tag{55}$$

$$\left(\hat{w_{Mn}}\right)^{\sigma_M} \left(\hat{L_{Mn}}\right)^{1-\alpha(\sigma_M-1)} = \sum_{j \in N_c} \pi_{M1nj} \left(\hat{\tilde{p}}_{Mj}\right)^{\sigma_M-1} \hat{w_{Mj}} \hat{L_{Mj}} + \sum_{j \in N} \pi_{M2nj} \left(\hat{\tilde{p}}_{Mj}\right)^{\sigma_M-1} \hat{w_{Fj}} \hat{L_{Fj}}, \tag{56}$$

where $\pi_{M1nj} = (1-\mu)\left(T_{nj}\right)^{1-\sigma_M} \frac{\tilde{p}_{Mj}^{\sigma_M-1} E_{Mj}}{p_{Mn}^{\sigma_M-1} E_{Mn}}$ and $\pi_{M2nj} = (1-\mu)\left(T_{nj}\right)^{1-\sigma_M} \frac{\tilde{p}_{Mj}^{\sigma_M-1} E_{Fj}}{p_{Mn}^{\sigma_M-1} E_{Mn}}$.

$$\left(\hat{w_{Fn}}\right)^{\sigma_F} \left(\hat{L_{Fn}}\right)^{\sigma_F(1-\gamma)+\gamma} \left(\hat{R_n}\right)^{-(1-\gamma)(\sigma_F-1)} = \sum_{j \in N_c} \pi_{F1nj} \left(\hat{\tilde{p}}_{Fj}\right)^{\sigma_F-1} \hat{w_{Mj}} \hat{L_{Mj}} + \sum_{j \in N} \pi_{F2nj} \left(\hat{\tilde{p}}_{Fj}\right)^{\sigma_F-1} \hat{w_{Fj}} \hat{L_{Fj}}, \tag{57}$$

where $\pi_{F1nj} = \mu\left(T_{nj}\right)^{1-\sigma_F} \left(E_{Fn}^{-1} p_{Fn}^{1-\sigma_F}\right) \tilde{p}_{Fj}^{\sigma_F-1} E_{Mj}$ and $\pi_{F2nj} = \mu\left(T_{nj}\right)^{1-\sigma_F} \left(E_{Fn}^{-1} p_{Fn}^{1-\sigma_F}\right) \tilde{p}_{Fj}^{\sigma_F-1} E_{Fj}$.

$$\begin{aligned} \hat{\tilde{V}}^{\nu} = & \sum_n \pi_{VMn} \left(\hat{w_{Mn}}^{\theta_M} \left(\hat{\tilde{p}}_{Fn}\right)^{-\mu\theta_M} \left(\hat{\tilde{p}}_{Mn}\right)^{-(1-\mu)\theta_M} \left(\hat{L_{Mn}}\right)^{\theta_M-1} \hat{H_n}^{1-\theta_M}\right)^{\nu} \\ & + \sum_n \pi_{VFn} \left(\hat{w_{Fn}}^{\theta_F} \left(\hat{\tilde{p}}_{Fn}\right)^{-\mu\theta_F} \left(\hat{\tilde{p}}_{Mn}\right)^{-(1-\mu)\theta_F} \left(\hat{L_{Fn}}\right)^{\theta_F-1} \hat{R_n}^{1-\theta_F}\right)^{\nu}, \end{aligned} \tag{58}$$

where $\pi_{VFn} = \frac{L_{Fn}}{\bar{L}}$ and $\pi_{VMn} = \frac{L_{Mn}}{\bar{L}}$.

$$\frac{\hat{w_{Mn}} \hat{L_{Mn}}}{\hat{H_n}} = \left(\hat{\tilde{p}}_{Fn}\right)^{\mu} \left(\hat{\tilde{p}}_{Mn}\right)^{(1-\mu)} \left(\pi_{r1n} - \pi_{r2n}\hat{R_n} - \pi_{r3n}\hat{H_n}\right)^{-\zeta}, \tag{59}$$

where $\pi_{r1n} = \frac{(1-\phi_{n2})\tilde{R}_n}{(1-\phi_{n2})\tilde{R}_n - R_n - H_n}$, $\pi_{r2n} = \frac{R_n}{(1-\phi_{n2})\tilde{R}_n - R_n - H_n}$ and $\pi_{r3n} = \frac{H_n}{(1-\phi_{n2})\tilde{R}_n - R_n - H_n}$.

$$\left(\hat{\tilde{p}}_{Fn}\right)^{\mu} \left(\hat{\tilde{p}}_{Mn}\right)^{(1-\mu)} \left(\pi_{r1n} - \pi_{r2n}\hat{R_n} - \pi_{r3n}\hat{H_n}\right)^{-\zeta} = \pi_{r4n} \frac{\hat{w_{Fn}} \hat{L_{Fn}}}{\hat{R_n}}, \tag{60}$$

where $\pi_{r4n} = \frac{\frac{1-\theta_F\gamma}{\gamma}\delta w_{Fn} L_{Fn}}{\tilde{p}_{Fn}^{\mu}\tilde{p}_{Mn}^{(1-\mu)} c_n \left(1-\frac{R_n+H_n}{(1-\phi_{n2})\tilde{R}_n}\right)^{-\zeta} \bar{R}_n + \frac{1-\theta_F\gamma}{\gamma}\frac{w_{Fn}L_{Fn}}{R_n}\delta\left(R_n - \bar{R}_n\right)}$.

In the extended model that incorporates N_r rural regions, there are 5 unknown variables $\{\hat{R_n}, \hat{L_{Fn}},$

$\hat{w_{Fn}}, \hat{\tilde{p}}_{Fn}, \hat{\tilde{p}}_{Mn}\}$ for each location pinned down by (52),(53),(55), (57), and the farmland supply function:

$$\left(\pi_{r6n} - \pi_{r7n}\hat{R_n}\right)^{-\zeta} - \pi_{r8n}\frac{\hat{w_{Fn}}\hat{L_{Fn}}\left(\hat{\tilde{p}}_{Fn}\right)^{-\mu}\left(\hat{\tilde{p}}_{Mn}\right)^{\mu-1}}{\hat{R_n}} = 0, \tag{61}$$

where $\pi_{r6n} = \frac{(1-\phi_n)\tilde{R}_n}{(1-\phi_n)\tilde{R}_n - R_n}$, $\pi_{r7n} = \frac{R_n}{(1-\phi_n)\tilde{R}_n - R_n}$, and $\pi_{r8n} = \frac{\frac{1-\theta_F\gamma}{\gamma}\delta w_{Fn}L_{Fn}}{\delta p_{Rn}\left(R_n - \bar{R}_n\right) + \tilde{p}_{Fn}^{\mu}\tilde{p}_{Mn}^{(1-\mu)}c_n\left(1-\frac{R_n}{(1-\phi_n)\tilde{R}_n}\right)^{-\zeta}\bar{R}_n}$.

Next, I outline the iterative algorithm used to solve for the equilibrium of the model.

1. Guess the initial values of $[\hat{\tilde{p}}_{Fn}, \hat{\tilde{p}}_{Mn}, \hat{L_{Mn}}, \hat{w_{Mn}}, \hat{L_{Fn}}, \hat{w_{Fn}}]^0$.

2. At given values of $[\hat{\tilde{p}}_{Fn}, \hat{\tilde{p}}_{Mn}, \hat{L_{Mn}}, \hat{w_{Mn}}, \hat{L_{Fn}}, \hat{w_{Fn}}]$, use (59) and (60) (and (61) in the extended model) to solve for the unique solution of $[\hat{R_n}, \hat{H_n}]^*$.

3. Given $[\hat{R_n}, \hat{H_n}]^*$, search for $[\hat{\tilde{p}}_{Mn}, \hat{\tilde{p}}_{Fn}, \hat{L_{Mn}}, \hat{w_{Mn}}, \hat{L_{Fn}}, \hat{w_{Fn}}, \hat{\bar{V}}]^*$ that satisfy the rest of equilibrium conditions.

4. Stop iteration if

   $||[\hat{\tilde{p}}_{Mn}, \hat{\tilde{p}}_{Fn}, \hat{L_{Mn}}, \hat{w_{Mn}}, \hat{L_{Fn}}, \hat{w_{Fn}}, \hat{H_n}, \hat{R_n}]^{t+1} - [\hat{\tilde{p}}_{Mn}, \hat{\tilde{p}}_{Fn}, \hat{L_{Mn}}, \hat{w_{Mn}}, \hat{L_{Fn}}, \hat{w_{Fn}}, \hat{H_n}, \hat{R_n}]^{t}|| < \epsilon_{tol}$.

   Otherwise, set $[\hat{\tilde{p}}_{Mn}, \hat{\tilde{p}}_{Fn}, \hat{L_{Mn}}, \hat{w_{Mn}}, \hat{L_{Fn}}, \hat{w_{Fn}}]^{t+1} = \epsilon_{iter}[\hat{\tilde{p}}_{Mn}, \hat{\tilde{p}}_{Fn}, \hat{L_{Mn}}, \hat{w_{Mn}}, \hat{L_{Fn}}, \hat{w_{Fn}}]^t + (1-\epsilon_{iter})[\hat{\tilde{p}}_{Mn}, \hat{\tilde{p}}_{Fn}, \hat{L_{Mn}}, \hat{w_{Mn}}, \hat{L_{Fn}}, \hat{w_{Fn}}]^*$ for some $\epsilon_{iter} \in (0,1)$ and go back to step 2.

Note that the equilibrium system is only defined to scale (it is homogenous of degree zero), I normalize the geometric mean of change in CPI to one.

A.5.2 Add a Trading Platform to the Economy

Having the trading platform introduces one more equilibrium condition and the system cannot be directly solved using hat algebra. Therefore, I simulate the counterfactual outcome by using the equilibrium conditions of the new setting directly. After expressing $[p_{Fn}, p_{Mn}, \tilde{p}_n, p_{Hn}, p_{Rn}, R_{Hn}, R_{Fn}]$ as functions of the other variables, there are 10 equilibrium conditions with 10 vectors: $[\bar{V}, \tilde{p}_{Fn}, \tilde{p}_{Mn}, L_{Mn}, L_{Fn}, w_{Mn}, w_{Fn}, H_n, R_n, c_F]$. I denote the upper bound of c_F as $c_{F,u}$ and set it to be the maximum of $c_{Rn} - p_{Rn}$ in the data. Correspondingly, the lower bound of c_F is denoted as $c_{F,l}$ and set to be the minimum of $c_{Rn} - p_{Rn}$ in the data. I outline the steps to search for the solution below.

1. Start with an initial guess of c_F between $c_{F,u}$ and $c_{F,l}$.

2. At given c_F, search for $[\bar{V}, \tilde{p}_{Fn}, \tilde{p}_{Mn}, L_{Mn}, L_{Fn}, w_{Mn}, w_{Fn}, H_n, R_n]^*$ that meet all the equilibrium conditions except (26).

   (a) Start with an initial guess of $[\tilde{p}_{Fn}, \tilde{p}_{Mn}, L_{Mn}, L_{Fn}, w_{Mn}, w_{Fn}]$.

   (b) At given $[\tilde{p}_{Fn}, \tilde{p}_{Mn}, L_{Mn}, L_{Fn}, w_{Mn}, w_{Fn}]$, solve for the unique $[R_n, H_n]^{**}$ using (22) and (25).

   (c) At given $[R_n, H_n]^{**}$, update $[\tilde{p}_{Fn}, \tilde{p}_{Mn}, L_{Mn}, L_{Fn}, w_{Mn}, w_{Fn}]$ using (27), (4), (6), (11) and (12).

(d) Given $[\tilde{p}_{Fn}, \tilde{p}_{Mn}, L_{Mn}, L_{Fn}, w_{Mn}, w_{Fn}]^{**}$, update $\bar{V}$ using (5).

(e) If $||[\bar{V}, \tilde{p}_{Fn}, \tilde{p}_{Mn}, L_{Mn}, L_{Fn}, w_{Mn}, w_{Fn}, H_n, R_n]^{**} - [\bar{V}, \tilde{p}_{Fn}, \tilde{p}_{Mn}, L_{Mn}, L_{Fn}, w_{Mn}, w_{Fn}, H_n, R_n]^t|| < \epsilon_{tol2}$, set $[R_n]^* = [R_n]^{**}$ and stop iteration. Otherwise, set $[\bar{V}, \tilde{p}_{Fn}, \tilde{p}_{Mn}, L_{Mn}, L_{Fn}, w_{Mn}, w_{Fn}]^{t+1} = [\bar{V}, \tilde{p}_{Fn}, \tilde{p}_{Mn}, L_{Mn}, L_{Fn}, w_{Mn}, w_{Fn}]^*$ and go back to step b.

3. Stop iteration if $|| \sum_n R_n^* - \sum_n \bar{R}_n|| < \epsilon_{tol}$. Otherwise, update $[c_{F,l}, c_{F,u}, c_F]^{t+1}$ through the following algorithm and go back to Step 2.

(a) If $\sum_n R_n \leq \bar{R} - \epsilon_{tol}$, update $c_{F,l}^{t+1} = c_F^t$, $c_{F,u}^{t+1} = c_{Fu}^t$ and $c_F^{t+1} = \frac{1}{2}\left(c_{F,u}^t + c_F^t\right)$.

(b) If $\sum_n R_n \geq \bar{R} + \epsilon_{tol}$, update $c_{F,l}^{t+1} = c_{F,l}^t$, $c_{F,u}^{t+1} = c_F^t$ and $c_F^{t+1} = \frac{1}{2}\left(c_F^t + c_{F,l}^t\right)$.

A.5.3 Add a Trading Platform with Price Differentiation Feature

To model the cap-and-trade platform with a price differentiation feature, I denote κ_n as the price premium faced by location n. κ_n follows the specification announced in Notice of the General Office of the State Council [2018] No.16. As before, the landlords receive additional payoff c_F from the trading platform for each unit of farmland above the minimum quantity $\bar{R}_n$. What is different is that, if the landlord develops less farmland than the minimum quantity of $\bar{R}_n$, she pays the platform $\kappa_n c_F$ for each unit of shortage. The price differentiation feature indicates that there is a positive profit left on the platform as long as at least one location facing $\kappa_n > 1$ buys farmland. I assume that the profit from the cap-and-trade platform is uniformly re-distributed between landlords.

In this alternative equilibrium, a location is in one of the following three cases. The first case is that the quantity of farmland is below the minimum quantity and she pays $\kappa_n c_F$ for each unit of farmland developed by another location:

$$c_{Rn} = p_{Rn} + \kappa_n c_F. \tag{62}$$

The second case is that she creates more than the required minimum amount of farmland and receives c_F for each extra unit of farmland:

$$c_{Rn} = p_{Rn} + c_F. \tag{63}$$

The last case is that it is not profitable either to create extra farmland or to reduce farmland to below the minimum quantity. In this case, the landlord just creates the minimum amount of farmland as required, and the following condition holds:

$$p_{Rn} + \kappa_n c_F > c_{Rn} > p_{Rn} + c_F. \tag{64}$$

c_F makes the total amount of farmland across locations equal the targeted minimum amount of farmland at the national level, and therefore (26) holds. Next, there is payment between landlords and I denote the payment to location n as E_{Rn}:

$$E_{Rn} = \left(R_n - \bar{R}_n\right)\left((\kappa_n - 1)1_{R_n < \bar{R}_n} + 1\right) c_F + e_R, \tag{65}$$

where e_R is the uniform redistribution of profit from the trading platform.

The algorithm to solve the new equilibrium is similar to the one when all the locations pay the same price. The only difference is at Step 2 (b). The farmland supply function depends on whether a location is selling, buying or neither buying or selling. I calculate both the farmland quantity if the location buy farmland hence use (62) to choose the farmland supply and the farmland quantity if the location sell farmland hence use (63) to choose the farmland supply. If farmland quantity in the first case is lower than the minimum level, the location buy on the platform. If farmland quantity in the second case is greater than the minimum level, the location sell on the platform. If none is the case, this location chooses the amount of farmland right at the minimum level. The amount of urban land can be derived once the farmland quantity is solved.

A.6 Model Extension: Landlord as a Monopoly in the Urban Land Market

In this model extension, I assume that the representative landlord in a location is a monopoly in the local urban land market. When making the urban land development decision, she takes into account that a lower urban land price would attract more workers to come to the local urban sector. This is to approximate the reality that the local government - the decision-maker of urban land development - would like to keep the urban land price low to attract firms to open new business (Yang et al., 2015). I assume that the representative landlord takes the farmland price as given, as in the baseline model.

For the representative landlord in location n, the demand for urban land is $p_{Hn}H_n = w_{Mn}L_{Mn}$. The landlord takes into account that the number of urban workers in a location, L_{Mn}, is endogenous to p_{Hn}. Specifically, from (6), $L_{Mn} = \bar{L}\bar{V}^{-\nu}\left(w_{Mn}\tilde{p}_n^{-\theta}\bar{B}_{Mn}\right)^{\nu} p_{Hn}^{-(1-\theta)\nu}$. Therefore, the demand for urban land becomes

$$p_{Hn}H_n^{\frac{1}{1+(1-\theta)\nu}} = k_n, \tag{66}$$

where $k_n \equiv \left(\bar{L}\bar{V}^{-\nu}\tilde{p}_n^{-\theta\nu}\bar{B}_{Mn}^{\nu}w_{Mn}^{1+\nu}\right)^{\frac{1}{1+(1-\theta)\nu}}$. k_n is assumed to be exogenous to the representative landlord, because all the components either depend on exogenous local fundamentals or the economic conditions of other locations in the economy.[60]

Without the Farmland Red Line Policy, the landlord's problem is to choose H_n and R_n to maximize the total land profit Π_n:

$$\Pi_n = p_{Hn}H_n + p_{Rn}R_n - \int_0^{H_n+R_n} \tilde{p}_n f(\psi_n)\left(1 - \frac{x}{(1-\phi_n)\tilde{R}_n}\right)^{-\zeta} dx, \tag{67}$$

subject to the urban land demand function (66) and the exogenous value of farmland price p_{Rn}. The first order conditions lead to the following urban land and farmland supply decisions:

$$p_{Hn} = \frac{1+(1-\theta)\,\nu}{(1-\theta)\,\nu}\tilde{p}_n f(\psi_n)\left(1 - \frac{H_n+R_n}{(1-\phi_n)\tilde{R}_n}\right)^{-\zeta}, \tag{68}$$

and

$$p_{Rn} = \tilde{p}_n f(\psi_n)\left(1 - \frac{H_n+R_n}{(1-\phi_n)\tilde{R}_n}\right)^{-\zeta}. \tag{69}$$

[60](66) is derived by plugging the supply function of the manufacturing labor into the urban land demand function.

When $\lambda_{Hn} = \frac{1+(1-\theta)\nu}{(1-\theta)\nu}$, the urban land supply function (22) is the same as (68).

A.7 Model Extension: Redistribute Land Development Profit between Workers

In this model extension, I assume a national portfolio that aggregates the land rents of the whole economy. The profit in the national portfolio is re-distributed equally among workers. This is an alternative way of incorporating the general equilibrium effects of the land profit without introducing heterogeneous wealth effects or inefficiencies due to the externalities in the labor migration (Redding and Rossi-Hansberg, 2017).

With this alternative assumption, the labor income now becomes

$$\tilde{w}_{sn} = w_{sn} + \pi_r, \tag{70}$$

where π_r is the transfer from the national portfolio. π_r satisfies the following relation:

$$\pi_r \bar{L} = \sum_n \left(p_{Rn} R_n + p_{Hn} H_n - CC_n \right), \tag{71}$$

and CC_n represents the total land development cost in location n:

$$CC_n = \int_0^{R_n + H_n} \tilde{p}_n c_n \left(1 - \frac{x}{(1-\phi_n)\tilde{R}_n} \right)^{-\zeta} dx. \tag{72}$$

The labor supply function becomes

$$L_{sn} = \bar{L} \bar{V}^{-\nu} \left(\tilde{w}_{sn} \tilde{p}_n^{-\theta} p_{H,sn}^{\theta-1} \bar{B}_{sn} \right)^{\nu}, \tag{73}$$

and the expected utility of a representative worker now is

$$\bar{V} = \left(\sum_{s' \in \{F,M\}} \sum_{n' \in N} \left(\tilde{w}_{s'n'} \tilde{p}_{n'}^{-\theta} p_{H,s'n'}^{\theta-1} \bar{B}_{sn} \right)^{\nu} \right)^{\frac{1}{\nu}}. \tag{74}$$

The aggregate demand for urban land in location n is

$$(1-\theta)\tilde{w}_{Mn} L_{Mn} = p_{Hn} H_n. \tag{75}$$

The aggregate demand for farmland used for residential purpose in location n is

$$(1-\theta)\tilde{w}_{Fn} L_{Fn} = p_{Rn} R_{Hn}. \tag{76}$$

Last, the tradable goods market clearing condition becomes

$$\frac{p_{Fn}Y_{Fn}}{\mu} = \theta \sum_{j \in N} \frac{(T_{jn})^{1-\sigma_F}}{\tilde{p}_{Fj}^{1-\sigma_F}} \tilde{w}_{Mj} L_{Mj} + \theta \sum_{j \in N} \frac{(T_{jn})^{1-\sigma_F}}{\tilde{p}_{Fj}^{1-\sigma_F}} \tilde{w}_{Fj} L_{Fj} + \sum_{j \in N} \frac{(T_{jn})^{1-\sigma_F}}{\tilde{p}_{Fj}^{1-\sigma_F}} CC_j, \tag{77}$$

$$\frac{p_{Mn}Y_{Mn}}{1-\mu} = \theta \sum_{j \in N} \frac{(T_{jn})^{1-\sigma_M}}{\tilde{p}_{Mj}^{1-\sigma_M}} \tilde{w}_{Mj} L_{Mj} + \theta \sum_{j \in N} \frac{(T_{jn})^{1-\sigma_M}}{\tilde{p}_{Mj}^{1-\sigma_M}} \tilde{w}_{Fj} L_{Fj} + \sum_{j \in N} \frac{(T_{jn})^{1-\sigma_M}}{\tilde{p}_{Mj}^{1-\sigma_M}} CC_j. \tag{78}$$

I take this alternative model to data and simulate the counterfactual equilibrium without the Farmland Red Line Policy. The quantitative results are shown in Table 48.

A.8 Model Extension: Make Farmland Density Affect the Cost of Urban Land Development

The baseline model is isomorphic to one in which urban land is converted from farmland or from a combination of farmland and unused land. This is because in the baseline model, to convert a piece of unused land plot into farmland or urban land takes the same cost. Therefore, assuming that urban land is converted from unused land is equivalent to assuming that the unused land is first converted into farmland and is then converted from farmland into urban land with no additional cost. And the latter is the same as assuming that urban land is converted from farmland at not additional cost. As a result, assuming that urban land is converted from farmland is the same as assuming that urban land is converted from unused land.

A consequence of this feature of the model is that introducing the Farmland Red Line Policy creates the same degree of distortion to city jurisdictions that use a different combination of farmland and unused land to develop urban land. However, as discussed in the empirical analysis, new urban land is converted from rural land surrounding the existing urban area in reality. If the rural land next to the existing urban area is unused land, the city creates new urban land using unused land only. In this case, the policy will not apply at all and thus has a limited impact on the local economy. In contrast, if the surrounding rural land is all farmland, then all the new urban land is converted from farmland. Without the policy, the farmland amount would decrease as analyzed earlier.

To incorporate this margin into the model, I assume that in location n, urban land production function is a Leontiff function of both unused land and farmland. Specifically, a unit of urban land is produced through converting λ_{Fn} unit of farmland and $1 - \lambda_{Fn}$ unit of unused land into urban land. Second, I assume that converting an unused land plot l into urban land costs $(f(\psi_n)x_{nl})^{\lambda_U}$ units of consumption goods bundle, where $\lambda_U < 1$, while converting the same land plot into farmland costs $f(\psi_n)x_{nl}$ units of consumption goods bundle as in the baseline model. Under this new assumption, converting a piece of land into urban land is less costly than converting the same piece of land into farmland. $\lambda_U < 1$ indicates that the cost to convert a land plot into farmland is more dispersed across land plots than the cost to convert unused land into urban land. This is because farmland is more demanding in terms of the subtle features of the soil compared to urban land. The suitability of land for crop growth depends on multiple conditions, such as suitable pH range, rich nutrition, functional water storage capacity, and flat terrain surface. A shortage in any one dimension would make it more costly to cultivate the land plot. However, urban land development mainly depends on the flatness of the terrain surface. Therefore, the marginal cost to convert unused land into farmland should increase more quickly than the marginal cost to convert unused land into urban land.

A second change is that the representative landlord only has the technology to convert the unused land into farmland. There is a new agent - a representative urban land developer - who has the Leontiff technology to combine farmland and unused land to produce urban land. The urban land developer buys farmland and unused land plots from the landlord to produce urban land. She gives λ_b fraction of the urban land development profit to the landlord as a payment for the unused land plots. Both the landlord and the urban land developer takes the land prices as given. Notice that the change here is a generalization of the baseline model, because when $\lambda_b = 1$, the model is equivalent to assuming that the landlord makes both the farmland development and urban land development decision.

The landlord has incentive to use low x_{nl} land plots for farmland production and rent the high x_{nl} land plots to the urban land developer. To see this, suppose there are two land plots l and l' with different land development costs, $f(\psi_n)x_{nl} < f(\psi_n)x_{nl'}$. Land plot l is rented by the urban land developer and converted into urban land while the land plot l' is converted into farmland. If she instead use the land plot l for farmland development and rent out land plot l, her revenue from farmland development is the same while the cost reduces by $f(\psi_n)x_{nl'} - f(\psi_n)x_{nl}$. On the other hand, her payoff from urban land decreases by $\lambda_b \left((f(\psi_n)x_{nl'})^{\lambda_U} - (f(\psi_n)x_{nl})^{\lambda_U})\right)$. Because $f(\psi_n)x_{nl'} - f(\psi_n)x_{nl} > \lambda_b \left((f(\psi_n)x_{nl'})^{\lambda_U} - (f(\psi_n)x_{nl})^{\lambda_U})\right)$ for any $\lambda_b \leq 1$, she would never use the land plot l' for farmland development and rent the land plot l to the urban land developer.

At the margin that the landlord is indifferent in converting the land plot into farmland and renting the same land plot to the urban land developer,

$$p_{Rn} - f(\psi_n)\left(1 - \frac{R_n + \lambda_{Fn}H_n}{(1-\phi_n)\tilde{R}_n}\right)^{-\zeta} = \lambda_b \frac{p_{Hn} - \lambda_{Fn}p_{Rn}}{1-\lambda_{Fn}} - \lambda_b \tilde{p}_n f(\psi_n)^{\lambda_U}\left(1 - \frac{R_n + \lambda_{F_n}H_n}{(1-\phi_n)\tilde{R}_n}\right)^{-\zeta\lambda_U}. \tag{79}$$

The left hand side of (79) represents the profit from developing the land plot into farmland. Notice that in this case, $\lambda_{Fn}H_n$ units of the farmland are later rented by the urban land developer and converted into urban land. Therefore, only R_n units of farmland are used in the agricultural sector. Next, the right hand side represents the proportion of the profit of using the same land plot in urban land development that goes to the landlord. The profit of converting the marginal land plot into urban land is derived in the following way. Given the unit urban land price, p_{Hn}, the proportion of farmland in urban land production, λ_{Fn}, and the unit price of farmland, p_{Rn}, the revenue left is $\frac{p_{Hn}-\lambda_{Fn}p_{Rn}}{1-\lambda_{Fn}}$. After subtracting the cost of converting that marginal land plot into urban land, the profit is $\frac{p_{Hn}-\lambda_{Fn}p_{Rn}}{1-\lambda_{Fn}} - \tilde{p}_n f(\psi_n)^{\lambda_U}\left(1 - \frac{R_n+\lambda_{F_n}H_n}{(1-\phi_n)\tilde{R}_n}\right)^{-\zeta\lambda_U}$.

Next, the urban land developer would keep renting unused land and farmland from the landlord and convert them into urban land until it is non-profitable. At the margin, it must be that the profit of developing the marginal unused land into urban land equals zero:

$$\frac{p_{Hn} - \lambda_{Fn}p_{Rn}}{1-\lambda_{Fn}} - \tilde{p}_n f(\psi_n)^{\lambda_U}\left(1 - \frac{R_n + H_n}{(1-\phi_n)\tilde{R}_n}\right)^{-\zeta\lambda_U} = 0. \tag{80}$$

(79) and (80) are the urban land supply function and the farmland supply function in the market equilibrium.

The Farmland Red Line Policy requires that the urban land developer has to guarantee that the farmland used in the agricultural production is not below $\bar{R}_n$ when converting farmland into urban land. If at the market price of farmland, p_{Rn}, the landlord has no incentive to supply above $\bar{R}_n$ units of farmland for agricultural use, the urban land developer needs to pay extra compensation for per unit of farmland developed. Therefore, condition (79) becomes the following complementary slackness condition:

$$\lambda_b \frac{p_{Hn} - \lambda_{Fn} p_{Rn}}{1-\lambda_{Fn}} - \lambda_b \tilde{p}_n f(\psi_n)^{\lambda_U} \left(1 - \frac{R_n + \lambda_{F_n} H_n}{(1-\phi_n)\tilde{R}_n}\right)^{-\zeta\lambda_U} \geq p_{Rn} - \tilde{p}_n f(\psi_n) \left(1 - \frac{R_n + \lambda_{Fn} H_n}{(1-\phi_n)\tilde{R}_n}\right)^{-\zeta},$$

$$R_n \geq \bar{R}_n,$$

$$\left(\frac{\lambda_b p_{Hn} - (1 + \lambda_b \lambda_{Fn} - \lambda_{Fn}) p_{Rn}}{1-\lambda_{Fn}} - \lambda_b \tilde{p}_n f(\psi_n)^{\lambda_U} \left(1 - \frac{R_n + \lambda_{F_n} H_n}{(1-\phi_n)\tilde{R}_n}\right)^{-\zeta\lambda_U} + f(\psi_n) \left(1 - \frac{R_n + \lambda_{Fn} H_n}{(1-\phi_n)\tilde{R}_n}\right)^{-\zeta}\right) \times \left(R_n - \bar{R}_n\right) = 0. \tag{81}$$

The marginal condition in which the urban land developer is indifferent between converting the land plot into urban land or keep it as undeveloped is the same as (80).

Before characterizing the general equilibrium of this extended version of the model, I use (80) to explain why this extension makes high λ_{Fn} locations have more distortion in urban land price when there is the Farmland Red Line Policy. The urban land price is a weighted average of the marginal cost of farmland development and the marginal cost of urban land development. A minimum farmland quantity constraint pushes up the marginal farmland development cost. It also pushes up the marginal cost to convert unused land into urban land, but increase is smaller given $\lambda_U < 1$. If a higher weight is loaded to the marginal farmland development cost, the urban land price would increase more when a minimum farmland quantity constraint is introduced. Therefore, urban land price is more distorted when λ_{Fn} is higher.

The general equilibrium is defined in the same way as in the baseline, except that the farmland supply functions and urban land supply functions changed into the (81) and (80) with the policy or (79) and (80) without the policy. This model extension introduces two parameters λ_U and λ_b. I specify λ_b to be zero to make the marginal farmland development cost equal the farmland price, which approximates the marginal condition of farmland development in reality, as discussed in 5.1.4. In the quantitative exercise, I provide counterfactual results under a series of values of λ_U from as low as 0.1 to as high as 0.9. The quantitative results are provided in Appendix Table 48.

B Data Appendix

B.1 Additional Tables

B.1.1 Additional Tables for Section 3

Table 12: Correlation between the Change in Urban land and the Change in Farmland

Depvar: DR_{00to10} DR	(1)	(2)	(3)
DH_{00to10}	-0.031	-0.065	
	(0.034)	(0.189)	
DH_{99to10}			-0.045
			(0.034)
Observations	631	631	631
R-squared	0.001	0.000	0.003

Notes. This table shows that there is little correlation between the change in urban land and the change in farmland across locations. The results suggest that the Farmland Red Line Policy successfully stops farmland decline during urbanization. The outcome variable in Column 1 and 3 are calculated using GeoExplorer II. The outcome variable in Column 2 is calculated using the MODIS farmland raster data. Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 13: Distance to the Administrative Boundary and the Conversion from Non-farmland to Farmland

Depvar: $1_{\text{nonfarm to farm}}$	(1)	(2)	(3)
Distance to the administrative boundary	-0.0005***	-0.0005***	-0.208***
	(0.0000)	(0.0000)	(0.012)
Observations	5,794,550	5,794,550	5,794,550
Specification	OLS	OLS	logit
City FE	No	Yes	No

Notes. This table shows that in a city jurisdiction, land grids closer to the administrative boundary are more likely to change from non-farmland to farmland during 2000 and 2010. Furthermore, 1 km further away from the administrative boundary reduces the probability of a grid to change from non-farmland into farmland by around 20%. It indicates that land grids more than 5 km away from the administrative boundary of a city jurisdiction is very unlikely to be used for new farmland development. To conduct the analysis, I use land cover and land use database at 500-m resolution to create a raster

dataset of whether at each grid, the land cover changes from non-farmland into farmland from 2001 to 2010. The data comes from NASA MODIS MCD12Q2. The database classifies every 500-m grid on the map into one of the 17 land use categories, such as farmland, urban land, grassland, and forest. Column 1 conducts a simple OLS regression without any controls. The results are robust to controlling for the city jurisdiction fixed effects (Column 2). Column 3 conducts a logit regression to make the interpretation of the coefficient transparent. The error terms are clustered at the city jurisdiction level. Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 14: Balance Test of the Land Conversion Barrier

	C_u below 50%	C_u above 50%	Difference (sd)
Population growth rate (1982 to 1990)	0.184	0.161	0.0231 (0.0373)
Employment growth rate (1982 to 1990)	0.294	0.265	0.0293 (0.0396)
Non-agricultural employment growth rate (1982 to 1990)	0.513	0.525	-0.0124 (0.0693)
Illiterate population growth rate (1982 to 1990)	-0.289	-0.316	0.0265 (0.0369)
College graduates growth rate (1982 to 1990)	0.125	0.0995	0.0251 (0.0860)
Population (ln of)	13.44	13.22	0.218*** (0.0605)
% college graduates	0.667	0.598	0.0688 (0.0711)
Illiteracy rate	14.96	14.85	0.114 (0.409)
Employment rate	83.75	81.65	2.099** (0.645)
% employment from non-agriculture	43.70	40.95	2.750 (1.916)
% employment from the construction sector	1.829	1.778	0.0516 (0.117)
% in-migration	5.012	5.106	-0.0943 (0.407)

Notes. This table shows that city jurisdictions with different land conversion barriers are balanced along various dimensions of economic and demographic characters in 1990. First of all, locations with different land conversion barriers are quite similar in terms of the growth of population, employment, economic structural change (growth of employment in the non-agricultural sector), and human capital accumulation (change in illiterate population and college graduates) during 1982 and 1990. Second, they are also quite similar over various measures of local economic characteristics in 1990, including the employment structure, education, and in-migration. Locations with a higher land conversion barrier had slightly fewer populations and a lower employment rate in 1990. The regression results in Section 4, 6, and 7 are robust to controlling the time-varying impacts of population and employment rate in 1990. All the regression results are also robust to controlling the time-varying effects of all the characteristics in 1990 reported in the balance test. Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 15: Farmland Density Prediction

Depvar: Actual farmland density	(1)	(2)
Available water storage	0.068***	
	(0.020)	
Water drainage	0.049***	
	(0.016)	
Gravel fraction	0.006**	
	(0.003)	
Organic carbon	0.267***	
	(0.078)	
PH	0.054***	
	(0.017)	
Gypsum	0.261**	
	(0.122)	
Salinity	0.025*	
	(0.014)	
Composite agricultural suitability		0.051***
		(0.016)
Observations	543	543
R-squared	0.067	0.019

Notes. The farmland density is regressed against the mean soil qualities of land within the buffer and get the predicted value of the dependent variable. The baseline specification includes the soil features most crucial for the soil suitability for cultivation, such as organic carbon, pH, salinity, gypsum, gravel fraction, water storage capacity, and soil drainage (Zhang et al., 2010). As shown in Column 1, all the critical chemical and physical features affect farmland density significantly. For robustness checks, instead of using individual soil characteristics measures, I use the nutrient availability index provided by the World Harmonized Soil Database to predict farmland density for robustness checks. This composite index represents the degree of soil suitability for low to intermediate level input farming activity. As shown in Column 2, this alternative index is also positively associated with the farmland density of rural land converted into urban land during the 2000s. The dependent variable is constructed using data from the City Development Yearbooks. Due to missing data issues in the yearbooks, the dependent variable is available for only 543 city jurisdictions. For each soil characteristic, I make a linear transformation such that the higher the value is, the more suitable that land grid is for cultivation. If the value of a soil character already preserves this feature, then no linear transformation is made. I then calculate the average value of each soil characteristic within the projected new urban area. For some characteristics such as organic carbon and pH, topsoil and bottom soil could have different values, and both are reported in the soil database. In this case, I use the minimum of the two as the value for that grid. Results are robust to using the average or the maximum of the two. Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

B.1.2 Impacts of the Land Conversion Barrier on Urban Land Supply, Robustness Checks

Table 16: Detect Urban Land Data Missing Issues and Measurement Errors

Depvar: ln(urban land)	(1)	(2)	(3)	(4)
CuxPost1999	-0.055***	-0.054***	-0.034*	-0.063***
	(0.018)	(0.019)	(0.020)	(0.020)
Observations	13,186	10,188	5,155	11,908
R-squared	0.994	0.995	0.997	0.994
Sample	Baseline	Cities since 1990	208 city subgroup	Exclude boundary change
N of jurisdictions	631	456	208	579

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 17: Drop Politically Favored Cities

Depvar: ln(urban land)	(1)	(2)	(3)	(4)
CuxPost1999	-0.055***	-0.052***	-0.057***	-0.059***
	(0.018)	(0.019)	(0.018)	(0.019)
Observations	13,186	12,582	13,092	12,849
R-squared	0.994	0.994	0.994	0.994
N of jurisdictions	631	606	627	617
Sample	Baseline	No provincial capitals	No prvn cities	No port cities

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 18: Alternative Ways to Predict the Farmland Density

Depvar: ln(urban land)	(1)	(2)	(3)	(4)	(5)	(6)
CuxPost1999	-0.055***	-0.054***	-0.052***	-0.052***	-0.053***	-0.046**
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Observations	13,186	13,186	13,186	13,186	13,186	13,186
R-squared	0.994	0.994	0.994	0.994	0.994	0.994
Soil measure construction	Min	Mean	Max	Min	Min	-
Soil vector	Baseline	Baseline	Baseline	Baseline	All	Composite index
Exclude undevelopable area	No	No	No	Yes	No	No

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 19: Alternative Ways to Measure the Land Ruggedness

Depvar: ln(urban land)	(1)	(2)	(3)	(4)
CuxPost1999	-0.055***	-0.040**	-0.058***	-0.060***
	(0.018)	(0.018)	(0.019)	(0.019)
Observations	13,186	13,186	13,186	13,186
R-squared	0.994	0.994	0.994	0.994
Slope measure	Baseline	> 6 degree slope	Terrain ruggedness index	Average slope

Notes. Column 1 uses the baseline specification for the independent variable. Column 2 uses percent of land grids with a slope of above 6 degrees to measure the land ruggedness. Columns 3 and 4 use the terrain ruggedness index and the average slope measure developed in Nunn and Puga (2012). Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 20: Alternative Specifications of the Projected New Urban Region and the Projected Farmland Development Region

Depvar: ln(urban land)	(1)	(2)	(3)	(4)	(5)
CuxPost1999	-0.055***	-0.055***	-0.055***	-0.058***	-0.049***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Observations	13,186	13,186	13,186	13,186	13,186
R-squared	0.994	0.994	0.994	0.994	0.994
Projected new urban region	2 km buffer	3 km buffer	1 km buffer	circular expansion	2 km buffer
Projected farmland development region	5 km buffer	5 km buffer	5 km buffer	5 km buffer	all

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 21: Control for the Land Ruggedness in the Projected New Urban Region

Depvar: ln(urban land)	(1)	(2)	(3)	(4)
CuxPost1999	-0.055***	-0.055***	-0.053***	-0.077***
	(0.018)	(0.018)	(0.018)	(0.025)
Observations	13,186	13,186	13,186	13,186
R-squared	0.994	0.994	0.994	0.994
Specification	Baseline	Exclude slope variation	Exclude Variation from Drainage and Gravel	Control slope time-varying effect

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 22: Control for the Railway Access, Add Time Trends, and Adjust Standard Errors

Depvar: ln(urban land)	(1)	(2)	(3)	(4)	(5)
CuxPost1999	-0.055***	-0.047***	-0.055***	-0.055***	-0.055***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.017)
Observations	13,186	13,186	13,186	13,186	13,186
R-squared	0.994	0.997	0.994	0.149	0.149
Distance cutoff	-	-	-	500 km	1000 km

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

B.1.3 Impacts of the Land Conversion Barrier on GDP, Robustness Checks

Table 23: Detect Missing Data Issues and Measurement Error Issues

	(1)	(2)	(3)	(4)	(5)
ln of GDP	All	Agriculture	Non-agriculture	Secondary	Tertiary
	Panel A. Balanced Panel				
CuxPost1999	-0.035**	-0.022	-0.043**	-0.061***	-0.023
	(0.017)	(0.021)	(0.018)	(0.021)	(0.018)
Observations	9,828	9,828	9,828	9,828	9,828
R-squared	0.998	0.984	0.998	0.995	0.996
N of jurisdictions	456	456	456	456	456
	Panel B. Jurisdictions without any Change in Administrative Boundary				
CuxPost1999	-0.037**	-0.016	-0.039**	-0.056***	-0.020
	(0.015)	(0.015)	(0.016)	(0.019)	(0.017)
Observations	12,419	12,419	12,419	12,419	12,419
R-squared	0.998	0.990	0.997	0.994	0.996
N of jurisdictions	579	579	579	579	579

Notes. The regression results are not affected by data missing issues or measurement error issues. In Panel A, I restrict the sample to the city jurisdictions that have data records during 1994 and 1998 and hence do not suffer from missing data problems. In Panel B, I exclude city jurisdictions that have changed the administrative boundary by incorporating a nearby county from the regression. These jurisdictions may bring in a spurious positive correlation between the growth of urban land area and economic growth simply because of the boundary change. In both cases, the estimated impact on GDP and population are close to the baseline results. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 24: Drop Politically Favored City Jurisdictions

ln of GDP	(1) All	(2) Agriculture	(3) Non-agriculture	(4) Secondary	(5) Tertiary
	Panel A. Drop 25 Provincial Capitals				
CuxPost1999	-0.035**	-0.020	-0.038**	-0.057***	-0.017
	(0.015)	(0.016)	(0.016)	(0.018)	(0.016)
Observations	13,016	13,016	13,016	13,016	13,016
R-squared	0.998	0.987	0.997	0.994	0.996
N of jurisdictions	606	606	606	606	606
	Panel B. Drop 4 Provincial Level Cities				
CuxPost1999	-0.040***	-0.028*	-0.042***	-0.060***	-0.022
	(0.014)	(0.017)	(0.015)	(0.018)	(0.016)
Observations	13,465	13,465	13,465	13,465	13,465
R-squared	0.998	0.987	0.997	0.995	0.996
N of jurisdictions	627	627	627	627	627
	Panel C. Drop Port Cities				
CuxPost1999	-0.040***	-0.023	-0.042***	-0.060***	-0.022
	(0.015)	(0.017)	(0.016)	(0.018)	(0.016)
Observations	13,252	13,252	13,252	13,252	13,252
R-squared	0.998	0.987	0.997	0.995	0.996
N of jurisdictions	617	617	617	617	617

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 25: Alternative Ways to Predict the Farmland Density

ln of GDP	(1) All	(2) Agriculture	(3) Non-agriculture	(4) Secondary	(5) Tertiary
	Panel A. Mean of the Top and Bottom Soil				
CuxPost1999	-0.038***	-0.030*	-0.041***	-0.059***	-0.020
	(0.015)	(0.017)	(0.016)	(0.018)	(0.016)
Observations	13,552	13,552	13,552	13,552	13,552
R-squared	0.998	0.987	0.997	0.995	0.996
	Panel B. Max of the Top and Bottom Soil				
CuxPost1999	-0.038**	-0.033*	-0.040**	-0.058***	-0.020
	(0.015)	(0.018)	(0.016)	(0.018)	(0.016)
Observations	13,552	13,552	13,552	13,552	13,552
R-squared	0.998	0.987	0.997	0.995	0.996
	Panel C. Exclude Undevelopable Areas				
CuxPost1999	-0.038***	-0.027	-0.040***	-0.059***	-0.020
	(0.014)	(0.016)	(0.015)	(0.018)	(0.016)
Observations	13,552	13,552	13,552	13,552	13,552
R-squared	0.998	0.987	0.997	0.995	0.996
	Panel D. Use All the Soil Characteristics				
CuxPost1999	-0.036**	-0.029*	-0.039**	-0.058***	-0.018
	(0.014)	(0.016)	(0.015)	(0.018)	(0.016)
Observations	13,552	13,552	13,552	13,552	13,552
R-squared	0.998	0.987	0.997	0.995	0.996
	Panel E. Use the Index of the Soil Suitability for Cultivation				
CuxPost1999	-0.036**	-0.038**	-0.039**	-0.056***	-0.021
	(0.015)	(0.017)	(0.016)	(0.019)	(0.016)
Observations	13,552	13,552	13,552	13,552	13,552
R-squared	0.998	0.987	0.997	0.995	0.996

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 26: Alternative Ways to Measure the Land Ruggedness

ln of GDP	(1) All	(2) Agriculture	(3) Non-agriculture	(4) Secondary	(5) Tertiary
	Panel A. % of surface with a >6 degree slope to measure ruggedness				
CuxPost1999	-0.044***	-0.034*	-0.049***	-0.065***	-0.029*
	(0.015)	(0.018)	(0.016)	(0.019)	(0.016)
Observations	13,552	13,552	13,552	13,552	13,552
R-squared	0.998	0.987	0.997	0.995	0.996
	Panel B. Ruggedness Measure from Nunn and Puga (2012)				
CuxPost1999	-0.037***	-0.023	-0.036**	-0.051***	-0.021
	(0.014)	(0.016)	(0.015)	(0.018)	(0.015)
Observations	13,552	13,552	13,552	13,552	13,552
R-squared	0.998	0.987	0.997	0.995	0.996
	Panel C. Average Slope from Nunn and Puga (2012)				
CuxPost1999	-0.037***	-0.022	-0.036**	-0.051***	-0.021
	(0.014)	(0.016)	(0.015)	(0.018)	(0.015)
Observations	13,552	13,552	13,552	13,552	13,552
R-squared	0.998	0.987	0.997	0.995	0.996

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 27: Alternative Specifications of the Projected New Urban Region and the Projected Farmland Development Region

ln of GDP	(1) All	(2) Agriculture	(3) Non-agriculture	(4) Secondary	(5) Tertiary
	Panel A.Specify a 1-km Outward Buffer as the Projected New Urban Region				
CuxPost1999	-0.039***	-0.027	-0.041***	-0.060***	-0.020
	(0.014)	(0.016)	(0.016)	(0.018)	(0.016)
Observations	13,552	13,552	13,552	13,552	13,552
R-squared	0.998	0.987	0.997	0.995	0.996
	Panel B. Specify a 3-km Outward Buffer as the Projected New Urban Region				
CuxPost1999	-0.038***	-0.027	-0.040**	-0.058***	-0.020
	(0.014)	(0.017)	(0.015)	(0.018)	(0.016)
Observations	13,552	13,552	13,552	13,552	13,552
R-squared	0.998	0.987	0.997	0.995	0.996
	Panel C. Circular Expansion of Urban Area				
CuxPost1999	-0.039***	-0.024	-0.041***	-0.059***	-0.020
	(0.015)	(0.016)	(0.016)	(0.018)	(0.016)
Observations	13,552	13,552	13,552	13,552	13,552
R-squared	0.998	0.987	0.997	0.995	0.996
	Panel D. All Rural Areas As the Farmland Development Region				
CuxPost1999	-0.037***	-0.029*	-0.040***	-0.058***	-0.020
	(0.014)	(0.017)	(0.015)	(0.018)	(0.016)
Observations	13,552	13,552	13,552	13,552	13,552
R-squared	0.998	0.987	0.997	0.995	0.996

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 28: Control for the Land Ruggedness in the Projected New Urban Area

ln of GDP	(1) All	(2) Agriculture	(3) Non-agriculture	(4) Secondary	(5) Tertiary
	Panel A. Exclude Land Ruggedness from the Predicted Farmland Density				
CuxPost1999	-0.039***	-0.027	-0.041***	-0.060***	-0.021
	(0.014)	(0.016)	(0.015)	(0.018)	(0.016)
Observations	13,552	13,552	13,552	13,552	13,552
R-squared	0.998	0.987	0.997	0.995	0.996
	Panel B. Exclude Variation from Gravel and Drainage				
CuxPost1999	-0.040***	-0.031*	-0.042***	-0.060***	-0.022
	(0.015)	(0.017)	(0.016)	(0.018)	(0.016)
Observations	13,552	13,552	13,552	13,552	13,552
R-squared	0.998	0.987	0.997	0.995	0.996
	Panel C. Control time-varying Effects of the Land Ruggedness in the Projected Urban Area				
CuxPost1999	-0.037**	-0.017	-0.037*	-0.061***	-0.015
	(0.018)	(0.025)	(0.020)	(0.023)	(0.019)
Observations	13,552	13,552	13,552	13,552	13,552
R-squared	0.998	0.987	0.997	0.995	0.996

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 29: Control for the Railway Access and Adjust Standard Errors

ln of GDP	(1) All	(2) Agriculture	(3) Non-agriculture	(4) Secondary	(5) Tertiary
	Panel A. Control for the time-varying Effects of Railway Access				
CuxPost1999	-0.039***	-0.027	-0.041***	-0.060***	-0.021
	(0.014)	(0.017)	(0.015)	(0.018)	(0.016)
Observations	13,552	13,552	13,552	13,552	13,552
R-squared	0.998	0.987	0.997	0.995	0.996
	Panel B. Apply the Conley Spatial Clustering				
CuxPost1999	-0.039***	-0.027	-0.041**	-0.060***	-0.021
	(0.015)	(0.017)	(0.016)	(0.018)	(0.017)
Observations	13,552	13,552	13,552	13,552	13,552
R-squared	0.072	0.074	0.057	0.058	0.052
distance cutoff	500 km	500 km	500 km	500 km	500 km

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

B.1.4 Impacts of the Land Conversion Barrier on Population, Robustness

Table 30: Drop Politically Favored City Jurisdictions

	(1)	(2)	(3)	(4)
	Panel A. Population			
CuxPost1999	-0.045***	-0.042***	-0.045***	-0.048***
	(0.012)	(0.012)	(0.012)	(0.012)
	Panel B. Employment			
CuxPost1999	-0.045***	-0.042***	-0.045***	-0.049***
	(0.012)	(0.012)	(0.012)	(0.012)
Observations	2,524	2,424	2,508	2,468
N of jurisdictions	631	606	627	617
Sample	Baseline	No provincial capitals	No prvn cities	No port cities

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 31: Alternative Ways to Predict the Farmland Density

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. Population					
CuxPost1999	-0.045***	-0.045***	-0.044***	-0.042***	-0.046***	-0.036***
	(0.012)	(0.012)	(0.013)	(0.012)	(0.012)	(0.014)
	Panel B. Employment					
CuxPost1999	-0.045***	-0.046***	-0.045***	-0.041***	-0.049***	-0.039***
	(0.012)	(0.012)	(0.013)	(0.012)	(0.012)	(0.014)
Observations	2,524	2,524	2,524	2,524	2,524	2,524
Soil measure construction	Min	Mean	Max	Min	Min	Min
Soil vector	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
Exclude undevelopable area	No	No	No	Yes	Yes	Yes

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 32: Alternative Ways to Measure the Land Ruggedness

	(1)	(2)	(3)	(4)
	Panel A. Population			
CuxPost1999	-0.045***	-0.054***	-0.037***	-0.039***
	(0.012)	(0.013)	(0.012)	(0.012)
	Panel B. Employment			
CuxPost1999	-0.045***	-0.062***	-0.032**	-0.034***
	(0.012)	(0.013)	(0.013)	(0.013)
Observations	2,524	2,524	2,524	2,524
Slope measure	Baseline	> 6 degree	Terrain ruggedness index	Average slope

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 33: Alternative Specifications of the Projected New Urban Region and the Projected Farmland Development Region

	(1)	(2)	(3)	(4)	(5)
	Panel A. Population				
CuxPost1999	-0.045***	-0.043***	-0.045***	-0.047***	-0.045***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
	Panel B. Employment				
CuxPost1999	-0.045***	-0.043***	-0.044***	-0.047***	-0.048***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Observations	2,524	2,524	2,524	2,524	2,524
new urban area	2km buffer	3km buffer	1km buffer	circular expansion	2km buffer
farm creation area	5 km buffer	5 km buffer	5 km buffer	5 km buffer	all

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 34: Control for the Land Ruggedness in the Projected New Urban Area

	(1)	(2)	(3)	(4)
	Panel A. Population			
CuxPost1999	-0.045***	-0.045***	-0.043***	-0.046***
	(0.012)	(0.012)	(0.013)	(0.017)
	Panel B. Employment			
CuxPost1999	-0.045***	-0.045***	-0.045***	-0.033*
	(0.012)	(0.012)	(0.013)	(0.017)
Observations	2,524	2,524	2,524	2,524
R-squared	1.000	1.000	1.000	1.000
Specification	Baseline	Exclude slope variation	Exclude Variation from Drainage and Gravel	Control slope time-varying effect

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 35: Control for the Railway Access and Adjust Standard Errors

	(1)	(2)	(3)	(4)
	Panel A. Population			
CuxPost1999	-0.045***	-0.045***	-0.045***	-0.045***
	(0.012)	(0.012)	(0.011)	(0.012)
	Panel B. Employment			
CuxPost1999	-0.045***	-0.045***	-0.045***	-0.045***
	(0.012)	(0.012)	(0.012)	(0.012)
Observations	2,524	2,524	2,524	2,524
Specification	Baseline	Control Transportation	500 km spatial corr	1000 km spatial corr

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 36: Impacts on Employment by Sector

	(1)	(2)	(3)	(4)
Depvar: ln employment of	All	Agriculture	Non-agriculture	Construction
	Panel A. Data from 1990 to 2010			
CuxPost1999	-0.064***	-0.038	-0.039**	-0.033
	(0.016)	(0.024)	(0.017)	(0.026)
Observations	1,893	1,893	1,893	1,893
R-squared	1.000	0.999	1.000	0.999
	Panel B. Data from 1982 to 2010			
CuxPost1999	-0.065***	-0.030	-0.050***	-0.033
	(0.016)	(0.026)	(0.018)	(0.026)
Observations	2,524	2,524	2,524	1,893
R-squared	1.000	0.999	1.000	0.999

Notes. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

B.1.5 Additional Tables for Section 4.3 and 4.4

Table 37: Association of the Land Conversion Barrier with Urban Land Prices, Robustness

Depvar $\ln P_{it}$	(1) All transactions	(2) Sold through auction	(3) Industrial land	(4) Commercial land
Cu	0.063***	0.128***	0.087***	0.098***
	(0.024)	(0.022)	(0.020)	(0.025)
Observations	4,978	4,906	4,758	4,583
R-squared	0.976	0.984	0.984	0.983
N of jurisdictions	608	606	601	605

Notes. In all the columns, the control variables include: (a) region time-varying effects, (b) the time-varying effects of population, employment rate, illiteracy rate, percent of college graduates, percent of employment from the agricultural sector, percent of employment from construction sector, and percent of immigrants from other city jurisdictions in 1990; (c) the time-varying effects of GDP, GDP from non-agricultural sector and population in 1996. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 38: Association of the Land Conversion Barrier with Urban Land Prices, Specification 2

Depvar $\ln P_{it} - \ln \hat{P}_{1996}$	(1) All transactions	(2) Sold through auction	(3) Industrial land	(4) Commercial land
Cu	0.051**	0.115***	0.068***	0.085***
	(0.024)	(0.023)	(0.021)	(0.026)
Observations	4,978	4,906	4,758	4,583
R-squared	0.955	0.970	0.965	0.970
N of jurisdictions	608	606	601	605

Notes. This table shows that the results are not sensitive to whether to control for the time-varying effect of the predicted urban land price in 1996 or not. Comparing with the baseline specification (Table 6), I do not control for the time-varying effect of the predicted urban land price in 1996 here. The results are quite similar between the two tables. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 39: Association of the Land Conversion Barrier and Floor-to-Area Ratio (FAR), Robustness

Depvar:	(1)	(2)	(3)
ln of the FAR of	Commercial land	Industrial land	Residential land
Cu	0.005	0.030***	0.039***
	(0.010)	(0.007)	(0.007)
Observations	4,077	3,521	4,627
R-squared	0.903	0.943	0.967
N of jurisdictions	594	589	603

Notes. This table shows that there is no correlation between the land conversion barrier and the FAR of newly developed commercial land during the 2000s (Column 1). For residential land and industrial land, the FAR is positively correlated with the land conversion barrier (Column 2 and Column 3). The average FAR of new buildings in each land use category is calculated as a land area-weighted average FAR across all land transactions in that land use category. I use the average of the maximum and the minimum of the legal FAR to approximate the FAR chosen by individual urban land developers. In all the regressions, the control variables include: (a) region time-varying effects; (b) the time-varying effects of population, employment rate, illiteracy rate, percent of college graduates, percent of employment from the agricultural sector, percent of employment from construction sector, and percent of immigrants from other city jurisdictions in 1990; (c) the time-varying effects of GDP, urban land supply, and population in 1996. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

B.1.6 Additional Tables for Section 6

Table 40: Estimation of the Price Elasticity of Farmland Supply

Depvar: lnc_{Rnt}	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$ln\left(1-\frac{H_{nt}+R_{nt}}{(1-\phi_n)\bar{R}_n}\right)$	-1.779***	-1.941***	-1.824***	-1.886***	-1.779***	-1.778***	-1.755***
	(0.323)	(0.344)	(0.338)	(0.335)	(0.323)	(0.323)	(0.322)
Wald-F	71.45	71.45	69.68	71.45	71.46	71.43	71.45
Observations	4,728	4,728	4,737	4,728	4,728	4,728	4,728

Notes. This table shows that the estimation of the price elasticity of farmland supply is robust to alternative ways to impute the farmland development cost, farmland supply during 2001 and 2004, and urban land supply for smaller city jurisdictions during 1999 and 2001. Column 1 provides the baseline results from Table 10. In Column 2, I assume that the compensation for farmland development, c_{Fn}, is positive in only one city jurisdictions in a prefecture. Specifically, it is only positive in the city jurisdiction that has the highest growth rate of urban land from 1999 to 2010. Correspondingly, ω_n is specified as $\frac{\Delta H_n}{\sum_{n'\in p}\Delta H_n'}$. The cost dispersion between locations from the same prefecture is higher when using this alternative imputation method. In Column 3, I assume that for locations that are not binding on the constraint, the farmland development cost in year t equals the maximum of farmland rent in that location up to year t. This alleviates the concern that if unforeseen negative demand shocks to the agricultural sector makes the farmland development cost in year t larger than the farmland rent. In Column 4, I use the change in urban land and farmland during 2000 and 2005 to determine the constraint binding condition. This deals with the concern that locations binding on the constraint in 2010 might not be binding back in the early 2000s. With this alternative binding condition information, I infer the marginal farmland development cost. In Column 5, I assume a constant increase in farmland amount between 2000 and 2005 instead of a constant growth rate of farmland. In Column 6, I assume a constant increase in urban land amount between 1997 and 2001 for these smaller cities instead of a constant growth rate of urban land. In Column 7, I assume that the interest rate is 6.5% when attenuating the farmland development cost to each year instead of 5.8% as in the baseline. 6.5% is the median interest rate faced by rural households calculated from the China Household Finance Survey (2011). Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 41: Structural Simulation of the Reduced Form Results, $\nu = 6$

	(1)	(2)	(3)	(4)
Depvar: Δln of	Urban land	GDP	$\text{GDP}_{nonagri}$	population
$\widetilde{c_{un}}$	-0.064***	-0.025**	-0.034***	-0.015**
	(0.021)	(0.010)	(0.012)	(0.008)
Observations	631	631	631	631
R-squared	0.089	0.089	0.094	0.077

Notes. This table shows the simulated impacts of the Farmland Red Line Policy on urban land supply, GDP and population using a quantitative model that has a large labor supply elasticity. In this quantitative model, the labor supply elasticity is specified to be 6, which is towards the high end of the estimates of the parameter in the literature. The simulated impacts on GDP and population are very close to the estimates using the realized data. The aggregate gain of real GDP and welfare if removing the policy is larger than the baseline case. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 42: Correlation of the Calibrated Urban Sector Productivity with FDI and Human Capital

Depvar: $\ln\bar{A}_M$	(1)	(2)	(3)	(4)	(5)
ln(FDI per worker)	0.0781***				
	(0.00874)				
ln(contracted FDI per worker)		0.0421***			
		(0.00810)			
% graduate degree			49.53***		
			(6.301)		
% college degree and above				2.924***	
				(0.350)	
Average education years					0.209***
					(0.0240)
Observations	631	631	631	631	631
R-squared	0.113	0.041	0.089	0.100	0.108

Notes. Standard errors in parentheses:*** p<0.01, ** p<0.05, * p<0.1

Table 43: Correlation of the Calibrated Urban Sector Amenity Level with Theaters and Library Book Collections

Depvar: $\ln\bar{B}_M$	(1)	(2)
ln(N of library books)	0.103***	
	(0.0161)	
ln(N of theaters)		0.164***
		(0.0215)
Observations	620	586
R-squared	0.063	0.090

Notes. Standard errors in parentheses:*** p<0.01, ** p<0.05, * p<0.1

Table 44: Correlation of Urban Land Price based on the Model and the Land Sales Data

Depvar: Calibrated urban land price	(1)	(2)	(3)	(4)
average price of all sales	0.186***			
	(0.0284)			
average price of land sold via auction		0.173***		
		(0.0301)		
average price of new urban land			0.135***	
			(0.0241)	
average price of existing urban land				0.163***
				(0.0235)
Observations	570	562	559	555
R-squared	0.070	0.056	0.053	0.080

Notes. This table shows that the model-calibrated urban land price is highly correlated with the price constructed using the land transaction data. Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 45: Correlation of Farmland Price based on the Model and the Rural Household Survey Data

Depvar: Market price of farmland	(1)	(2)	(3)
Calibrated farmland price	0.348***	0.699***	0.976***
	(0.100)	(0.161)	(0.230)
Constant	4.091***	1.898	-0.044
	(0.806)	(1.219)	(1.753)
Observations	5,902	3,078	354
R-squared	0.023	0.046	0.095

Notes. The error term is clustered at location level (city jurisdiction or rural region). Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

B.1.7 Additional Tables for Section 7

Table 46: Sensitivity of Counterfactual Outcomes to Alternative Parameter Values

Specification	(1) Baseline	(2) $\alpha = 0.02$	(3) $\alpha = 0.08$	(4) $\theta_M = 0.59$	(5) $\theta_M = 0.76$	(6) $\nu = 2$	(7) $\nu = 4$
Workers' welfare	5.78	5.66	5.90	7.09	3.85	5.27	6.13
Manufacturing output	4.95	4.80	5.11	5.80	3.66	4.31	5.35
Agricultural output	-2.79	-2.79	-2.80	-3.22	-2.17	-2.51	-3.01
Urban population	5.20	5.18	5.22	6.19	3.75	4.40	5.78
Urban land	39.58	39.14	40.08	40.69	37.90	37.35	41.41
Farmland	-6.67	-6.68	-6.65	-6.57	-6.82	-6.60	-6.74

Specification	(8) $\zeta = 1.25$	(9) $\zeta = 2.52$	(10) $\sigma_M = 4$	(11) $\sigma_M = 9$	(12) $\sigma_F = 4$	(13) $\sigma_F = 9$	(14) $\sigma_F = 7$
Workers' welfare	5.07	6.33	5.65	5.82	5.59	5.74	5.79
Manufacturing output	4.69	5.16	4.85	4.99	4.72	4.91	4.97
Agricultural output	-2.64	-2.89	-2.54	-2.84	-2.80	-2.80	-2.79
Urban population	4.95	5.40	5.07	5.25	4.97	5.15	5.22
Urban land	35.31	43.07	37.84	40.24	37.87	39.26	39.72
Farmland	-8.02	-5.50	-6.69	-6.65	-6.44	-6.62	-6.69

Notes. The baseline results are reported in Column 1 for comparison. The counterfactual results are robust to the value of α in the range of $[0.02, 0.08]$ (Column 2 and 3), θ_M in the range of $[0.59, 0.76]$ (Column 4 and 5), ν in the range of $[2, 4]$ (Column 6 and 7), ζ in the range of $\zeta \in [1.25, 2.52]$ (Column 8 and 9), σ_M in the range of 4 to 9 (Column 10 and 11), σ_F in the range of 4 to 9 (Column 12 and 13) as well as when it is specified to be the same as σ_M (Column 14).

Table 47: Sensitivity of the Counterfactual Outcomes to the Imputation of τ

Specification	Baseline	Specification 2	Specification 3	Specification 4	Specification 5
Workers' welfare	5.78	5.82	7.46	5.93	6.60
Manufacturing output	4.95	4.99	6.51	5.08	5.66
Agricultural output	-2.79	-2.82	-3.67	-1.98	-3.21
Urban population	5.20	5.26	6.50	5.21	6.00
Urban land	39.58	40.45	53.78	55.06	47.34
Farmland	-6.67	-6.71	-7.20	-4.01	-7.66
Mean of τ	1.43	1.45	1.56	1.73	1.54
SD of τ	0.62	0.80	0.82	2.55	0.96

Notes. This table shows that the estimates of aggregate effects of the policy are robust to alternative ways of imputing the cost-to-rent ratio of farmland. Across all the methods of imputing the cost-to-rent ratio, baseline specification returns the lowest mean cost-to-rent ratio. Therefore, this method generates the most conservative estimates about the aggregate cost of the policy in terms of workers' welfare. In the **baseline specification**, I assume that the value is the same as in 2004, and the cost-to-rent ratio of farmland in 2004 is calculated based on the assumption that all the locations within the same prefecture that are binding on the quantity constraint have the same gap between the farmland development cost and farmland rent. In 7 locations, the cost-to-rent ratio is above five, and this might be due to measurement errors in the data. To avoid too large an increase in urban land in individual jurisdictions, potentially due to measurement errors, I adjust the cost-to-rent ratio of the top 1% locations to the value at 99 percentile. **Specification 2** uses the same data as in Specification 1, except that the values from the top 1% locations are not replaced with the 99 percentile value. The aggregate results are very close to the baseline results. In **Specification 3**, I infer the farmland development cost in 2010 using data about the change in the percentage of land development from 2005 to 2010, the change in national average farmland development cost from 2004 to 2011 and the farmland development cost in 2004. I then calculate the cost-to-rent ratio using the imputed farmland development cost and farmland rent in 2010. The counterfactual results based on this alternative imputation method generate slightly larger welfare gain. This is because, overall, the cost-to-farmland ratio increases over time. In **Specification 4**, I assume that the cost-to-rent ratio is the same as in 2004 again. But the cost-to-rent ratio in 2004 is constructed based on the assumption that within a prefecture, the cost-to-rent ratio is only positive in the city jurisdiction that has the highest percentage of land development around the 2000s. Under this specification, the constraint is only binding in less than 300 locations, but the distortion is relatively high in these locations. Hence we observe that urban land increases by 55.06%. This does not generate additional welfare gain, because the increase in urban land concentrates in a few places and hence the marginal impact from these additional increase in urban land is minimal. In **Specification 5**, I assume that the interest rate used to amortize the farmland development cost is 7.0% instead of 5.8% as in the baseline. A higher interest rate increases the amortized farmland development cost and therefore increases the cost-to-rent ratio.

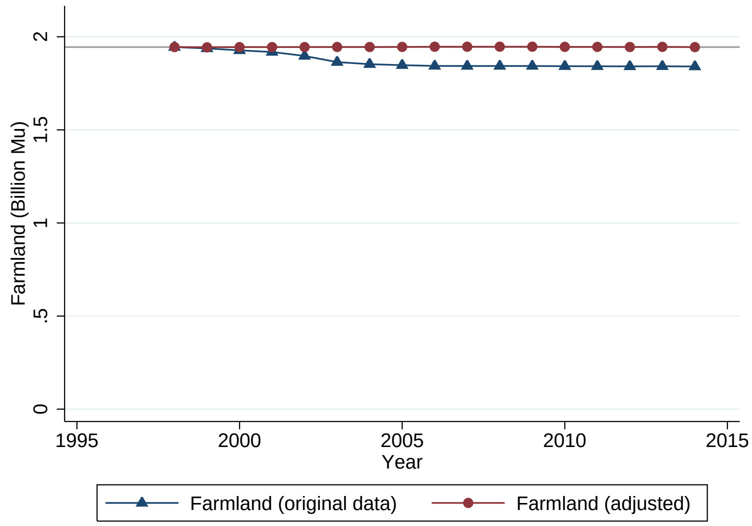
Table 48: Sensitivity of Counterfactual Outcomes to Model Extensions

Specification	(1) Baseline	(2) Drop rural regions	(3) $\lambda_U = 0.1$	(4) $\lambda_U = 0.5$	(5) $\lambda_U = 0.9$	(6) Re-distribute land profit
Workers' welfare	5.78	7.78	6.02	5.95	5.43	6.54
Manufacturing output	4.95	2.04	2.79	2.77	2.57	5.13
Agricultural output	-2.79	-6.83	2.80	2.78	2.59	-3.21
Urban population	5.20	2.33	5.30	5.23	4.83	5.86
Urban land	39.58	35.31	41.75	40.91	35.99	40.68
Farmland	-6.67	-11.22	-6.42	-6.42	-6.43	-6.56

Notes. This table shows that the estimates of the aggregate cost of the policy are qualitatively similar when using alternative model specifications. **The first extension** is to drop the rural regions that only have an agricultural sector in it. By doing so, I prevent the trade and migration flow between the city jurisdictions and the rest of rural regions and only focus on the welfare of workers inside city jurisdictions. The first consequence is that more urban land increases the utility of workers already residing in the city jurisdictions but does not increase the total population across city jurisdictions. As a result, the welfare increase is higher, but the urban population increase is smaller than the baseline outcome. The second consequence is that city jurisdictions cannot shift agricultural production to the rest of rural regions. Therefore, city jurisdictions cannot specialize in manufacturing production as much as in the baseline. This can be seen from a smaller increase in output from the manufacturing sector comparing to the baseline. Finally, despite that the decrease of agricultural output and farmland is larger, it is smaller than the corresponding change in city jurisdictions in the baseline. **The second extension** is to make the increase in urban land in the counterfactual equilibrium increase with the farmland density in urban land development. λ_U governs the relative cost of developing unused land into urban land to developing the same unused land into farmland. The baseline estimates are close to the results from the extended version for a wide range of values of λ_U. Note that the welfare impact gets larger when λ_U is close to zero because the cost of urban land mainly depends on the cost of farmland as λ_U goes to 0. Therefore, reducing the cost of using farmland to develop urban land would reduce urban land cost even more. **The third extension** is to assume a national portfolio that collects all the land development profit and redistribute equally among the workers. The counterfactual results based on the alternative model specification are quite similar to the baseline outcome.

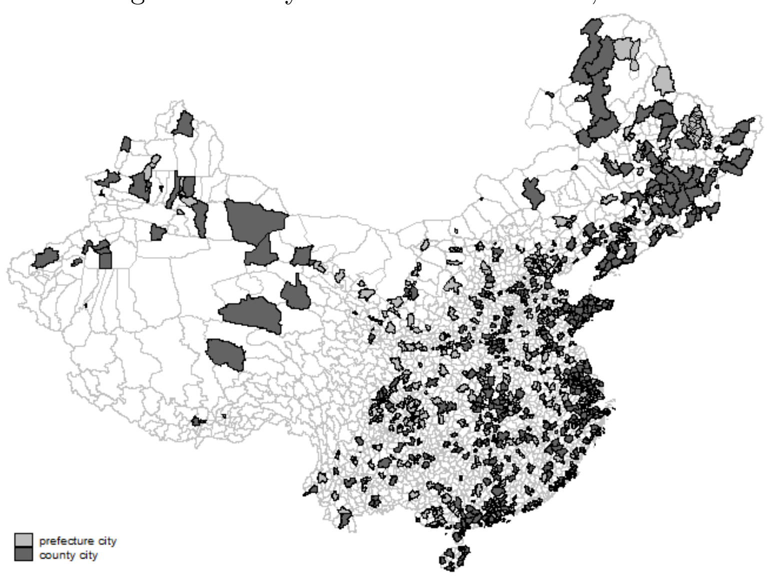
B.2 Additional Figures

Figure 9: The National Total Quantity of Farmland



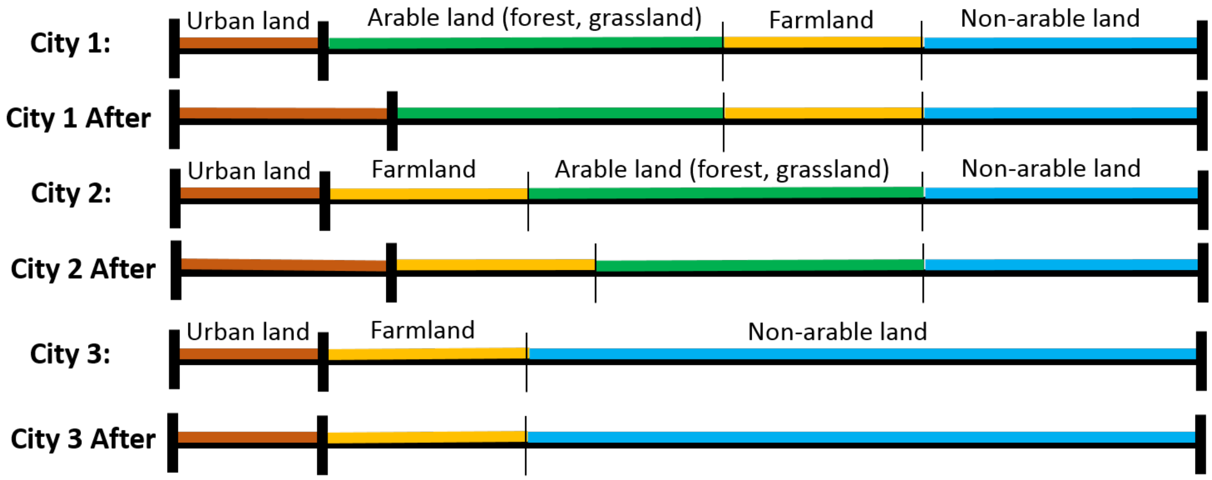
Notes. The Grain for Green National Project during the 2000s converted farmland into unused land (grassland and forest). After adding back the reduced farmland due to this project, the farmland quantity at the national level has barely changed since the Farmland Red Line Policy was adopted. The data comes from the China Land Resource Yearbooks. Grain for Green National Project is announced in 1999 and implemented during the first half of the 2000s. The project aims to restore an ecological balance to the western parts of the country by converting low-yield farmland back into forests and pasture. In the figure, the blue line represents the quantity of farmland without adding back the farmland converted into unused land during the Grain for Green National Project. The red line represents the quantity of farmland after adding back the reduced farmland. The solid grey horizontal line represents the quantity of farmland in 1998 at the national level.

Figure 10: City Jurisdictions in China, 2010



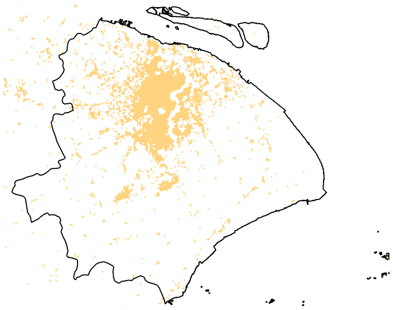
Notes. This figure shows the geographical coverage of city jurisdictions.

Figure 11: Examples of Urban Land Expansion

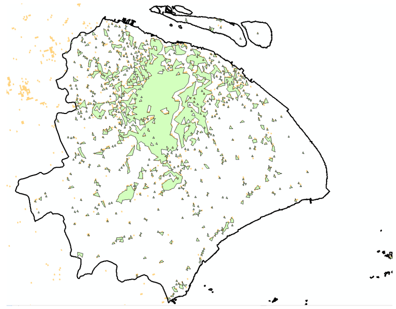


Notes. This figure illustrates how the local geographical features determine the additional cost of urban land development since the Farmland Red Line Policy was adopted. First, it is more costly to develop new urban land if the rural land surrounding the existing urban land is farmland. For example, In City jurisdiction 1, outside the urban area is unused land, while in City jurisdiction 2, outside the urban area is farmland. If the local government plans to increase a unit of urban land at the urban fringe, City 1 will convert unused land into urban land, while City 2 will convert farmland into urban land. Therefore, City 2 bears an additional cost of creating a unit of new farmland, while City 1 does not. Ceteris paribus, City jurisdiction1 increases urban land more easily than City 2 does. As in reality, urban land expansion happens in a two-dimensional world, and this variation becomes the percent of rural land around the existing urban areas being farmland. The higher the share is, the more difficult it is to create new urban land at the urban fringe. Second, it is more costly to develop new urban land if a city jurisdiction is geographically constrained to create new farmland. For example, in both City jurisdiction 2 and 3, if the local government plans to increase a unit of urban land at the urban fringe, farmland is converted into urban land. While City 2 has a lot of unused land suitable for cultivation, City 3 does not have any. Therefore, the additional cost from farmland creation is low in City 2 but is close to infinity in City 3. As a result, City 2 can expand urban land while City 3 cannot.

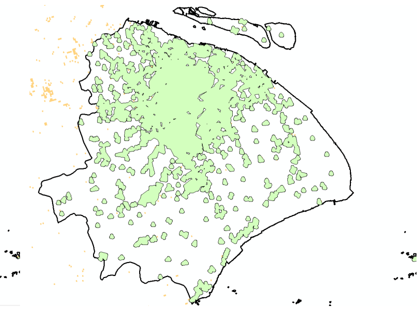
Figure 12: Construction of Urban Area Boundary



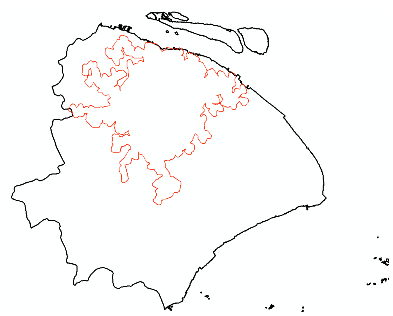
(a) Urban Land Cover Raster



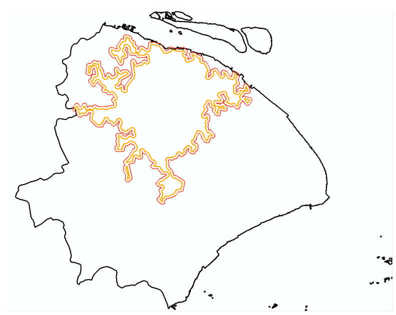
(b) Step 1



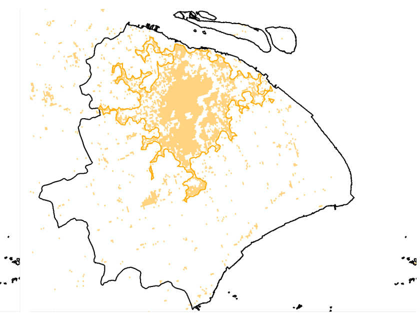
(c) Step 2



(d) Step 3



(e) Step 4

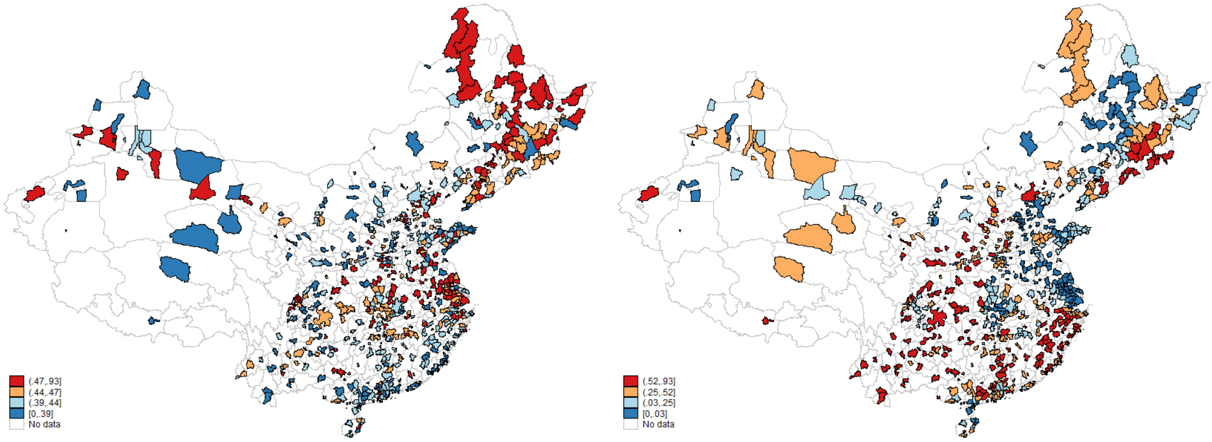


(f) Final Boundary

Notes. This figure illustrates the steps to draw the boundary of an urban area in 1995 for a city jurisdiction. I use Shanghai as an example. In Step 1, I classify geographically connected urban land points into a single polygon. To do so, I convert urban land cover raster data for 1995 into polygon. Next, I expand the boundary of polygon outward by 0.5 miles and group merge polygons that are overlapped into one polygon. In step 3, I choose the largest polygon within the jurisdiction boundary and derive its boundary. In the last step, I move all points on the boundary drawn in Step 3 inward by 0.5 miles and define the new boundary as the final urban area boundary. In Subfigure f, I overlay the original urban land cover raster data with the boundary drawn in Step 4 to show the urban area boundary properly captures the integrated urban area.

Figure 13: Land Ruggedness and Farmland Development

Figure 14: Decomposition of the Land Conversion Barrier



(a) Predicted Farmland Density

(b) Land Ruggedness in the Projected Farmland Development Region

Figure 15: Histogram of the Land Conversion Barrier

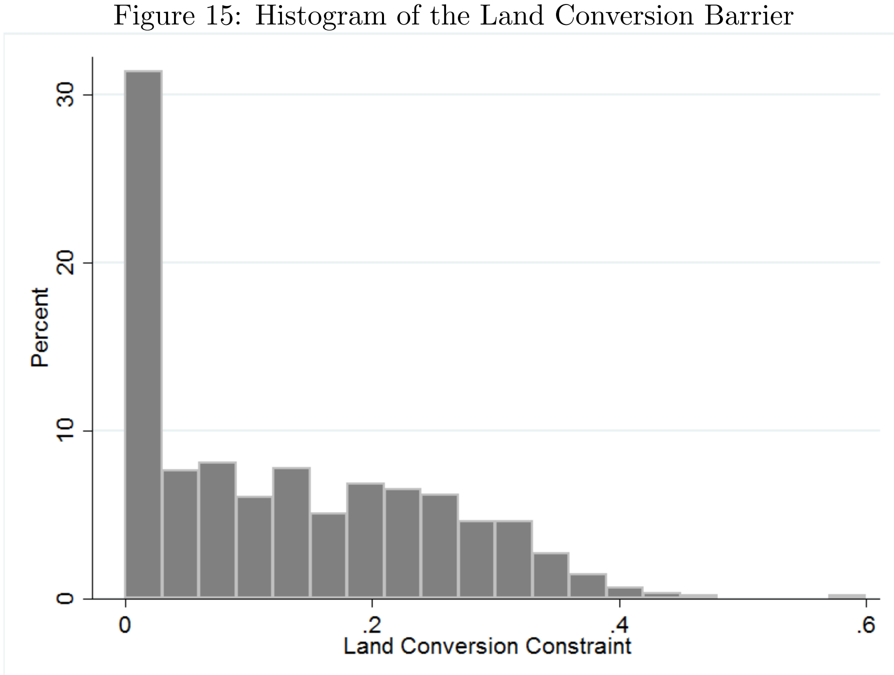
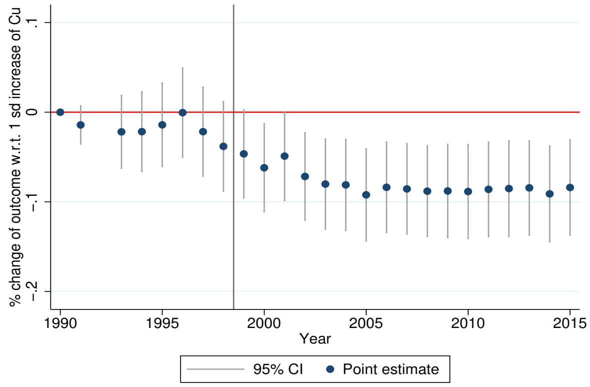
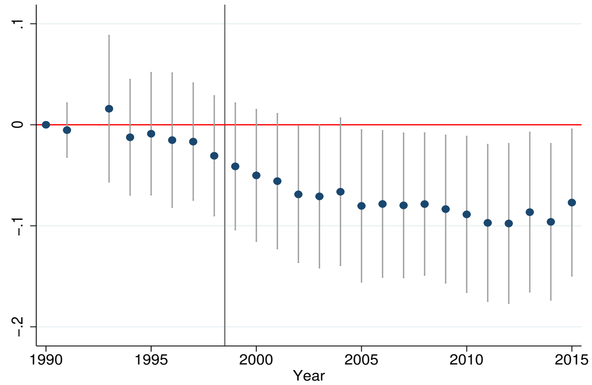


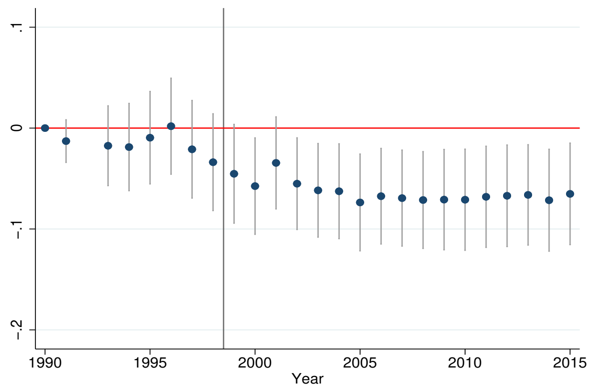
Figure 16: Impacts of the Land Conversion Barrier on Urban Land Supply, Robustness Check



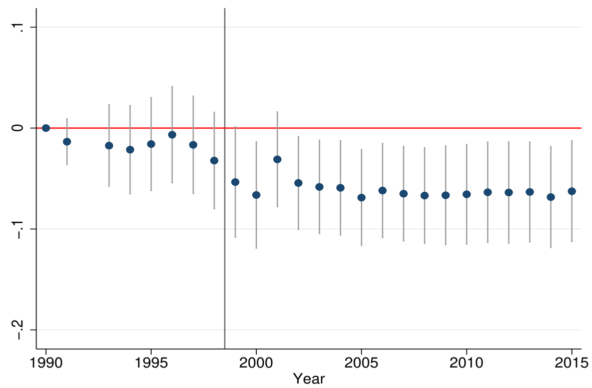
(a) Baseline



(b) 208 City Subgroup



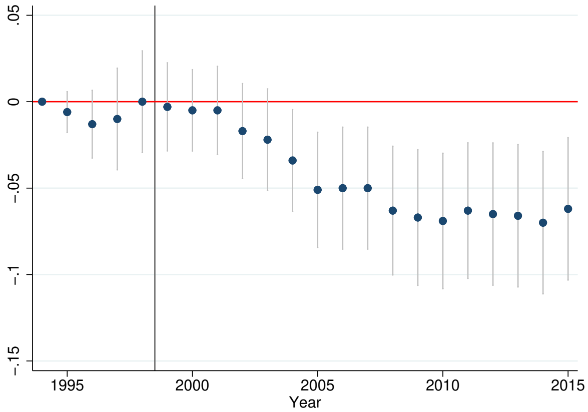
(c) Control for the time-varying effects of 1990 population and employment rate



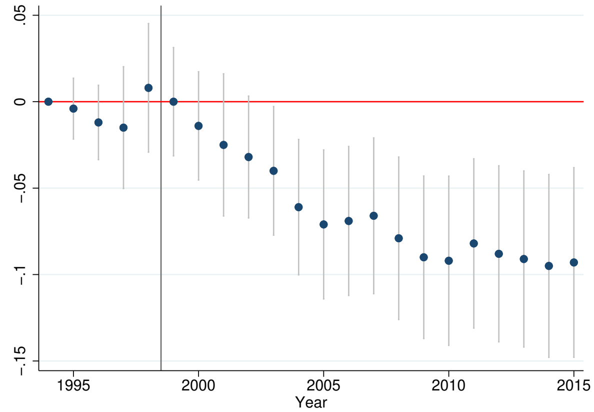
(d) Full controls

Notes. The results are robust to using a subgroup of 208 city jurisdictions that always report urban land area data since 1990, as shown in Subfigure b. Next, the results are robust to controlling for the time-varying impacts of population and employment rate in 1990 (Subfigure c) and controlling for the time-varying effects of the full set of economic characteristics of a city jurisdiction in 1990 (Subfigure d). Note that in Subfigure b, the confidence intervals of coefficients are larger, which is expected since only one-third of city jurisdictions are used in this regression. Point estimates of βs before 1998 are even closer to 0, which suggests that the missing data issues for urban land during 1997 and 2001 would worsen the parallel trend outcomes if anything.

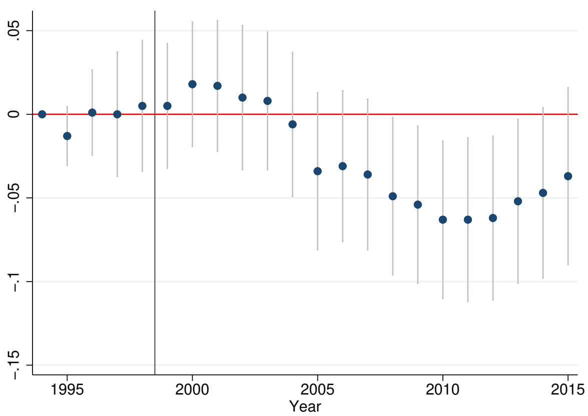
Figure 17: Impacts of the Land Conversion Barrier on GDP by Sector, Robustness Check 1



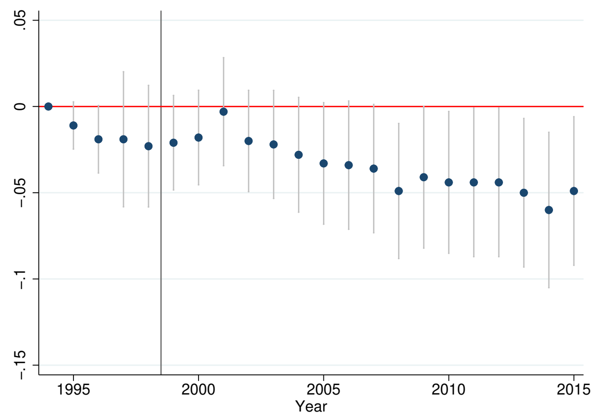
(a) Impacts on GDP



(b) Impacts on GDP from the secondary sector

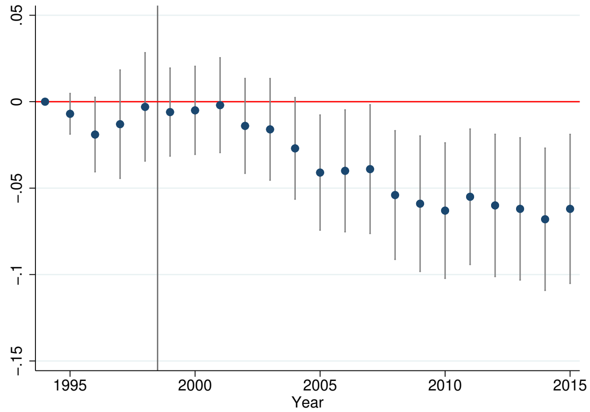


(c) Impacts on GDP from the agricultural sector

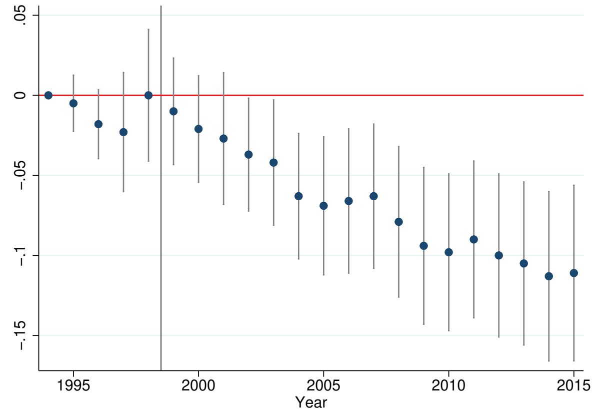


(d) Impacts on GDP from the service sector

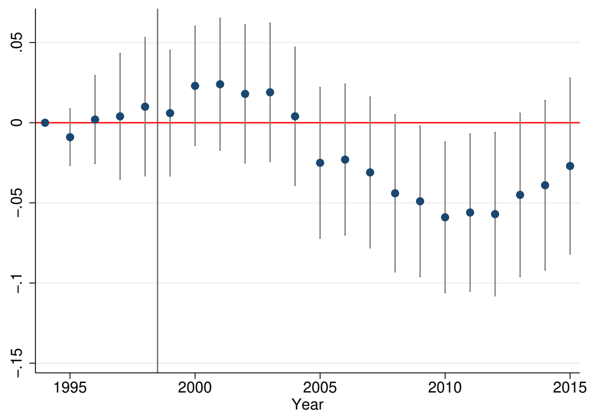
Figure 18: Impacts of the Land Conversion Barrier on GDP by Sector, Robustness Check 2



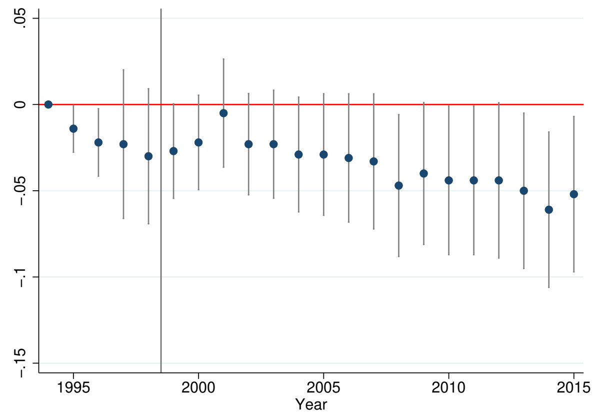
(a) Impacts on GDP



(b) Impacts on GDP from the secondary sector

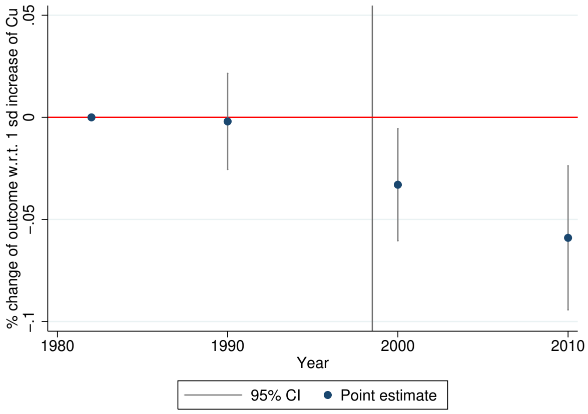


(c) Impacts on GDP from the agricultural sector

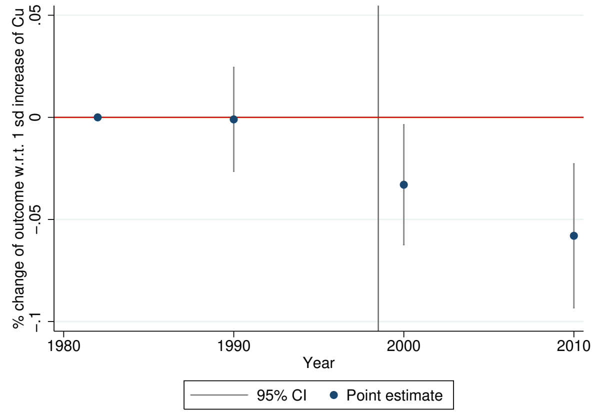


(d) Impacts on GDP from the service sector

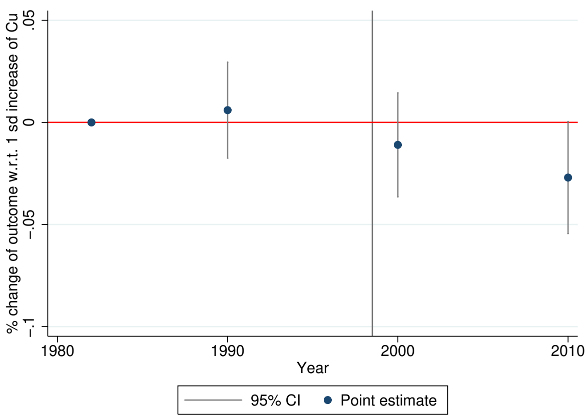
Figure 19: Impacts of the Land Conversion Barrier on Population and Employment, Robustness



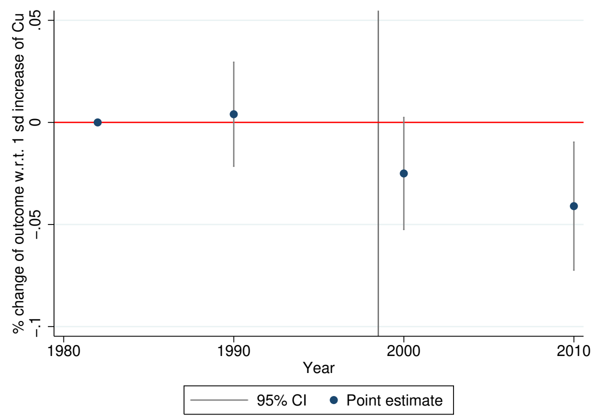
(a) Impacts on population, Control 1



(b) Impacts on employment, Control 1

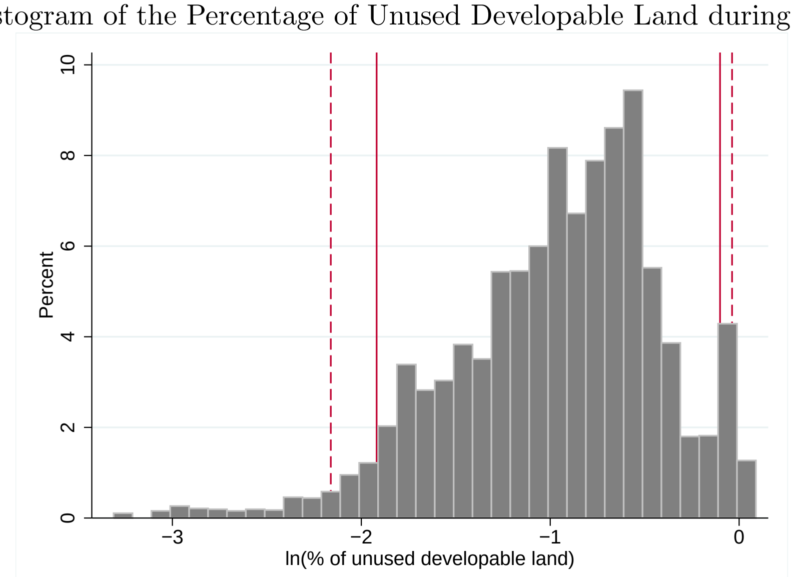


(c) Impacts on population, Control 2



(d) Impacts on employment, Control 2

Figure 20: Histogram of the Percentage of Unused Developable Land during 1999 and 2004



B.3 Description of Datasets

This subsection describes the datasets used in the empirical analysis. First, to construct a panel dataset of city jurisdictions with their economic outcomes from 1990 to 2015, I assemble data from multiple yearbooks and census. They include the China City Statistical Yearbooks (1990 - 2015), the China City Development Yearbooks (2002 - 2015), China Data Online, and the population census. Variables in the panel data at the annual frequency include GDP, GDP breakdown by sector, urban land, government expenditure per capita, and the number of hospital beds. Variables in the panel data at the decadal frequency include population, employment, and employment by sector.

The City Statistical Yearbooks provide data of urban land up to 2001 and the rest of annual frequency data up to 1996. The number of city jurisdictions included in the City Statistical Yearbooks kept increasing during the 1990s. This is because some city jurisdictions were too small in terms of their non-agricultural activities, hence not included in the City Statistical Yearbooks until the 1990s. One consequence is that the panel data is unbalanced for all the annual frequency data (urban land supply and GDP) during the 1990s. As shown in Section 4, the main regression results are not sensitive to whether to use a full sample or to use a balanced subsample of data.

Two city jurisdictions might merge into one if their urban areas become economically integrated. To make the geographical unit of analysis comparable over time, I use 2010, the latest census year, as the benchmark. If two city jurisdictions merged into one during 1990 and 2010, I treat them as one city jurisdiction throughout the period of my study and aggregate the data correspondingly. Very few city jurisdiction merges happened around 2015. For these cases, I treat two city jurisdictions as separated and remove both city jurisdictions from the dataset after they merged. Moreover, 9% of city jurisdictions changed jurisdiction boundaries by incorporating a neighbor county from 1990 to 2015. Unavailability of annual county-level data during the 1990s makes it impossible to construct a time-consistent geographical unit of analysis for these 9% city jurisdictions. I provide robustness checks by excluding these cities from regression analysis whenever applicable. I find that whether including them in the regression or not does not affect the main results. It suggests that potential measurement errors in the data brought by these city jurisdictions is not a concern.

The population census has several advantages comparing to the population data reported in the City Statistical yearbooks. First, the population census provides a more accurate accounting of the number of residents. Yearbooks only count people who register their Hukou in that city jurisdiction. Yet, working migrants may work in one city jurisdiction for a long time while keeping their Hukou registration at their birthplace. In contrast, in the population census, whoever stays in a city jurisdiction for more than six months by the time of census survey counts as a resident in that city jurisdiction. Next, the population census data allows me to construct comparable variables over time. For a few city jurisdictions that changed the administrative boundary by incorporating a nearby county during 1982 and 2010, the population of that county is always added into the city throughout the time period of my study.

Second, several additional databases are used to construct variables for supplementary analysis. The first database is the transaction data of urban land sales published on the LandChina.com. This database covers the majority of the urban land transactions in China during 2007 and 2015. I use the database to construct the average urban land price and the average Floor-to-Area ratio at the level of city jurisdiction - year - land use (commercial land, industrial land, residential land, and others). The second dataset is the urban land use by category during 2002 and 2015 from the China City Development Yearbooks. The categories include business use (industrial use and commercial use), residential use, public facility (including schools, government buildings, hospitals, libraries, stadiums, etc), and transportation and green areas (such as roads, bus and train stations, and parks).[61] Third, I use the urban land use raster data at 30-m resolution for 1995 and 2015 to draw the boundary of the urban area in each city jurisdiction. Detailed explanation about the construction of the urban boundary is in the footnote of Figure 12. The polygon of urban boundary is then used to calculate the remoteness index of the urban areas each year.

Third, to construct the land conversion barrier, I take the data of soil qualities and the terrain slope feature at 1 km resolution from the World Harmonized Soil Database (Fischer et al., 2008b). In the World Harmonized Soil Database, the soil qualities of land grids in China is based on the 1:1 million soil map of China, provided by the Chinese Academy of Sciences (Nachtergaele et al., 2010). The granular data about soil qualities and land ruggedness were collected during the national soil survey conducted during 1989 and 1993 (Fischer et al., 2008a). Second, the actual farmland density of the rural land converted into urban use in each city jurisdiction during 2002 and 2015 is from the China City Development Yearbooks. For 13.9% of city jurisdictions (88 out of 631), the amount of farmland converted into urban land is always missing from 2002 to 2015. Such city jurisdictions are dropped out of the prediction regression. Third, variables from the 1982 and 1990 population census are used to conduct the balance test for the land conversion barrier.

Fourth, to impute the marginal cost of farmland development after 1999, I calculate the government payment per unit of new farmland during 1999 and 2004 using data from the China Land Resource Yearbooks. The yearbooks report the total government payment on farmland development and the hectares of new farmland developed aggregated at the prefecture level. Therefore, the prefecture average government payment per unit of new farmland is the total government payment on farmland development divided by the hectares of new farmland developed. Second, the data of the farmland amount in each city jurisdiction in 2000, 2005, and 2010 come from GeoExplorer II. I also use the MODIS 500m land cover data from 2001 and 2011 and the GLC30m (Global Land Cover at 30-meter resolution) from 2010 for robustness checks. Starting from 1999, the Grain-for-Green program in China converts farmland in very mountainous regions back to unused land (such as grassland and forest). The program makes the change in farmland in the places involved in the program negative, even though it is not due to urbanization. The GeoExplorer reports the amount of farmland on each type of land. Therefore, I can detect the reduction of farmland on mountainous land during 2000 and 2010. I treat the reduction of the farmland on mountainous land as caused by this program and make a deduction on the amount of farmland in 2000. After the adjustment, the change in the farmland from 2000 to 2010 is solely due to the interaction of the Farmland Red Line Policy and the location governments' land development decisions.

Fifth, to impute the marginal cost of new farmland development after 1999, I calculate the gov-

[61] The original classification has more than four categories and the classification adjust in 2008. To make the classification consistent across time, I aggregate them into four categories.

ernment payment per unit of new farmland during 1999 and 2004 using data from the China Land Resource Yearbooks. The yearbooks report the total government payment on farmland development and the hectares of new farmland developed aggregated at the prefecture-level. Therefore, the prefecture average government payment per unit of new farmland is the total government payment on farmland development divided by the hectares of new farmland developed.

Sixth, data about the amount of farmland in each city jurisdiction in 2000, 2005, and 2010 come from GeoExplorer II. I also use the MODIS 500m land cover data from 2001 and 2011 and the GLC30m (Global Land Cover at 30-meter resolution) from 2010 for robustness checks. Starting from 1999, the Grain-for-Green program in China converts farmland in very mountainous regions back to unused land (such as grassland and forest). The program makes the change in farmland in the places involved in the program negative, and it is not a violation of the Farmland Red Line Policy. The GeoExplorer II reports the amount of farmland on each type of land. Therefore, I can detect the reduction of farmland in mountainous areas during 2000 and 2010. I treat the reduction of the farmland in mountainous areas as caused by this program and make a deduction on the amount of farmland in 2000. After the adjustment, the change in the farmland from 2000 to 2010 is solely due to the interaction of the Farmland Red Line Policy and the rural residents' land development decisions.

Finally, I disclose the source of the variables used to test the model fit in Section 6.5. The percentage of graduate degree holders, the percentage of the population with a college degree and above, and the average education years of population are calculated using the 2010 population census. Second, the number of theaters and the number of books collected by the public libraries in 2012 come from the China City Statistical Yearbook. Third, the FDI investment per capita in 2007 is also from the China City Statistical Yearbook. 2007 is the closest year to 2010 for which the data of FDI investment at the city jurisdiction level is provided. Next, I use the Chinese Household Finance Survey from 2012 to calculate the expected market value of the farmland in 210 locations (city jurisdictions and the rest of rural regions).

B.4 Construction of the Land Conversion Barrier

The construction of the land conversion barrier takes four steps. In Step 1, I define the projected new urban area as the rural land surrounding the existing urban area right before the policy began. In Step 2, I calculate the soil qualities in the projected new urban area and use them to predict the percentage of urban land converted from farmland since the policy was adopted. In Step 3, I define the projected farmland development region as land close to the administrative boundary. In Step 4, I calculate the ruggedness of land in the projected farmland development region and use it to approximate the unit cost of new farmland development. Figure 21 uses the Shanghai City Jurisdiction to illustrate where the projected new urban land and the projected farmland development region are in the baseline specification.

Projected new urban area. The projected future urban area in a city jurisdiction is defined as a 2-km outward buffer of the boundary of the urban area in 1995. 1995 is the most recent year before 1999 for which urban land cover raster data at a fine level is available. Urban land cover raster data is from Liu et al. (2018). The boundary of an urban area is defined as the largest continuous urban

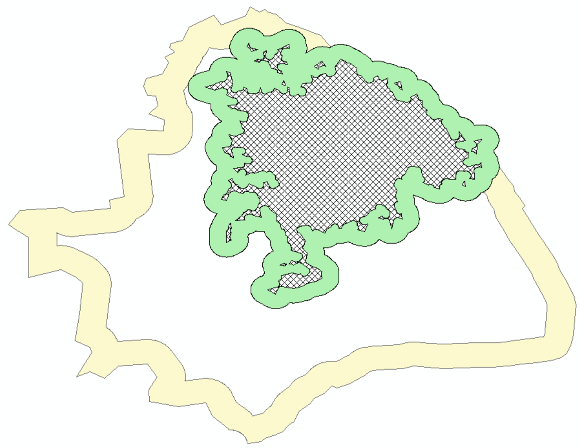


Figure 21: Projected New Urban Area and Projected Farmland Development Region

Notes. I use Shanghai City Jurisdiction as an example to illustrate the projected future urban area and the projected farmland development region. The grey area represents the urban area in Shanghai in 1995 based on the raster data of urban land for 1995. The green area represents the projected new urban area and the yellow area represents the projected farmland development region.

built-up area within the administrative boundary of the city. Two blocks of the built-up area are treated as connected if the distance is within 1 mile. Details of the boundary construction process are explained in Figure 12.

Predicted farmland density. The predicted farmland density is the predicted percentage of new urban land converted from farmland since 1999. I regress the actual percentage of new urban land converted from farmland during 2002 and 2015 against the soil qualities of the projected new urban area.[62] The predicted value of the dependent variable is used as the predicted farmland density. It is essentially a weighted average of soil characteristics of the projected future urban area. Soil features used in the baseline specification include organic carbon, pH, salinity, gypsum, gravel fraction, water storage capacity, and soil drainage.[63] In agricultural engineering literature, these features are the most critical ones that determine soil suitability for cultivation (Zhang et al., 2010). The results are robust to using all soil features provided by the World Harmonized Soil Database (Fischer et al., 2008a) to predict the farmland density.

Projected farmland development region. I define the projected farmland development region to

[62] The regression results are displayed in Table 15 Column 1. All the critical chemical and physical features significantly affect the farmland density. For robustness checks, instead of using individual soil characteristics measures, I use the nutrient availability index provided by the World Harmonized Soil Database to predict the farmland density for robustness checks. This composite index represents the degree of soil suitability for low to intermediate level input farming activity. As shown in Column 2, this alternative index is also positively associated with the farmland density of rural land actually converted into urban land during the 2000s.

[63] For each soil characteristic, I make a linear transformation such that the higher the value is, the more suitable that land grid is for cultivation. If the value of a soil character already preserves this feature, then no linear transformation is made. I then calculate the average value of each soil characteristic within the projected new urban area. For some characteristics such as organic carbon and pH, topsoil and bottom soil could have different values, and both are reported in the soil database. In this case, I use the minimum of the two as the value for that grid. Results are robust to using the average or the maximum of the two.

be a 5-km inward buffer of the administrative boundary of city jurisdiction. In Section 4, I show that the results are not sensitive to assuming all the rural land outside the existing or projected future urban area to be farmland development region.

Ruggedness of land in the projected farmland development region. I use land ruggedness in the projected farmland development region to approximate the cost of farmland development (Nunn and Puga, 2012). In agriculture engineering literature, the ruggedness of land is a crucial determinant for low cost cultivation of land. On the flat land (as in Figure 13a), low cost standard cultivation techniques are applicable immediately; in the rugged area (as in Figure 13b), flat surface along the slopes have to be created first to make the land cultivable, hence involving a lot more labor input and agricultural engineering technique. The soil quality of the land is less of a constraint here because when the raw soil does not provide enough nutrition, an easy solution and hence a common practice would be to put the top soil of the farmland that is converted into urban land onto to the newly developed farmland.[64] Finally, in the baseline, I chose the percentage of land grids with a slope of 15 degrees or more as a measure of land ruggedness.[65] The results are not sensitive to a various different ways to measure the land ruggedness and alternative buffer width.

B.5 Additional Robustness Checks of the Empirical Analysis

This subsection provides a series of robustness checks to show that the main empirical findings in Section 4 are not sensitive to regression specifications, sample selections, and alternative specifications of the land conversion barrier.

First, the causal impact of the land conversion barrier on urban land is not affected by urban land data missing issues or potential measurement errors in urban land area, as discussed in Appendix B.3. In Appendix Table 16 Column 2, I only include city jurisdictions that exist in the yearbook since 1990. In Column 3, I only include the 208 cities with urban land supply data available in the yearbooks from 1997 to 2001. In Table 16 Column 4, I exclude the city jurisdictions that have adjusted city jurisdiction boundary through incorporating a neighbor county. All the results are similar to the baseline results, suggesting that the measurement errors in the urban land supply data is not a concern.

Similarly, the causal impact of the land conversion barrier on GDP is not contaminated by the sample selection problem (Appendix Table 23 Panel A) or data measurement errors due to boundary inconsistency in a few city jurisdictions (Appendix Table 23 Panel B).

Second, I do not find empirical evidence supporting that urban land expansion in politically favored city jurisdictions are less affected by the land conversion barrier. Politically favored cities may face a lower political cost to delegate farmland creation tasks to other locations. If these city jurisdictions are less affected, by excluding them from the regression, β should be more negative than the baseline regression outcome. In Appendix Table 17, I exclude 26 provincial capital cities in Column 2, 4 provincial-level city jurisdictions in Column 3, and 14 coastal port city jurisdictions in Column 4.[66]

[64] Since there is construction in the underground of the new urban land (to build in electricity line, water system, etc.), the top soil has been removed anyway. Hence the only additional cost here is to ship the top soil to where the new farmland is. See National Land Resource Department Annual Report, 2006.

[65] This is because the central government explicitly discourages farmland from being created at surface with a slope of 15 degrees or above.

[66] Chen et al. (2017) suggests that provincial capital cities have better political connections with the central government

The coefficients from these subgroup regressions are very close to the baseline outcome reported in Column 1. Similarly, the impact on GDP and population does not change after dropping politically favored cities (Appendix Table 30 and Table 24).

Third, the results are not sensitive to the way the land conversion barrier is specified. There are three decisions to be made when defining the land conversion barrier. First is what variables are used to predict the farmland density. Second is how to measure the land ruggedness in the projected farmland development region. Third is which areas are defined as the projected future urban area and the projected farmland development region. In Table 18, Table 25 and Table 31, I use alternative sets of soil quality variables to predict the farmland density. Results are all similar to the baseline outcomes. In Table 19, Table 26 and Table 32, I use different definitions of land ruggedness to approximate the farmland development costs and again results are similar to the baseline results. In Table 20, Table 27 and Table 33, a series of ways to define the projected new urban area and the projected farmland development region are used. All results are close to one another, suggesting that the result is robust to alternative ways to define the projected new urban area and the projected farmland development region.

Finally, I apply HAC spatial clustering to adjust the error term and find it generates similar results as when clustering the error term at the city jurisdiction level. The results for urban land supply are reported in Appendix Table 22 Column 4 and 5. The results for GDP are reported in Appendix Table 29 Panel B and the results for population are reported in Appendix Table 35 Column 3 and 4.

B.6 Construct Variables for Model Quantification

This subsection introduces the procedure to construct geographical features of a location in the quantitative model. First of all, ψ_n is a vector of average soil qualities of land inside the administrative boundary of location n. The soil qualities capture the major dimensions that matter for crop growth, including pH, organic carbon, gravel percentage, water storage capacity, water drainage, and soil electrical conductivity. Next, ϕ_n is the percentage of undevelopable land in location n. It is defined as the percentage of land grids in location n with a local slope above 15 degree.[67] Finally, the total amount of land in a location, $\tilde{R}_n$, is calculated using the shapefile of administrative boundaries of counties in China in 2010.

The minimum farmland quantity constraint $\bar{R}_n$ equals the amount of farmland in location n right before the Farmland Red Line Policy was implemented. The nearest year for which the farmland quantity is available for all the locations is 2000. Therefore, I use the amount of farmland in 2000 to approximate the corresponding value in 1999. One concern is that if a positive demand shock for agricultural land happened during 1999 and 2000, the amount of farmland in 2000 would be above the actual minimum farmland quantity. I use a sub-sample of 360 county-level city jurisdictions to show

and thus have access to cheaper investment cost. Note that 26 out of 27 provincial capital cities are included in the baseline regression. This is because Lasa has serious data missing issues for years after 1996 and thus dropped out of the sample. The four provincial cities enjoy higher political hierarchy than the rest of city jurisdictions. The 14 coastal port cities have a more favorable foreign investment policy as well as urban development policy since 1984.

[67] It is technically difficult to develop land with a slope above 15 degree into farmland or urban land (Saiz, 2010; Nunn and Puga, 2012).

that farmland barely increased in most of the county cities.[68] Specifically, only 12.2% of locations have a growth rate of farmland greater than 1%, and another 17.4% of locations have a growth rate of farmland in less than 1%.[69] Given that county cities have a higher share of agricultural employment, the growth of farmland in county cities are expected to be higher than prefecture cities nearby if there is a regional positive agricultural demand shock during 1999 and 2000.

The quantity constraint on farmland is binding in a city jurisdiction if and only if the quantity of farmland in 2010 equals the minimum quantity of farmland and the amount of urban land increased during 1999 and 2010.[70] In some locations, the quantity of farmland in 2010 is less than the minimum farmland quantity. It may exist in practice and hence observed in the data for two reasons. First of all, in some locations, the farmland near urban areas has been converted into new urban land, while the new farmland development cannot be finished until another year.[71] These locations will have the quantity of farmland temporarily below the required minimum amount. Second, in some cases, a city jurisdiction might ask the neighbor rural region within the same prefecture to help create some of the new farmland, when the marginal land development cost in the city jurisdiction is relatively high. As a result, the quantity constraint gets permanently lower in the city jurisdiction while permanently higher in the neighbor rural region.[72]

For locations with a quantity of farmland below the minimum farmland quantity, I adjust the minimum quantity of farmland to the quantity of farmland in 2010. This adjustment likely reduces the estimated aggregate welfare cost of the policy, because this adjustment essentially reduces the quantity constraint in the model used to approximate the reality. Hence in the counterfactual equilibrium without the constraint, the welfare gain is smaller. Next, for rural regions that help create new farmland, the minimum quantity of farmland in the rural region is updated to the observed farmland quantity level in 2010. Furthermore, the minimum farmland quantity constraint is treated as binding.[73] The rest of rural regions are treated as not binding on the constraint, and the minimum farmland quantity constraint in such rural area is adjusted to the level of farmland quantity in 2010 if it is higher than the level of farmland in 2010.[74] The median adjustment of $\bar{R}_n$ is -0.54%.

At the end of this subsection, I describe the data used for estimating the supply elasticity of

[68]The data comes from the China City Statistical Yearbooks.

[69]20.0% of locations have a growth rate of less than -1%. This can be explained by that farmland development takes a few years in locations with difficulty to create new farmland. In such locations, the farmland growth rate could be temporarily below 0. In this case, the farmland quantity in 2010 is below the minimum amount, hence bias downward the calibrated minimum quantity constraint. This makes the simulated welfare gain in the counterfactual equilibrium a lower bound of the actual welfare gain. This is because some locations that are actually binding on the constraint in 2010 are treated as not affected by the policy based on the calibrated criteria. Therefore, in the counterfactual analysis, these locations will not create much productivity gain.

[70]If the amount of urban land did not change from 1999 to 2010, which accounts for 5.8% of city jurisdictions, it is treated as not binding on the constraint. This is to make the constraint condition assignment conservative.

[71]For example, as documented in the China Land Resource Yearbook (2008), the Shanghai government could not meet the minimum farmland quantity goal in 2007 because they use coastal land to create farmland, and it cannot be finished in a year. Such a case is not against the regulation, though it is discouraged by the central government.

[72]5.2% city jurisdictions fall into the second case. The existence of such a situation reduces the estimated impact of the farmland conversion barrier to urban land supply and other outcome variables if anything.

[73]Typically the city jurisdiction pays the rural region to help create new farmland. This makes the marginal cost of land development higher than the farmland rent, which indicates that the constraint is binding.

[74]Note that these rural regions have a minimum farmland quantity constraint as well. The constraint will bind if a negative farmland demand shock occurs. I treat them as non-binding to make the estimates towards the conservative end.

developed land. First, data about the amount of farmland and urban land at annual frequency during 1999 and 2004 are needed. One challenge is that the data of farmland, R_{nt}, is only available for 2000 and 2005. Furthermore, urban land amount is not available for county cities during 1997 and 2001. To impute the data of farmland during the data missing years, I use farmland amount in 2000 and 2005 to infer the amount of farmland for the rest of the years. In the baseline, I adopt the constant growth rate assumption and for robustness check use constant quantity change assumption. Similarly, for the missing urban land data, I assume a constant growth rate of urban land between 1997 and 2001 to impute the farmland amount in the corresponding year. The estimates are robust to assuming that the change in urban land amount is constant across years.

Another variable used for estimating the supply elasticity of unused land is the government payment per unit of new farmland, which is used to calibrate the marginal cost of farmland development during 1999 and 2004. As introduced in B.3, the data exists at the prefecture-level, which includes an average of 2 city jurisdictions and a rural region. Therefore, the observed prefecture-level data c^p_{Fg} is a weighted average of unit government spending on new farmland across all the locations within a prefecture. Specifically, for prefecture g, $c^p_{Fg} = \sum_{n \in g} \omega_n c_{Fn}$, where ω_n is the percentage of new farmland developed in location n to meet the policy requirement. In the baseline, I assume that $c_{Fn} = c^p_{Fg(n)}$ in all the locations that are binding on the farmland quantity constraint by 2010 while c_{Fn} is 0 for the rest of locations. In Appendix Table 40, I show that the results are robust to using various assumptions to impute the marginal farmland development cost across locations. The error term is clustered at the prefecture level as any measurement error in c^p_{Fg} affects all the locations within the same prefecture. Given that the farmland rent tends to be similar, this imputation assumption indicates that the marginal farmland development costs are similar across locations inside the same prefecture.

B.7 Model Fitness: Out-of-Sample Test

This subsection provides a detailed discussion on the fit of the model by testing the correlation between recovered unobservables with proxy variables not used in the calibration. The corresponding tables are provided in Appendix B.1.6.

First, I show that FDI per worker around 2010 is strongly positively correlated with the calibrated urban sector productivity (Table 42 Column 1). FDI is commonly used to explain the productivity, especially in developing countries, because it represents the local access to the frontier technology in production and management (Haskel et al., 2007). Next, labor skills are associated with the productivity of an urban area (Simon and Nardinelli, 2002). In Table 42 Column 3 to 5, I show that the calibrated urban sector productivity is positively associated with the percentage of graduate degree holders, the percentage of population with a college degree and above, and the average education years of population.

Second, I show that the model calibrated amenities of the urban sector is positively associated with the characteristics that make the location more desirable to live, including the presence of theaters and the scale of public library collections. Table 43 Column 1 suggests that the number of books collected by the public libraries positively correlates with the calibrated local amenity level. Next, Table 43

Column 2 shows that the number of theaters is positively associated with the location amenity level.

Third, the calibrated farmland price is close to its counterpart, calculated using the Chinese Household Finance Survey. The survey reports both the amount of farmland owned by a household and the expected market value of the farmland, and the farmland price can be computed.[75] I regress the expected farmland price at the household level against the calibrated farmland price of that location and cluster the error term at the level of city jurisdictions/rural regions. A coefficient close to 1 and a constant close to 0 would indicate that the calibrated farmland price is a good approximation of the average farmland price in a location. The regression associated with Column 1 includes all the rural households who own farmland. The coefficient in front of the calibrated farmland price is significantly positive, though smaller than 1. This can be explained by that rural households in more remote rural areas, and those who do not rent out farmland are not as informed about the actual market value of their farmland. As shown in Column 2, if only including the rural households inside the city jurisdictions, the coefficient gets much closer to 1, and the constant is not significantly different from zero. If we further restrict the sample to those who rent out their farmland in the survey year, the coefficient is even closer to 1, and the constant is very close to zero, as displayed in Column 3.

Finally, I test the correlation between model calibrated urban land price and the average urban land price based on urban land transaction data in 2010 from LandChina.com. As shown in Table 44 Column 1, there is a strong positive correlation between the model calibrated data and the actual data. The results are robust to the urban land price based on urban land plots sold through auction only (Column 2), newly developed urban land (Column 3), and existing urban land (Column 4). Note that the calibrated price and the one calculated using the transaction data are not expected to be along a 45-degree line, and hence the coefficient is not necessarily 1. This is because the calibrated urban land price refers to the value of land after building structures are put on it while the urban land price reflects the value of urban land with basic infrastructure access (such as electricity and water system) but no real estate construction yet.

[75]The survey conducted in 2012 covers 210 out of 889 rural regions included in my analysis. Despite that the survey started in 2010 and was conducted every two years, I do not use the first round of the survey because the unit of expected market value of the farmland is not specified in the survey, hence causing large measurement errors.