Title
Geographic determinants of China's urbanization

Authors
Christensen, Peter
McCord, Gordon C

Publication Date
2016-07-01

DOI
10.1016/j.regsciurbeco.2016.05.001

Peer reviewed

Geographic Determinants of China's Urbanization*

Peter Christensen[†]

Gordon C. McCord[‡]

May 19, 2016

Abstract

This study exploits a unique set of satellite and other spatially explicit data to examine the role of three exogenous geographic factors in shaping and constraining urbanization: biophysical land suitability for agriculture, distance to major ports and terrain slope. The setting is China in the 1990's, the most expansive process of urban growth in history. Our empirical results suggest that these geographical factors explain nearly half of the variation in urbanization levels. However, controlling for long-run levels, we find a weakly negative relationship between agricultural land suitability and urban expansion from 1990-2000, which is consistent with the theoretical expectations that rising opportunity costs affect the development of fertile lands. We examine heterogeneity in the effects of geography using interactions with province fixed effects and, even more flexibly, with a localized regression technique (geographically weighted regression). Our results indicate that agricultural land suitability has opposing effects in different regions, for example leading to increased urban expansion in the Pearl River Delta and restricting urban expansion along the northern coast. These results should caution scholars against assuming homogeneous effects of physical geography across regions when doing empirical analysis of urban dynamics.

Keywords: urbanization, China, satellite, spatial econometrics, geography

---

*The authors offer a special thank you to Dr. Jiyuan Liu and colleagues at the Chinese Academy of Sciences for contributing critical expertise and data on land use/land cover change in China. Helpful feedback from Ken Gillingham, Joshua Graff Zivin, Matthew Grant, Gordon Hanson, Solomon Hsiang, Chandra Kiran Bangalore Krishnamurty, and William Masters was much appreciated, as well as from participants at seminars in UC San Diego and Cornell's Dyson School of Applied Economics.

[†] pchrist@illinois.edu, Department of Agricultural and Consumer Economics, University of Illinois, Urbana-Champaign

[‡] gmccord@ucsd.edu, School of Global Policy and Strategy, University of California, San Diego

1 Introduction

Many now refer to the 21st century as the urban century. The global urban population is expected to grow from 2.8 billion in 2000 to 6.25 billion by 2050 before stabilizing at the end of the century (United Nations Population Division, 2012). It is clear that much of this growth will occur in emerging Asia and Africa, however far less is known about the distribution and types of settlements that are developing in those regions. Rapid urbanization around the developing world and the trade-offs over land use have generated interest in the study of how geographic features influence the spatial distribution of urban growth. There is evidence that locational fundamentals are first-order determinants of city size rank (Davis and Weinstein, 2002). Recent research has examined the effect of geography on the timing of urban development (Motamed et al., 2009), urban activity as a function of access to primate cities (Storeygard, 2014), housing market responses to land constraints (Saiz, 2010), and how physical geography conditions urban sprawl in the U.S. (Burchfield et al., 2006). This paper employs a unique set of satellite and spatially explicit data on land quality, ports, rivers, groundwater and topography to study the impact of market access and land suitability for agriculture in determining urban location and expansion during the most rapid build-out in history: China in the 1990's. We employ a variety of econometric specifications to identify the total effect of these exogenous geographic characteristics, providing new estimates on how the biophysical landscape conditions human society.

Our work builds upon models that relate agglomeration and urban dynamics to underlying geography. Given that modern societies locate most of the labor, production, and high-productivity activity in cities, economics has long been concerned with the geographic constraints that determine the location and growth of urban areas. One feature that is extensively discussed in the literature is access to markets. The "new economic geography" tradition implicitly linked geography to the virtuous cycles of agglomeration, postulating that declining transport costs can stimulate industrial production by increasing the purchasing power of rural consumers and the market for manufactured goods (Fujita and Thisse, 1996; Fujita et al., 1999; Krugman, 1991). Importantly, the growth of the industrial sector and further reduction in transport costs ultimately fuels a spatial re-organization as industrial firms respond to returns associated with agglomeration. A growing urban center is the product of this emerging production hub and the simultaneous migration of workers (Murata, 2008; Ottaviano et al., 2002). A somewhat distinct and long-standing body of theory emerged to explain the spatial patterns of growth within and across urban regions (Alonso, 1964; Mills, 1967; Muth, 1969), as well as the timing of agricultural land conversion (Capozza and Helsley, 1989). As with the new economic geography models, the growth of urban areas will depend on access to global and regional markets – trade will increase both the incomes of urban producers as well as potential in-migrants.

A second geographical feature relevant for urban dynamics is agricultural production potential. One strand of the literature that addresses agriculture and urban growth focuses on the transition from agricultural to industrial production, explaining the agglomeration of firms and population as a part of a multifaceted structural change described by Kuznets (1973) and Lucas (2000). When agricultural productivity rises beyond subsistence, labor is freed up for the manufacturing and service sectors, both of which achieve higher productivity in urban areas, particularly in those with access to local and global markets to competitively buy inputs and produce exports. The effect of geography in affecting the timing of structural change and urbanization was proposed by Gollin et al. (2002) and empirically tested by Motamed et al. (2009), who find that areas with higher agricultural productivity and navigable river access were more likely to urbanize sooner. On the other hand, the Alonso-Mills-Muth models and Capozza and Helsley (1989) predict that highly productive agricultural areas present a high opportunity cost to urbanization. Cities grow away from a Central Business District (CBD) and trade off benefits of urban activity (urban wages net of rent and commuting costs to the CBD) with the agricultural land rents forgone by expanding the urban area. Empirical work has measured a negative relationship agricultural land rents and urbanization, including in China (Deng et al., 2008).

The question of how geographical features constrain urbanization is particularly relevant for China, where contemporary economic reforms have catalyzed urbanization at an unparalleled rate. China added more that 150 million inhabitants to its cities between 1990-2000 and another 452 million are expected between 2000 and 2030 (United Nations Population Division, 2012). As Table 1 shows, by 2000 six provinces had urbanized over 10% of their land area (Shanghai, Tianjin, Beijing, Jiangsu, Shandong, Henan). The first three have smaller land areas but competing interests for land use has already led to social tension. The latter three are much larger provinces and their high percentage of urban land cover implies larger impacts of land use change on agriculture and ecosystems. At the national level, China's urban areas covered 1.74% of land by 2000, as compared to 1.92% in 1992 in the United States (Burchfield et al., 2006). Given China's ongoing structural transformation and the fact that its population is four times that of the U.S., urban land cover is likely to expand significantly in the coming decades.

Among policymakers, attention has been paid to determining the optimal location for urban growth within the country as well as preserving farmland around rapidly growing urban centers. A growing literature has examined China's urbanization processes and the policies that regulate them (Lichtenberg and Ding, 2008; Ding and Lichtenberg, 2011). Deng et al. (2008) test and confirm classic predictions from urban economic theory about what drives the spatial structure of urban growth, including the role of agricultural land value. Our paper extends this line of inquiry by exploring the role of exogenous physical characteristics affecting agricultural potential, providing an alternative to some of the measures used in the literature that are likely endogenous to urban processes. Au and Henderson (2006a) examine the relationship between worker output and the spatial size of cities, finding that China's cities appear to be undersized relative to the optimum. Other studies have looked at the impacts of political incentives (Lichtenberg and Ding, 2009) and institutional constraints such as the Hukou system (migration restrictions) on China's process of urbanization (Au and Henderson, 2006b; Bosker et al., 2012). Ongoing work focuses on the effects of transportation infrastructure on city population decentralization (Baum-Snow et al., 2015) and on incomes (Banerjee et al., 2012; Faber, 2014).

This study contributes to the literature by exploiting a unique set of satellite and other spatially explicit data to examine the role of three exogenous geographic factors in shaping and constraining the location of China's new urban lands: biophysical land suitability for agriculture, distance to major ports and terrain slope. Our empirical results suggest that these three factors explain around half of the variation in urbanization levels in China. However, controlling for long-run levels, we find a weakly negative relationship between agricultural land suitability and urban expansion from 1990-2000, which suggests that the opportunity costs of developing fertile lands are not negligible. Our parsimonious model of geographic constraints does not account for many possible omitted variables. We address possible impact of confounding variables through a variety of additional specifications: controlling for other geographical variables emphasized in related empirical work, controlling for long-run levels in a growth specification, controlling for major policy changes, adding province fixed effects, and examining heterogeneity at the province level and with a localized regression technique (geographically weighted regression, discussed in the online appendix). Our results suggest that agricultural land suitability has opposing effects in different regions, leading to increased urban expansion in the Pearl River Delta and restricting urban expansion along the northern coast. These results should caution scholars against assuming homogenous effects of physical geography across regions when doing empirical analysis of urban dynamics.

2 Specifications

Two-sector models[1] predict that city genesis will occur in or very near agriculturally favorable lands, particularly in historic times when transport costs were high. If urban growth is a process that exhibits some increasing returns to agglomeration, then we would expect to see a strong association between urban area location and agricultural land quality. Moreover, to the extent that land quality is a spatially autocorrelated physical attribute, then the fact that city genesis differentially occurs near good land means that we expect

---

[1] The long theoretical tradition on agriculture-driven structural change dates back to Rostow (1960) and Johnston and Mellor (1961). Mathematical formulations are presented by Matsuyama (1992); Laitner (2000); Hansen and Prescott (2002); Gollin et al. (2002, 2007) and others.

to continue to see urban growth occurring in areas better suited for agriculture. On the other hand, to the extent that land use change on the margins of cities takes into account the opportunity cost of urbanizing agriculturally productive land, we might expect that better lands would be differentially less urbanized as city growth proceeds.

In addition to measuring the association of urbanization to agricultural suitability, we also include cost-distance to ports and terrain slope as variables of interest. Urban areas with less costly access to national and international markets are more likely to thrive and grow. Meanwhile, extremely rugged terrain is likely to slow urban expansion and encourage compact cities, although mild hills may encourage sprawl in some contexts such as the U.S. (Burchfield et al., 2006).

We employ the following reduced-form specification to measure the association between urban land area and our geographical characteristics of interest:

$$urbanarea_i = \beta_0 + \beta_1 \cdot landsuit_i + \beta_2 \cdot ln(costdistance_i) + \beta_3 \cdot slope_i + \varepsilon_i \tag{1}$$

where the percentage of urbanized land in area i is modeled as a function of the land suitability for agriculture, the cost-weighted distance to the nearest sea ports (as explained below, cost-weighting takes into account topography and ocean-navigable rivers), and terrain slope. Our prior is that β_2 will be negative, as we expect the profitability of urban development to decline at a faster rate with distance from the coast than the profitability of agriculture. We have an ambiguous prior on β_1, however, since better land suitability for agriculture means that the area likely urbanized earlier and could achieve higher density while contemporary city growth might avoid the opportunity cost of urbanizing the best agricultural land. The specification is in reduced form, of course, since many other variables will influence the location of urban land (infrastructure decisions, zoning, broader policy environments such as hukou[2]) and these endogenous variables are partly determined by the geographical characteristics of interest. For our purposes of measuring associations and eventually doing prediction, however, we are interested in the total effect of the geographical variables on the urban land distribution as opposed to the effects conditional on partly endogenous variables such as infrastructure.

Similarly, we estimate these relationships in the case of urban expansion during 1990-2000, conditional on the initial distribution of urban land in 1990:

---

[2] Land use right reform in China has proceeded at varying speeds across provinces Deng et al. (2008), while Chinese urbanization and broader demographic dynamics are constrained by the hukou system. Since 1958, this household registration system has tightly controlled how many rural workers move into urban areas, and workers outside of their authorized area could not access government- or employer-issued rations, housing, or health care. As part of its economic liberalization, China has relaxed the hukou restrictions over the last two decades (rural residents can now buy urban residency permits). Nevertheless, the hukou system remains an obstacle to the free movement of people within the country, with measurable consequences for agglomeration (Au and Henderson, 2006b).

$$urbangrowth_i = \beta_0 + \beta_1 \cdot urbanarea_i^{1990} + \beta_2 \cdot landsuit_i + \beta_3 \cdot ln(costdistance_i) + \beta_4 \cdot slope_i + \varepsilon_i \qquad (2)$$

3 Data

We measure urban expansion using satellite observations of urban land conversion in 1X1km grid cells across China. Following Capozza and Helsley (1989), we assume that urban land conversion at any time occurs in locations where the rental value of land in urban use exceeds the rental value in agricultural use. The quantity of urban land cover observed in a 1X1km grid cell at time (t) provides a continuous measure of the quantity of land within that grid cell for which productivity in urban use is greater than productivity in agricultural use at time (t) or any time previous (given the longevity of urban infrastructure). While other work on city growth has generally focused on urban population growth, our approach has the distinct advantage of allowing for direct tests of changes in urban productivity relative to agricultural productivity. Satellite data provide consistent measurements across large regions and jurisdictional boundaries[3], and recent work has measured urban expansion using measurements from multiple sensors (Zhang and Seto, 2011; Deng et al., 2008; Burchfield et al., 2006). At decadal time scales, measurements of urban land expansion from the Landsat and MODIS sensors have generally produced more consistent measurements than DMSP/OLS ("nightlights") due to well-known overglow and saturation effects (for many dense urban areas the light sensor saturates at a maximum reading above which variations in light emissions are not detectable), inconsistent time-series resulting from sensor degradation, and inconsistencies across nightlights data captured by different sensors (each spanning between 2-8 years) (Small and Elvidge, 2013).

The China Land Cover Dataset (CLCD) provides spatially explicit measurements of urban land cover change for the period 1990-2000 and is considered the country's primary national inventory of land cover (Liu et al., 2005). The dataset was derived from a mosaic of scenes from the Landsat TM/ETM sensors – 524 scenes from 1988-1990 and 512 scenes from 1999-2000. Scenes were selected for the CLCD based on low atmospheric effects and from the peak agricultural season in order to maximize the vegetation signal (Liu et al., 2005). The Landsat TM data have been aggregated into a 1X1 km^2 grid at the national level. The dataset contains measurements of total existing urban land cover (defined as a proportion of the 1X1 km^2 cell) for the years 1990 and 2000.[4]

Other researchers have examined the relationship between agricultural fertility and urban expansion in China and elsewhere. In particular, we build on the work of Deng et al. (2008) who estimate the relationship

---

[3] When scenes from high and medium resolution sensors are compiled in mosaics, discontinuities are sometimes observed across tile boundaries. The location of tile boundaries are independent of the processes of economic interest and can be treated as classical measurement error.

[4] A value of 1 = 1% of a 1km^2 grid cell = 10,000 m^2

between agricultural land rents[5] and urban land expansion using the CLCD measurements. Causal inference based on these estimates is limited by the fact that agricultural land values are almost certainly endogenous to demand for agricultural products emanating from the urban areas themselves. The identification strategy employed in this paper is designed to deal with the endogeneity of agricultural rents in the process of urban expansion. We employ an index of land suitability for agriculture (which we will shorten to "land suitability" or "agricultural suitability" throughout the paper) created by Ramankutty et al. (2002) that combines data on climate (temperature and moisture availability) and soil (soil carbon density and soil pH) to model inherent suitability for cultivation. Specifically, the authors use 1960-1991 climate data to calculate growing degree days calculated on 5-degree basis and the ratio of actual to potential evapotranspiration as a measure of moisture availability to plants. Together with global data on soil carbon density and soil pH, they use flexible functional forms from agronomic science to fit these climate and soil measures to high-resolution data on observed croplands around the world (using data from circa 1992) and finally generate an index that measures the probability that each 0.5 degree grid cell is cultivated. This index does not measure agricultural production itself (which could be affected by demand from nearby urban centers), but instead captures the exogenous physical determinants of agricultural productivity. The index is mapped in Figure 1 for China.

We proxy for access to global markets using a measure of the cost-weighted distance to sea ports, which reflects the transport costs associated with travel across China and captures topographical constraints. This measure is only a function of topography and location relative to ports and rivers and is thus independent of infrastructure. The cost distance is calculated to the world's largest ports (226 ports) using Containerisation International's rank of the largest ports by traffic volume (Fossey, 2008). For each land cell, we calculate the nearest distance to a port, and allow transport from a location to a port to happen over land, over sea, or on a navigable river or lake. In order to find the optimal path to a port, the relative transport cost between land and water transport is required. Limão and Venables (2001) use cost data on shipping a standard 40-foot container from Baltimore to different destinations around the world in 1990, and find that an extra 1,000 km by sea adds $190, whereas 1,000 km by land adds $1,380. This indicates roughly a 1:7 ratio between the cost of sea and land travel, which we use to construct the index. Land cells are assigned a transport cost of 7, while ocean, navigable lakes and rivers are assigned a cost of 1. The cost is increased proportional to terrain slope (using an FAO/IIASA map of median terrain slope in Figure 3), up to a fourfold cost increase[6] at the steepest slope category of >45%. This cost surface is then used to calculate the cost-minimizing path from each land cell to the nearest port.

The cost distance for China is mapped in Figure 2, and highlights the geographic isolation of China's

---

[5] Deng et al. (2008) use a measure of investment in the agricultural sector.

[6] Results are all robust to reducing the penalty for steepest terrain from fourfold to twofold.

western regions as well as the importance of river systems which open up parts of the interior of the country to oceanic trade. It is possible that port location is partially determined by city location, which poses a problem for identification. We develop a separate parameter that measures the cost-weighted distance to coastlines. The correlation between the ports variable and the coastlines variable is .98 and results are robust to these two versions of the variable, as well as to using the straight-line distance to ports instead of the minimized cost-distance.[7]

For the terrain slope variable, we employ the higher-resolution (90X90 meter) Shuttle Radar Topography Mission data from NASA (Jarvis et al., 2008). In addition to the three main geographic variables of interest, we employ a set of alternative geographical variables to match those used in Burchfield et al. (2006) to study urban sprawl in the U.S. Using monthly weather data at 0.5 decimal degree resolution from Matsuura and Willmott (2012), we use average 1980-1990 monthly weather to construct cooling degree-months (the total number of degrees above 18ºC summed across months of the year) and heating degree-months (the total number of degrees below 18ºC summed across months of the year). We use data from BGR/UNESCO (2008) on hydrogeology as a measure of groundwater availability for agriculture.

All of our variables are aggregated to half-degree cells across China (around 50km by 50km), since this is the resolution of the index of land suitability for agriculture. Table 2 presents the summary statistics for the variables in the study. Urban land cover and growth are expressed as percent of cell area. Note that many cells have very small (or zero) urban land cover, and the highest value for urban land cover is 35.5% (assuaging econometric concerns of upper censoring in the dependent variable). While a tremendous amount of urban land expansion occurred between 1990-2000, there are inherent limitations in modeling growth processes on decadal time scales. In particular, the variance in urban land cover growth is quite small ($s^2 = .48$) relative to that of the total observed land cover in 1990 ($s^2 = 12.18$). Table 3 shows correlations in the data. The correlation coefficient between land suitability and cost-distance to ports after the premultiplication for weighted least squares is -0.53, assuaging potential concerns of multicollinearity.

3.1 Policy Variables

Prior to the year 2000, China's central government shaped the geographic distribution of urban growth primarily through three regulatory policies: an agricultural land leasing program, deregulation and economic stimulus in coastal provinces, and agricultural land preservation regulations. If these policies are correlated with the geographic variables we focus on, then one might be concerned that the coefficients we report are biased if we do not include the policy variables. However, we agree with Démurger et al. (2002) that

---

[7] Results available upon request. We opt for using the cost-distane measure since it incorporates navigable rivers and topography which will certainly affect access to ports and international markets.

province policies (including on land use changes) are likely endogenous to the underlying geography of the provinces themselves. Thus, partialing out policies would result in an underestimation of the effect of China's geography on urban dynamics. For completeness, we include a specification in both the levels and growth analysis controlling for the relevant policies.

The land leasing system allows municipal governments to compensate farmers for agricultural land and sell long-term leases on that land to developers. This program is the primary instrument for generating revenue at the municipal level and has been shown to have a critical impact on urban growth. It was developed experimentally in four cities in 1988 (Shenzhen, Guangzhou, Tianjin, and Shanghai) and was then expanded to the rest of the country in 1992 (Wang et al., 2011). The second set of policies are the preferential policies on economic development in China's coastal provinces. Several papers have been written about their effect, though a majority of these papers proxy for these policies using a dummy variable for coastal provinces (Jian et al., 1996). We use an index developed by Démurger et al. (2002) to quantify preferential policy treatment based on the establishment of different types of special economic or other development zones. The third relevant policy is a system of agricultural land controls that restrict the conversion of agricultural land. These controls were introduced through the Basic Farmland Protection Regulation in 1994 and further developed in the Land Administration Law of 1998 (Lin and Ho, 2005). Christensen (2015) shows that the design of this policy affects its impact on agricultural conversion across provinces. Since these were introduced after 1990, we control for them in the growth regression but not in the levels regression.

4 Results

We begin studying the spatial distribution of urban land area in 1990, followed by analysis on spatial heterogeneity in the coefficients of interest by interacting them with province dummies. In the online appendix, we further examine heterogeneity by employing geographically weighted regression, as well as an MLE-based general spatial model as an alternative estimation method that explicitly models the spatial structure of the data. Our results provide robust evidence confirming the basic prediction that urban areas tend to locate where land is better suited for agriculture and with lower transport costs to ocean-based trade. However, we find some evidence that agricultural suitability is negatively associated with urban expansion during the contemporary period (1990-2000). The models that allow for spatial heterogeneity show a more nuanced picture: the positive association between urban area and agricultural productivity occurs mainly in the coastal provinces and an opposing relationship is found in some of the most urbanized provinces. This is consistent with the hypothesis that the opportunity cost of urbanizing good agricultural land increases at higher levels of urban coverage. Province-level and local Geographically Weighted Regression estimates suggest that the power of the geographical variables in explaining urban growth from 1990-2000 is limited compared to their power in explaining the distribution of urban areas in 1990.

4.1 Urban Land Cover in 1990

We estimate the relationship between geographic constraints and urban area in 1990 using the model specified in equation 1. Table 4 column (i) reports least squares estimates, weighting observations by the land area in the grid cell.[8] The least squares model is likely to present biased estimates and/or inflated t-statistics in the presence of spatial dependence and autocorrelation of errors. A Lagrange Multiplier test provides evidence of spatial dependence in the OLS residuals (significant to 1%) by comparing models with and without a spatial dependence term. Meanwhile, the spatial Durbin-Watson statistic, which tests the correlation of residuals to nearest neighbors, finds strong presence of spatial autocorrelation (values below 2.0 indicate positive autocorrelation (Gujarati, 2003) and as a rule-of-thumb values below 1.0 indicate problematic autocorrelation). The Moran's I statistic also indicates strong positive autocorrelation in both levels and growth models (within 1% statistical confidence). We report standard errors clustered by province to allow for spatial autocorrelation at the province level. In addition, we follow Conley (1999) in adjusting standard errors using a nonparametric estimator that is spatial heteroskedasticity autocorrelation consistent (SHAC) and find that our results are robust to this correction (we use a 5 decimal degree kernel, or 10 grid cells in our data, in the spatial autocorrelation adjustment). Conley standard errors are presented in square brackets, and inference throughout the paper is discussed using these standard errors when available.

The coefficient on the land suitability indicator suggests that a cell with optimal biophysical conditions for agriculture (in terms of temperature, moisture, and soil type) has more urban land cover by 2.39% of land area (0.7σ) compared to a completely infertile cell. The coefficient on cost-distance to ports indicates that a doubling of the cost-distance to a cell is associated with a .86 percentage point or .25 standard deviation reduction in urban land cover.[9] Terrain slope has a coefficient of -0.06, which means that a cell with a one standard deviation higher slope (14.6%) will have 0.9 percentage points less urban land cover, or 0.26 of a standard deviation. This very parsimonious model captures 52% of the variation in the urban land cover

---

[8] We present weighted least squares because there is significant variation in cells' land area (fixed 0.5 degree cells decrease in land area in proportion to the cosine of latitude, moreover coastal cells and border cells will have smaller areas, sometimes almost no land). It is likely that these smaller areas exhibit greater variance in outcomes, so we weigh each observation by its land area. Results are qualitatively unchanged if Ordinary Least Squares is used instead, and those results are available upon request.

[9] As an alternative (unreported but available upon request), we estimate a polynomial specification where the ln(cost distance) variable is replaced by linear and quadratic terms. The resulting coefficients of -0.0002 and 5.3e-09 indicate increasing marginal effects of distance on urban land cover over the positive domain. Though the quadratic specification yields strong significance, we opt for the log specification because it yields higher R-squared values and is thus a tighter fit for prediction. Results are also robust to using great-circle distance instead of cost-distance, though conceptually we prefer cost-distance as a measure of access to the coast since it incorporates topography and navigable rivers that open up the interior of the country.

data, suggesting that physical geography can explain a large part of the variation in urban land patterns.[10]

Our dependent variable of interest – urbanized percentage of the cell land area – is a limited dependent variable with a minimum value of zero and a theoretical maximum value of 100, although no observations are actually 100% urbanized. In column (ii), we report the results of a Tobit specification with a lower bound of zero to evaluate the sensitivity of our results to functional form assumptions. All coefficients remain strongly significant and increase in magnitude, especially in the case of land suitability.

Column (iii) includes province fixed effects that absorb any province-level omitted variables that might be correlated with our independent variables of interest and that might affect urbanization. All coefficients remain significant at the 1 percent level. The coefficient on land suitability decreases to 1.48, though the coefficient on cost-distance to ports increases appreciably to -1.53. These results suggest that geography powerfully explains urban land patterns even within Chinese provinces. Given that many grid cells have zero urbanization, column (iv) uses a logit estimate to explore the extensive margin of urbanization (in which the dependent variable is whether the grid cell has any urban land cover). The coefficients on land suitability and cost-distance to ports remain strongly predictive of whether a cell has any land cover, but terrain slope is not statistically significant in this estimate. Column (v) explores the intensive margin of urbanization by running least squares with provice fixed effects only on cells with nonzero urbanization. The results are very similar to the fixed effects specification in column (iii).

Columns (vi)-(viii) test the robustness of our estimates to the spatial scale of analysis and provide suggestive evidence about whether spillovers are at play. For example, if urban areas depend on having fertile agricultural areas within 150 km, then the analysis using 50 km grid cells will mismeasure the relationship by only considering land fertility within 50km of each observation. Column (vi) continues with the same specification, including province fixed effects, but the data has been aggregated to China's 344 prefectures. The results are noisier, but remain consistent. The coefficient on land suitability is significant at a 10 percent level and its magnitude doubles to 3.34 (compared to 1.48 in the comparable regression of column iii). This larger magnitude is consistent with the interpretation that land suitability for agriculture affects urban land patterns across an area larger than 50 km. The coefficient on cost-distance to ports continues to have a negative sign and similar magnitude, though it is not statistically significant when using prefectures. Terrain slope, on the other hand, is strongly significant and has a coefficient four times larger than in column (iii).

As opposed to column (vi), column (vii) increases the spatial resolution to 10km. This results in 85,666 observations and a higher proportion of cells with no urban land cover, so column (viii) explores the intensive margin by excluding cells with zero urbanization. In both specifications, the three geographical variables are

---

[10] The fixed effects model has an R-squared of 0.62, but the within R-squared excluding the explanatory power of the province dummies is 0.14

strongly significant and have magnitudes comparable to the analogous estimates in columns (iii) and (v). Overall, our results are robust across different spatial scales. Increasing the size of the unit of observation suggests some degree of spatial spillover in how land suitability for agriculture favors urbanization.

While we cannot rule out the role of other omitted variables, all estimates are consistent with the hypothesis that urban areas are more likely to locate on land that is characterized by prime agricultural conditions (climate and soil), access to international markets (low transport costs to sea ports), and smooth terrain. This suggests that agglomeration forces driving cities to expand into nearby agricultural land have empirically dominated the opportunity cost effects that might have steered urban expansion away from fertile agricultural land.

Table 5 presents the results of specifications that add alternative biophysical variables that have been used to study urban sprawl in the United States (Burchfield et al., 2006). These include cooling degree-months (above 18ºC) and heating degree-months (below 18ºC) as measures of amenity value of warm weather as well as energy costs for temperature control. Note that temperature is a component of the land suitability index (growing degree days), but these degree-months look to capture non-agricultural effects of temperature. Following Burchfield et al. (2006), we also include dummy variables for whether the grid cell exhibits complex hydrogeology or has shallow aquifers (the omitted category being a major groundwater basin). Column (i) shows that adding these variables to the least squares regression has no effect on the coefficients of our three variables of interest. Urban levels are not significantly correlated with cooling and heating degree months. Meanwhile, the coefficients on complex and shallow hydrogeology are strongly significant and negative, indicating that the cells over a major groundwater basin (and therefore easier access to groundwater) are likely to have more urban land area. This is consistent with the findings from Burchfield et al. (2006), who argue that the expense of expanding city water infrastructure makes groundwater availability an important contributor to the cost of urban expansion. An alternative interpretation, in line with our intuition on the role of agriculture in determining city location, is that groundwater allows for more irrigation and thus increases agricultural potential.[11]

Column (ii) adds average precipitation to the model, and finds that it is not statistically associated with urbanization after controlling for our land suitability index (we remind the reader that the index includes the evapotranspiration ratio, which quantifies the amount of moisture available to the plant). Column (iii) adds province fixed effects. The coefficients on our three variables of interest are consistent with the analagous regression in Table 4 column (iii), suggesting that they are robust to controlling for cooling and heating degree-months and precipitation. Finally, column (iv) replaces the province fixed effects with province-level

---

[11] Note that groundwater availability is not captured by the land suitability index, which uses the evapotranspiration ratio to quantify the moisture availability for plants.

policy variables that may affect urban land patterns in 1990. The first is the land leasing policy and the second is whether the province has preferencial policies for development as measured by Démurger et al. (2002). Neither policy variable is statistically significant, and the coefficients on our three variables of interest remain significant and consistent.

4.2 Urban Land Cover Growth

The results above suggest that agricultural potential, market access, and terrain slope explain a large part of how urban areas expanded historically. However, whether these factors continue to determine where growth occurs in the modern period is a distinct and important empirical question. Models that emphasize the importance of agricultural fertility as a catalyst of urban agglomeration generally feature a regional economy where trade is limited. In regions undergoing fundamental urbanization processes in the 20th and 21st centuries, modern transport costs have altered the tradeoff between natural growth of cities in fertile areas and the opportunity cost of urbanizing fertile areas since food can be brought in from afar. The majority of China's urban growth has occurred in the modern period (in terms of production and transportation technologies, and in a context of international trade), raising an important question about the extent to which agricultural fertility has defined agglomeration processes and urbanization patterns.

Column (i) of Table 6 reports estimates from equation 2. The data indicate that urban growth was more likely to occur in cells that already contained urban areas and is more likely to occur in areas with good market access (close to large ports). Unlike the levels regression above, however, the land suitability coefficient has a negative sign, indicating that areas with favorable conditions for agriculture were less likely to experience urban growth from 1990-2000. This result reflects the ambiguity expressed in our motivational discussion. Our findings are consistent with the idea that urbanization has historically occurred in areas where food could be grown plentifully nearby, resulting in a positive correlation between urban location and agricultural productivity. Large improvements in transport technology during the last century may have altered this constraint. The opposing force is the opportunity cost of converting the most productive agricultural lands, where agricultural activities have concentrated. The observed effect is that lands with higher agricultural potential in regions near existing urban centers are differentially slower to urbanize than equally proximate and less agriculturally productive lands. This illustrates a critical tension in China's land economy – productive agricultural lands must compete with increasing returns to the growth of historical centers whose initial location depended on agricultural fertility. Column (ii) adds the variables analagous to those used in Burchfield et al. (2006), and suggests that urban land conversion during this period occurred at a faster rate in regions with fewer heating degree months. Column (iii) adds precipitation and finds it not be associated to urban growth. Note that our variables of interest remain consistent in magnitude after adding these controls. Column (iv) adds the three relevant provincial policy measures relevant to urbanization and indicates that the provinces where the agricultural land controls were in place are those that experienced the fastest urban growth. This likely reflects the endogeneity of the policy as a response to the loss of agricultural land to expanding urban areas. Column (v) adds province dummies to absorb province-wide omitted variables. The result differs from column (i) in that land suitability is no longer significant, and the coefficient on cost-distance to ports doubles in magnitude and remains significant to 5 percent levels. Terrain slope is now statistically significant as well, and its negative sign suggests that steeper terrain slowed urban expansion. Columns (vi) and (vii) report estimates from specifications at different spatial scales. Column (vi) collapses the data to prefecture level. The distance to ports variable continues to be significant and increases in magnitude. Column (vii) uses 10km grid cells instead. These results suggest that urban growth occurs primarily in areas with pre-existing urban areas and that both distance to ports and steep terrain slope reduce urban growth. Column (viii) drops the grid cells with no urban land cover in 2000. The results for urban land cover remains unchanged, though the estimate of the effect of distance to ports doubles.

Overall, these results suggest that after conditioning on the amount of urban land cover in the first period, our geographic variables of interest have a less robust association with urban land cover growth. This implies that the effect of geography on urban land cover growth on a decadal time scale occurs largely through the distribution of urban land in the initial period. Both cost-distance to ports and terrain slope are sometimes significant, but the result is not robust across specifications. It may be the case that more than a single decade of observations of land cover change is required to construct a sufficiently powerful test. This will be possible with the release of the next decade of CLCD observations. Another explanation is that regional heterogeneity in the effect of agricultural fertility on urban growth means that the regressions in Table 6 miss more nuanced effects of geography on urban land cover growth. We explore this empirically below and find that this is indeed the case.

4.3 Spatial Heterogeneity in Effects

The theory that motivates our study expresses ambiguity in the empirical relationship between agricultural suitability and urban development. Results presented above suggest that the role of agricultural fertility in determining the geography of urban expansion has changed in China. However, there are a range of unobserved variables that vary across China's diverse regions that might bias these estimates, such as provincial government economic and social policies, variation in enforcement of national laws, or variation in social and cultural norms (some provinces were created along cultural boundaries). We control for some of these explicitly either by using province fixed effects or including controls for policy, but now we explore this heterogeneity more directly. We first estimate the model with province fixed effects, adding interactions of province dummies with our variables of interest in order to look at spatial heterogeneity in marginal effects. The results are displayed in Figure 5.

The differences across provinces are stark. Land suitability for agriculture has a higher effect in the eastern part of the country, with the province-level coefficient going as high as 42.2 (for comparison, the estimates above which assume spatial homogeneity were on the order of 1.5 in the least squares model with province fixed effects). Moreover, three coastal provinces indicate a negative association, which suggests that the opportunity cost of urbanizing good land is a first-order concern. It is notable that these three provinces (Beijing, Jiangsu and Tianjin, with 9%, 12% and 14% urban land cover, respectively) are three of the six most urbanized provinces. This suggests a threshold of urban land cover above which the opportunity cost of urbanizing good agricultural land truly constrains spatial patterns of urban growth. The cost-distance to ports also displays a heterogeneous association with urban land cover in 1990. The variable is significant in almost all provinces, but the elasticity is stronger at the coasts by an order of magnitude (up to -18 in Beijing, compared to the -1.5 estimated from the fixed effects model and -1.0 in the MLE spatial models in the online appendix). Clearly, empirical models assuming homogenous effects across the country overlook heterogeneity in the relationship between urban expansion and physical geography. The difference in coefficients across provinces is statistically significant; for example, the 25th percentile coefficient for land suitability is 1.3 with a standard error of 0.2, while the 75th percentile coefficient is 7.3 with a standard error of 0.4. In the case of the cost-distance to ports, the 25th percentile coefficient is -2.0 with a standard error of 0.2, while the 75th percentile coefficient is 0.6 with a standard error of 0.3. Note that this heterogeneity may be due to confounding variables that interact with our variables of interest or due to nonlinear relationships (for example, our functional form would lead to heterogenous marginal effects if opportunity costs of urbanizing fertile lands rise at high levels of urban land cover). Regardless, documenting the spatial variation in these relationships is important for prediction, policy making, and to suggest future avenues of theory and empirical research to fully characterize urbanization processes in China.

Province-level results for growth from 1990-2000 are presented in Figure 6. The spatial patterns are less systematic than in the case of the levels regression, but continue to show spatial variation. The difference in coefficients across provinces is statistically significant; for example, the 25th percentile coefficient for land suitability is -0.3 with a standard error of 0.09, while the 75th percentile coefficient is 0.5 with a standard error of 0.1. In the case of the cost-distance to ports, the 25th percentile coefficient is -0.4 with a standard error of 0.01, while the 75th percentile coefficient is -0.04 with a standard error of 0.02. Favorable land suitability for agriculture is positively associated with urban growth in around half of the provinces, though northern provinces and Beijing, Tianjin and Shandong display negative associations. As in the analysis on urban levels, the fact that some of these provinces are the most urbanized means the negative relationship between agriculture and urban growth could indicate higher opportunity costs to urbanizing fertile lands in already urbanized provinces. Meanwhile, cost-distance to ports is negatively associated with urban growth in many provinces (with Beijing, Shanghai and Jiangsu having the steepest elasticities), but much of the country has no significant relationship. Shandong, Hainan and Xinjiang have positive coefficients that defy the theories we have discussed here, though the magnitude is very small compared to the negative counterparts.

These results are consistent with the hypothesis that geographical variables play a strong role in determining city location (as in Motamed et al., 2009). The distribution of existing urban areas is clearly correlated with geographic features, while conditional on the distribution of China's urban lands in 1990, decade-on-decade urban expansion has a weaker relationship to physical geography. In both cases, these relationships vary substantially across regions.

5 Conclusion

The results of this study provide important new evidence on the relationships between agricultural suitability, access to international markets, and terrain slope on the current-day distribution of urban land cover. We highlight ambiguity in the literature regarding the expected empirical relationship between biophysical agricultural suitability and urban development. We use the case of China to distinguish between estimating the role of geography in levels of urban land area (which reflect both city genesis and historical growth) and the role in modern-day urban expansion. Urban development in China has tended to locate in fertile areas with good access to international markets. The marginal effect of port access decreases in western China, but it has played a critical role in shaping urban centers in the East.

In the case of modern-day urban expansion, the direct impact of access to ports is less substantial and the opportunity cost of agricultural fertility may outweigh its impact as a catalyst for agglomeration. We highlight this distinction because it has important implications for understanding land use policy. Our results are consistent with the idea that fertile lands exert a significant influence over city location, but that the opportunity cost associated with converting them becomes important as economies mature. Increasing returns to city growth in areas surrounded by prime farmland ultimately create a tension in the land market, a dynamic that is exemplified by the enormous political debate over the conversion of prime farmlands around productive urban centers. The central government has attempted to regulate the process of farmland conversion, and other research suggests that these interventions have had impacts on productivity (Lichtenberg and Ding, 2008; Ding and Lichtenberg, 2011).

We acknowledge some important limitations of our data, particularly with the measurement of urban land change within a single decade. Estimation will improve substantially as optical satellite archives continue to grow. We present a series of models and tests demonstrating that our primary findings are robust to spatial autocorrelation as well as potential confounding variables (such as China's land policies) and variable mis-specification (distance to coast or distance to ports vs. cost-distance to ports). We find a great deal of heterogeneity in the marginal effect of geographic constraints across China, both within and across provinces. This suggests that estimation of average effects at the national level might have some important limitations, the most critical being conflation of lack of statistically significant effects and opposing effects across regions. The literature on forecasting patterns of urban land cover change currently focuses on national level population and productivity growth, without a sophisticated parsing of constraints and catalysts that determine the spatial distribution within a country. We find that this parsimonious model of exogenous geographical determinants of urban growth explains a substantial amount of the variation in urban land cover across China, suggesting that the geographical landscape continues to play a profound role in the spatial distribution of human population and economic activity.

References

Alonso, William, *Location and Land Use*, Cambridge, MA: Harvard University Press, 1964.

Au, Chun-Chung and J. Vernon Henderson, "Are Chinese Cities too Small?," *Review of Economics and Statistics*, 2006, *74*, 549–576.

_ **and** _ , "How migration restrictions limit agglomeration and productivity in China," *Journal of Development Economics*, 2006, *80* (2), 350–388.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian, "On the Road: Access to Transportation Infrastructure and Economic Growth in China," Working Paper, National Bureau of Economic Research 2012.

Baum-Snow, N., L. Brandt, J.V. Henderson, M.A. Turner, and Q. Zhang, "Roads, Railroads and Decentralization of Chinese Cities," Working Paper 2015.

BGR/UNESCO, "World-wide Hydrogeological Mapping and Assessment Programme (WHYMAP)," Technical Report 2008.

Bosker, M., S. Brakman, H. Garretsen, and M. Schramm, "Relaxing Hukou: Increased labor mobility and China's economic geography," *Journal of Urban Economics*, 2012, *72* (2-3), 252–266.

Burchfield, M., H.G. Overman, Diego Puga, and M.A. Turner, "Causes of Sprawl: A Portrait from Space," *Quarterly Journal of Economics*, 2006, *121* (2), 587–633.

Capozza, Dennis R. and Robert W. Helsley, "The Fundamentals of Land Prices and Urban Growth," *Journal of Urban Economics*, 1989, *26* (1), 295–306.

Christensen, Peter, "The cost of conservation in an emerging economy: Satellite evidence from China's Farmland Protection Program," Working Paper 2015.

Conley, T.G., "GMM estimation with cross sectional dependence," *Journal of Econometrics*, 1999, *92*, 1–45.

Davis, Donald R. and David E. Weinstein, "Bones, Bombs, and Break Points: The Geography of Economic Activity," *The American Economic Review*, 2002, *92* (5), 1269–1289.

Démurger, Sylvie, Jeffrey D. Sachs, Wing Thye Woo, Shuming Bao, Gene Chang, and Andrew Mellinger, "Geography, Economic Policy, and Regional Development in China," *Asian Economic Papers*, 2002, *1* (1), 146–197.

Deng, Xiangzheng, Jikun Huang, Scott Rozelle, and Emi Uchida, "Growth, Population and Industrialization, and Urban Land Expansion of China," *Journal of Urban Economics*, 2008, *63* (1), 96–115.

Ding, Chengri and Erik Lichtenberg, "Land and Urban Economic Grwoth in China," *Journal of Regional Science*, 2011, *51* (2), 299–317.

Faber, Benjamin, "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System," *Review of Economic Studies*, 2014, *81* (3), 1046–1070.

Fossey, John, *Containerisation International Yearbook 2008*, London, UK: Informa Maritime & Transport, 2008.

Fujita, M. and J.F. Thisse, "Economics of Agglomeration," *Journal of the Japanese and International Economies*, 1996, *10*, 339–378.

_ , Paul Krugman, and Anthony Venables, *The Spatial Economy: Cities, Regions and International Trade*, Cambridge, MA: MIT Press, 1999.

Gollin, Douglas, Stephen Parente, and Richard Rogerson, "The Role of Agriculture in Development," *The American Economic Review: Papers and Proceedings*, 2002, *92* (2), 160–164.

_ , _ , **and** _ , “The food problem and the evolution of international income levels,” *Journal of Monetary Economics*, 2007, *54*, 1230–1255.

Gujarati, D.N., *Basic econometrics, 4th ed.*, Boston, MA: McGraw-Hill, 2003.

Hansen, Gary D. and Edward C. Prescott, “Malthus to Solow,” *American Economic Review*, 2002, *92* (4), 1205–1217.

Jarvis, A., H.I. Reuter, A. Nelson, and E. Guevara, “Hole-filled SRTM for the globe, Version 4,” *CGIAR-CSI SRTM 90m Database*, 2008.

Jian, Tianlun, Jeffrey D. Sachs, and Andrew M. Warner, “Trends in Regional Inequality in China,” *China Economic Review*, 1996, *7* (1), 1–21.

Johnston, Bruce and John W Mellor, “The Role of Agriculture in Economic Development,” *American Economic Review*, 1961, *51* (4), 137–151.

Krugman, Paul, “Increasing Returns and Economic Geography,” *Journal of Political Economy*, 1991, *99*, 483–489.

Kuznets, Simon, “Modern economic growth: Findings and reflections,” *The American Economic Review*, 1973, *63* (3), 247–258.

Laitner, John, “Structural Change and Economic Growth,” *Review of Economic Studies*, 2000, *67*, 545–561.

Lichtenberg, Erik and Chengri Ding, “Assessing Farmland Protection Policy in China,” *Land Use Policy*, 2008, *25* (1), 59–68.

_ **and** _ , “Local officials as land developers: Urban spatial expansion in China,” *Journal of Urban Economics*, 2009, *66* (1), 57–64.

Limão, Nuno and Anthony J. Venables, “Infrastructure, Geographic Disadvantage, Transport Costs, and Trade,” *The World Bank Economic Review*, 2001, *15* (3), 451–479.

Lin, George C.S. and Samuel P.S. Ho, “The state, land system, and land development processes in contemporary China,” *Annals of the Association of American Geographers*, 2005, *95* (2), 411–436.

Liu, J., H. Tian, M. Liu, D. Zhuang, J. Melillo, and Z. Zhang, “China’s changing landscape during the 1990s: Large-scale land transformations estimated with satellite data,” *Geophysical Research Letters*, 2005, *32*, 2405.

Lucas, Robert E., “Some macroeconomics for the 21st century,” *Journal of Economic Perspectives*, 2000, *14* (1), 159–168.

Matsuura, Kenji and Cort J Willmott, “Terrestrial air temperature: 1900–2012 Gridded Monthly Time Series,” *Center for Climatic Research, Dep. Of Geography, University of Delaware, Newark.*, 2012.

Matsuyama, Kiminori, “Agricultural Productivity, Comparative Advantage, and Economic Growth,” *Journal of Economic Theory*, 1992, *58*, 317–334.

Mills, Edwin S., “An Aggregative Model of Resource Allocation in a Metropolitan Area,” *American Economic Review*, 1967, *57* (2), 197–210.

Motamed, M., R. Florax, and W. Masters, “Geography and Economic Transition: Global Spatial Analysis at the Grid Cell Level,” mimeo 2009.

Murata, Y., “Engel’s Law, Petty’s Law, and agglomeration,” *Journal of Development Economics*, 2008, *87* (1), 161–177.

Muth, R., *Cities and Housing*, Chicago, IL: University of Chicago Press, 1969.

Ottaviano, G.I.P., T. Tabuchi, and J.T. Thisse, “Agglomeration and Trade Revisited,” *International Economic Review*, 2002, *43* (2), 409–435.

Ramankutty, Navin, Jonathan A. Foley, John Norman, and Kevin McSweeney, “The global distribution of cultivable lands: current patterns and sensitivity to possible climate change,” *Global Ecology and Biogeography*, 2002, *11*, 377–392.

Rostow, W. W., *The Stages of Economic Growth: A Non-Communist Manifesto*, Cambridge: Cambridge University Pressl, 1960.

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

Small, Christopher and Christopher D. Elvidge, “Night on Earth: Mapping decadal changes of anthropogenic night light in Asia,” *International Journal of Applied Earth Observation and Geoinformation*, 2013, *22*, 40–52.

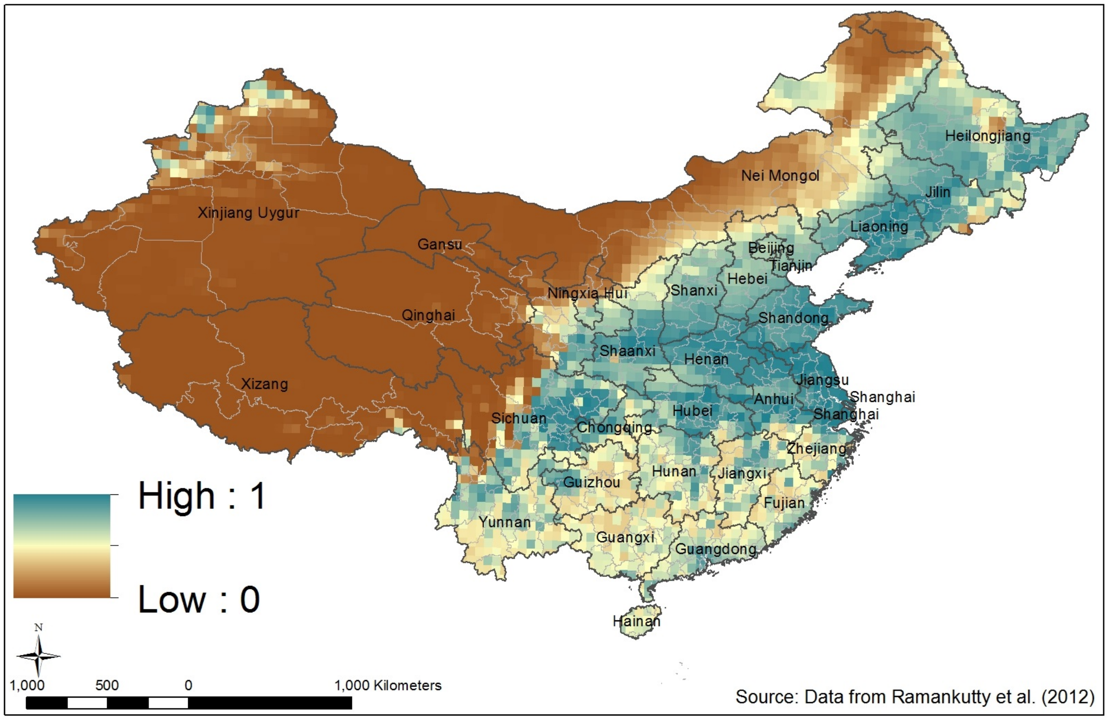
Storeygard, Adam, “Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa,” Working Paper 2014.

United Nations Population Division, *World Population Prospects: The 2010 Revision and World Urbanization Prospects: the 2011 Revision*, United Nations Publications, 2012.

Wang, De, Li Zhang, Zhao Zhang, and Simon Xiaobin Zhao, "Urban Infrastructure Financing in Reform-era China," *Urban Studies*, 2011, *48* (14), 2975–2998.

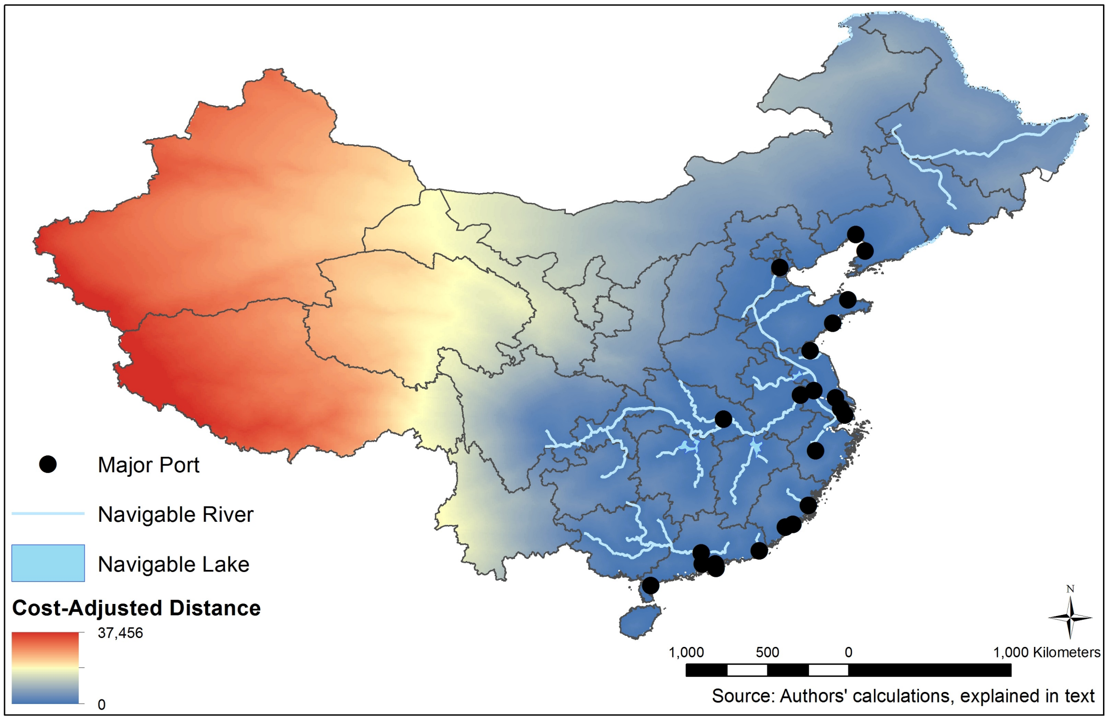
Zhang, Qingling and Karen C. Seto, "Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data," *Remote Sensing of the Environment*, 2011, *115* (9), 2320–2329.

Figure 1: Index of Land Suitability for Agriculture



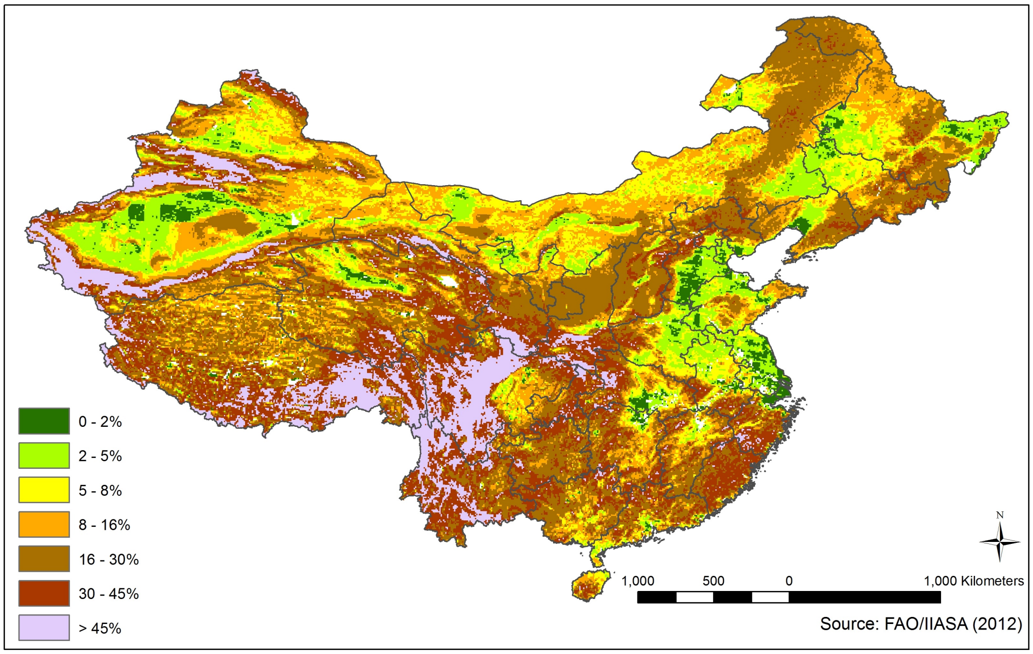
Source: Data from Ramankutty et al. (2012)

Figure 2: Cost-Adjusted Distance to Ports



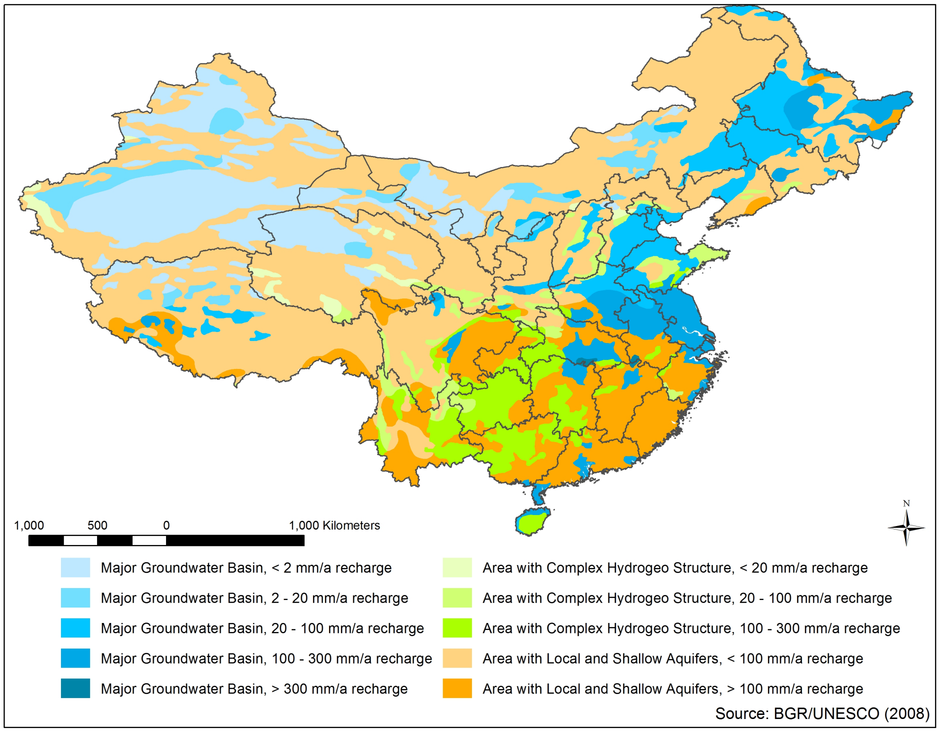
Source: Authors' calculations, explained in text

Figure 3: Terrain Slope



Source: FAO/IIASA (2012)

Figure 4: Hydrogeology



Source: BGR/UNESCO (2008)

Figure 5: Marginal Effects on Urban % of Cell in 1990 (by province)

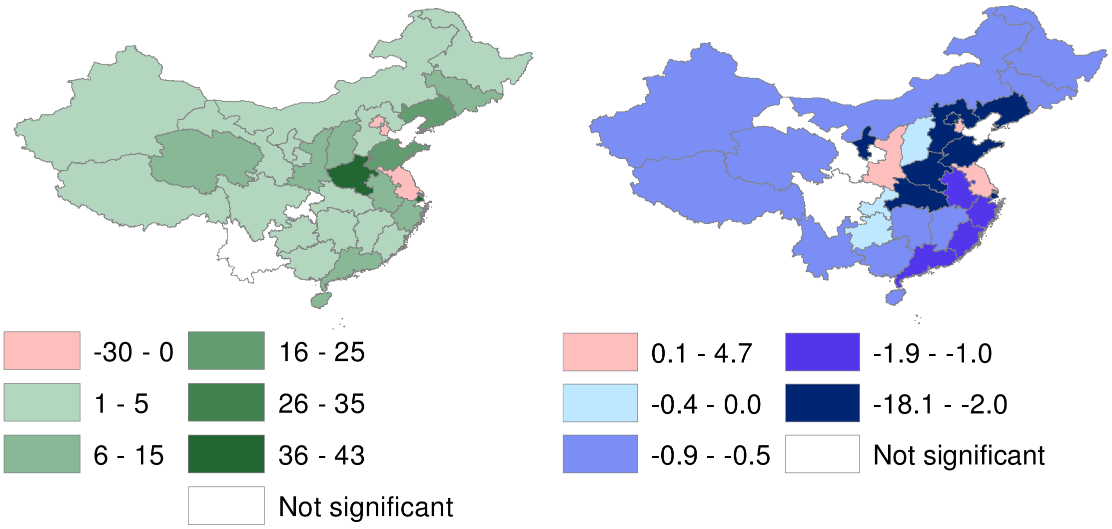


Figure 6: Marginal Effects on 1990-2000 Urban Growth as % of cell area, by province

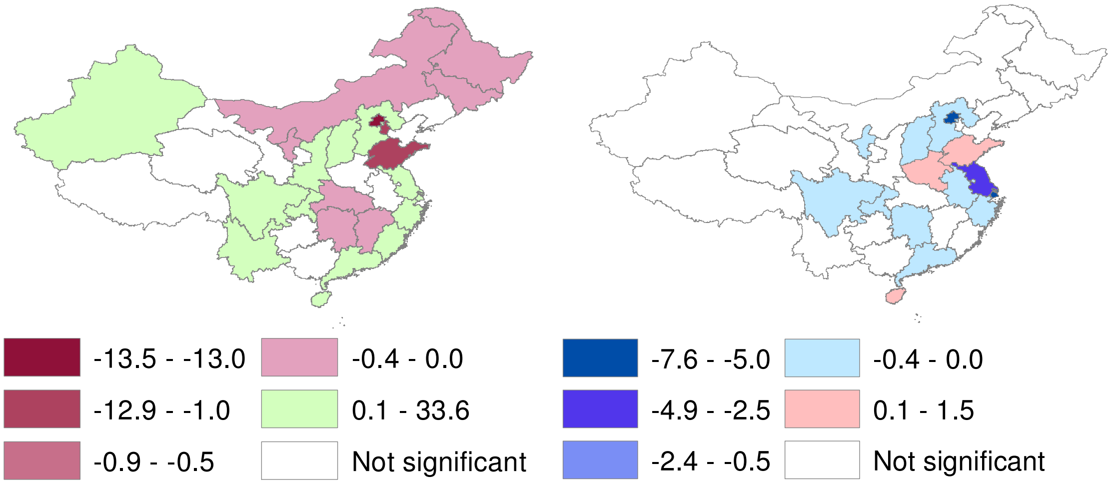


Table 1: Urban Land Cover, by province

		Urbanization in 1990	**Urbanization in 2000**	
Province	**# 0.5-degree Grid Cells**	%	%	km sq.
Tianjin	6	14.36%	15.52%	2,070
Shanghai	5	13.12%	17.88%	1,877
Jiangsu	41	11.80%	13.99%	14,186
Shandong	69	11.58%	12.39%	20,057
Henan	64	9.66%	10.40%	17,039
Beijing	8	9.04%	13.50%	2,563
Anhui	52	7.23%	8.02%	10,865
Hebei	77	5.94%	7.12%	12,718
Liaoning	71	5.51%	5.73%	8,716
Guangdong	69	3.53%	4.24%	7,303
Jilin	93	3.10%	3.18%	6,533
Zhejiang	40	2.51%	3.08%	2,866
Shanxi	65	2.36%	2.59%	4,145
Hubei	73	2.32%	2.52%	4,892
Heilongjiang	236	1.81%	1.85%	9,139
Ningxia Hui	21	1.68%	1.92%	986
Hainan	3	1.68%	1.91%	656
Fujian	48	1.60%	1.73%	1,874
Guangxi	89	1.58%	1.70%	4,199
Jiangxi	63	1.40%	1.50%	2,581
Shaanxi	82	1.28%	1.47%	3,054
Hunan	72	1.09%	1.22%	2,394
Nei Mongol	561	0.91%	0.93%	11,777
Gansu	166	0.68%	0.74%	3,044
Chongqing	31	0.48%	0.78%	644
Yunnan	153	0.45%	0.52%	2,172
Sichuan	186	0.45%	0.58%	2,880
Guizhou	65	0.29%	0.34%	602
Xinjiang Uygur	735	0.25%	0.29%	4,792
Qinghai	285	0.12%	0.13%	930
Xizang	463	0.01%	0.01%	128
China	**3,992**	**1.57%**	**1.74%**	**168,370**

Table 2: Summary Statistics

	Obs	Mean	Standard Deviation	Min	Max
Area (kms squared)	3992	2,337.3	482.3	41.9	2,863.4
Urban Land Cover in 1990 (% of grid cell)	3992	1.58	3.49	0	35.5
1990-2000 Urban Land Cover Increase (% of grid cell)	3992	0.17	0.69	0	15.6
Land Suitability for Agriculture	3992	0.35	0.36	0	1
Cost-Adjusted Distance to Ports	3992	12,187.4	10,495.8	38.3	37,118.8
Slope (%)	3992	16.3	14.6	0.4	74.6
Cooling Month-Days (above 18°C, per year)	3956	0.13	0.15	0	0.91
Heating Month-Days (below 18°C, per year)	3956	1.54	0.82	0	4.07
Precipitation (m, average per month)	3956	0.05	0.04	0.001	0.295

Table 3: Raw Correlations

	Area	Urban Cover	Urban Increase	Land Suitability	Distance to Ports	Slope	Cooling Month-Days	Heating Month-Days	Precipitation
Area (kms squared)	1.0000								
Urban Land Cover in 1990 (% of grid cell)	0.0670	1.0000							
1990-2000 Urban Increase (% of grid cell)	0.0601	0.5666	1.0000						
Land Suitability for Agriculture	0.1828	0.5405	0.3181	1.0000					
Cost-Adjusted Distance to Ports	0.5651	-0.4177	-0.2750	-0.5310	1.0000				
Slope (%)	0.2987	-0.2944	-0.1528	-0.0084	0.2507	1.0000			
Cooling Month-Days (above 18°C, per year)	0.2876	0.3746	0.2774	0.4688	-0.3345	-0.1612	1.0000		
Heating Month-Days (below 18°C, per year)	-0.0092	-0.2961	-0.2174	-0.5444	0.5151	0.0574	-0.7536	1.0000	
Precipitation (m, average per month)	0.3715	0.1540	0.1380	0.5575	-0.3048	0.3696	0.5459	-0.5238	1.0000

Table 4: Urban Land Cover in 1990

Dependent variable:	Urban Land %	Urban Land %	Urban Land %	Urban > 0	Urban % if Urban > 0	Urban Land %	Urban Land %	Urban % if Urban > 0
Independent Variables:	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Land Suitability	2.39***	4.15***	1.48***	3.86***	1.42***	2.93*	1.95**	2.05***
	(0.88), [0.78]	(1.03)	(0.64), [0.44]	(0.96), [0.73]	(0.66), [0.47]	(1.50)	(0.73), [0.49]	(0.93), [0.60]
ln(Cost Distance to Ports)	-0.86***	-1.19***	-1.53***	-1.45***	-1.49***	-0.79	-1.00***	-1.00***
	(0.23), [0.25]	(0.26)	(0.39), [0.29]	(0.48), [0.28]	(0.41), [0.32]	(0.56)	(0.32), [0.20]	(0.36), [0.29]
Slope	-0.06***	-0.07***	-0.04***	-0.002	-0.08***	-0.11***	-0.05***	-0.15***
	(0.02), [0.017]	(0.02)	(0.02), [0.01]	(0.008), [0.008]	(0.02), [0.02]	(0.04)	(0.02), [0.01]	(0.03), [0.02]
N	3,992	3,992	3,992	3,992	2,721	344	85,666	39,132
Spatial Resolution	50 km	50 km	50 km	50 km	50 km	Prefecture	10 km	10 km
Adjusted / Pseudo R-squared	0.52	0.16	0.64	0.43	0.62	0.77	0.43	0.40
Province Fixed Effects			YES		YES	YES	YES	YES
Obs weighted by land area	YES							

Regressions include a constant, not reported.

Standard errors in parentheses, clustered by province

Conley SHAC standard errors in square brackets, *** significant to 1%, ** significant to 5%, * significant to 10% levels

Model (ii) is a Tobit specification with a lower bound of zero.

Model (iv) is a Logit specification for a binary dependent variable

Table 5: Alternative Geographic Variables

	Urban Land %			
Independent variables:	(i)	(ii)	(iii)	(iv)
Land Suitability	2.42***	2.33***	1.81***	2.60***
	(0.88), [0.75]	(0.74), [0.62]	(0.61), [0.44]	(0.86), [0.77]
ln(Cost Distance to Ports)	-0.99***	-1.21***	-1.31***	-0.71***
	(0.21), [0.23]	(0.33), [0.28]	(0.34), [0.27]	(0.24), [0.23]
Slope	-0.03***	-0.01	-0.02**	-0.06***
	(0.011), [0.009]	(0.01), [0.01]	(0.01), [0.01]	(0.02), [0.02]
Cooling Degree-Months	1.67	-0.39	-1.06	
	(2.99), [3.02]	(2.92), [2.81]	(2.39), [1.92]	
Cooling Degree-Months Squared	0.40	4.34	6.38**	
	(4.43), [4.28]	(4.08), [3.93]	(3.30), [2.96]	
Heating Degree-Months	1.28	-0.04	0.43	
	(1.08), [0.79]	(1.64), [0.97]	(0.87), [0.53]	
Heating Degree-Months Squared	-0.15	0.12	-0.04	
	(0.21), [0.16]	(0.35), [0.21]	(0.18), [0.11]	
Complex Hydrogeological Structure^	-2.58***	-2.53***	-1.62***	
	(0.72), [0.66]	(0.77), [0.65]	(0.66), [0.43]	
Local & Shallow Aquifers^	-2.32***	-2.12***	-1.43***	
	(0.76), [0.64]	(0.75), [0.58]	(0.59), [0.37]	
Precipitation (mm)		-7.67	-3.60	
		(14.39), [13.25]	(8.21), [7.54]	
Precipitation squared		-76.04	-21.91	
		(66.63), [78.60]	(27.38), [31.64]	
Land Leasing Policy				-1.14
				(2.59), [1.92]
Preferential Policies for Development				0.60
				(0.88), [0.64]
N	3956	3956	3956	3992
Spatial Resolution	50 km	50 km	50 km	50 km
Adjusted / Pseudo R-squared	0.58	0.59	0.66	0.52
Province Fixed Effects			YES	
Observations weighted by land area	YES	YES		YES

Regressions include a constant, not reported.

Standard errors in parentheses, clustered by province

Conley SHAC standard errors in square brackets, *** significant to 1%, ** 5%, * 10% levels

^ Omitted category: Major groundwater basin

Table 6: Urban Growth from 1990-2000

	Percentage Point Increase in Urban Land Cover from 1990-2000							
Independent variables:	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Urban Land Cover % in 1990	0.11*** (0.02), [0.02]	0.10*** (0.03), [0.02]	0.10*** (0.03), [0.02]	0.10*** (0.02), [0.02]	0.08*** (0.03), [0.02]	0.04 (0.06)	0.10*** (0.02), [0.02]	0.10*** (0.02), [0.02]
Land Suitability	-0.19* (0.12), [0.11]	-0.18 (0.12), [0.12]	-0.15 (0.12), [0.12]	-0.14 (0.11), [0.10]	-0.01 (0.10), [0.07]	0.68 (0.55)	-0.06 (0.11), [0.07]	-0.04 (0.15), [0.12]
ln(Cost Distance to Ports)	-0.08** (0.05), [0.04]	-0.07 (0.06). [0.06]	-0.10* (0.08), [0.06]	-0.04 (0.05), [0.04]	-0.17** (0.11), [0.07]	-0.40* (0.21)	-0.15*** (0.08), [0.06]	-0.28** (0.13), [0.11]
Slope	0.0006 (0.001), [0.001]	0.0003 (0.001), [0.001]	0.001 (0.001), [0.001]	-0.0016 (0.0018), [0.0012]	-0.0017** (0.0011), [0.0007]	0.001 (0.009)	-0.0014** (0.0007), [0.0007]	-0.002 (0.002), [0.002]
Cooling Degree-Months		-0.50 (0.52), [0.57]	-0.76 (0.54), [0.61]					
Cooling Degree-Months Squared		0.70 (0.91), [1.13]	1.04 (0.89), [1.15]					
Heating Degree-Months		-0.22** (0.12), [0.11]	-0.32*** (0.16), [0.13]					
Heating Degree-Months Squared		0.050** (0.029), [0.025]	0.069** (0.039), [0.028]					
Complex Hydrogeological Structure^		-0.12 (0.14), [0.11]	-0.10 (0.14), [5.10]					
Local & Shallow Aquifers^		-0.05 (0.10), [0.08]	-0.04 (0.10), [3.27]					
Precipitation (mm)			-2.45 (2.10), [1.70]					
Precipitation squared			6.02 (9.07), [8.88]					
Land Leasing Policy				0.24 (0.23), [0.29]				
Preferential Policies for Development				0.0001 (0.07), [0.05]				
Agricultural Land Controls				0.18*** (0.07), [0.06]				
N	3,992	3,956	3,956	3,220	3,992	344	85,666	39,424
Spatial Resolution	50 km	50 km	50 km	50 km	50 km	Prefecture	10 km	10 km
Adjusted R-squared / Pseudo R-squared	0.37	0.37	0.38	0.38	0.44	0.45	0.23	0.23
Province Fixed Effects					YES	YES	YES	YES
Observations weighted by land area	YES	YES	YES	YES				

Regressions include a constant, not reported.

Standard errors in parentheses, clustered by province. Conley SHAC standard errors in square brackets, *** significant to 1%, ** significant to 5%, * significant to 10% levels

^ Omitted category: Major groundwater basin

Column (viii) limits sample to cells with nonzero urban land cover in 2000.