Sunlight, Urban Density and Information Diffusion*

Qing Zhang†

January 9, 2019

Job Market Paper

See the latest version here

Abstract

Does urban density facilitate the diffusion of information? This paper exploits plausibly exogenous variation generated by a unique national policy in China that requires all residential buildings to receive sufficient hours of sunshine. The policy creates higher degrees of restriction on density at higher latitudes, where longer shadows require buildings to be further apart. Data on individual housing projects across China reveal that the cross-latitude variation in regulatory residential Floor Area Ratio can be described quite well by a formula linking structure density to latitude through the solar elevation angle. These differences in building density further induce differences in population density and land prices across latitudes. Using differential topic dynamics on a national petition platform to measure information diffusion, this paper shows that people respond to shifts in government attention with varying speeds across latitudes. Increases in local government reply rate to a topic raises the volume of subsequent posts on the same topic, exhibiting an S-shaped time trajectory consistent with local information diffusion about shifting government priorities. These responses are systematically faster in southern cities, where density is higher. Survey evidence further indicates that otherwise similar individuals are more likely to gossip about public issues in a southern city.

*I am grateful to W. Bentley MacLeod, Suresh Naidu, and Amit Khandelwal for their continuous support and guidance. I want to thank Alain Bertaud, Don Davis, Rebecca Dizon-Ross, Junyan Jiang, Enrico Moretti, Dan O'Flaherty, Bernard Salanié, Tobias Salz, Shang-Jin Wei, David Weinstein, Jack Willis, Haaris Mateen, Anurag Singh, Yue Yu, Danyan Zha, and seminar participants at Columbia and MIT for valuable comments that greatly improved this paper. All errors are my own.

†Department of Economics, Columbia University. Email: qz2208@columbia.edu.

1 Introduction

Density is the defining characteristic of cities. The exchange of information, knowledge and ideas facilitated by urban density is believed to result in higher productivity (Jacobs, 1969) and constitute an important source of urban agglomeration economies (Rosenthal and Strange, 2004). In the words of Marshall (1890), "The mysteries of the trade become no mysteries; but are as it were in the air."

However, empirically testing for the effect of density in accelerating information transmission is challenging due to both difficulties in measuring information diffusion and a paucity of exogenous sources of variation in density. On the one hand, information is typically unobservable to the researcher. On the other hand, people choose where to live and firms choose where to locate, making density endogenous to local conditions. Despite some notable contributions, researchers are still searching for compelling evidence that establishes a link between density and information diffusion (Glaeser and Gottlieb, 2009).

This paper tackles both challenges. In response to the endogeneity challenge, I exploit a unique national policy in China that requires all residential buildings to receive two hours of sunshine a day in winter. This policy creates varying degrees of restriction on the density of buildings across latitudes, as northern buildings have to be further apart from each other due to longer shadows at higher latitudes. The resulting differences in building density exogenously induce differences in population density across latitudes. China thus provides an ideal setting to study the implications of urban density, due to the presence of this source of variation and the country's expansive territory.

To measure information diffusion, this paper uses observed dynamics of user activity on Local Leader Message Board, a national online petition platform, where a separate board exists for each city, and all postings and government replies are publicly viewable. I study how the volume of postings on a given topic responds to local government's reply rate to the same topic. The advantage of this setting is that I observe the same information shocks as the users under study, which allows me to trace out gradual increases in postings over time consistent with diffusion of information about local government's shifting priorities.

To understand the implications of the sunlight policy, I start by characterizing the behavior of developers in choosing building height given the policy. Under the assumption that it is increasingly more costly to build an extra floor, profit-maximizing developers will choose to lay out buildings in such a fashion that allows just enough sunshine to reach each building. The developer's optimal building height is then determined by a balance between increased sales from more floor space and the corresponding increases in construction costs. The increase in floor space associated with a unit increase in height is connected to the angle of sunlight, and turns out to be higher in the south. Developers in the south are thus predicted to choose to build higher buildings, all else being equal.

This prediction is tested with 2005 1% census data, which record whether a household lives in a high-rise (with at least seven floors) and the construction year of the building. Before the policy was implemented, there is no systematic difference in the prevalence of high-rises across latitudes. The prevalence of high-rises starts to diverge in 1993, exactly the year when the sunlight policy was enacted. Households living in buildings constructed after 1993 are more likely to be in a high-rise in the south than in the north, after a host of residence, household and city characteristics are controlled for. This result is robust to the inclusion of province-specific trends. That is, within the same province, southern cities constructed more high-rises than northern cities after 1993. Moreover, southern and northern cities did not experience systematically different growth rates in urban population, employment or GDP, alleviating the concern that differential economic growth may be driving the differential growth of high-rises.

The diverging prevalence of high-rises suggests that the policy was enforced. Compared to building height, we are more interested in how much floor space is allowed to be built, which determines housing supply. The density of floor space can be succinctly summarized in the value of Floor Area Ratio (FAR), defined as the amount of floor space built on one unit area of land. Since developers' optimal FAR is linked to the solar elevation angle, the optimal FAR level varies across latitudes.

The Chinese government uses FAR upper limits as the primary tool to restrict urban density (Brueckner et al., 2017), as is also the case in India (Bertaud and Brueckner, 2004; Harari, 2017), Japan (Asami et al., 2006), and many other countries around the world (Brueckner and Lall, 2015). If the government seeks to maximize revenues from land sales, as could be a reasonable first-order approximation in the Chinese context (Chen and Kung, 2016), it will set regulatory FAR at developers' optimal level. A comprehensive dataset of land transactions covering a majority of new housing development in urban China during 2007-2016 and containing FAR upper limits for individual projects reveals that the cross-latitude variation in regulatory FAR can be described quite well by a formula implied by developers' profit-maximizing behavior. Including a rich set of controls at project and city levels does little to change the correlation between observed FAR and that predicted by the model.

The strong correlation between observed and predicted FARs suggests that the sunlight requirement is often the binding contraint on density, surpassing other considerations such as ventilation, views, parking space, fire safety and households' demand for open space. To see whether the sunlight requirement is *always* a binding constraint on density, I explore deviations from the policy, and find that the regulatory FAR of housing projects on larger plots and in richer cities has a significantly weaker correlation with that predicted by the sunlight policy. These projects are subject to stricter development restrictions (i.e., lower FARs) than required by the sunlight policy, possibly reflecting higher demand for open space by richer households and better transportation infrastructure in richer cities. All in all, my estimates suggest that moving across the interquartile range of latitudes, regulatory FAR would increase by 15%, all else being equal.

Under these exogenously set FAR restrictions, similarly dense 1-km^2 grid cells in 1990 have diverged in terms of population density by 2010, with southern cells becoming significantly denser. This holds within provinces and when controlling for initial city characteristics. A standard urban economics model with homogeneous households predicts higher land prices where population density is higher. Consistent with this prediction, I find that among grid cells with similar population densities in 1990 and similar nighttime light luminosities in 1992, those in the south had higher residential land prices during 2007-2016, the sample period of my land transactions data. Cells in the south also had higher commercial land prices, consistent with a model featuring an integrated land market where land is allocated to its most profitable use. In a world where the residential and commercial land markets are separate due to zoning restrictions, this is consistent with greater market access to consumers in denser places being capitalized into commercial land rents, which has been modeled in Ahlfeldt et al. (2015) and Donaldson and Hornbeck (2016), among others. Interestingly, I find that commercial FAR is also higher in the south, even though the sunlight policy does not apply to commercial development. This can be explained by an elastic supply of commercial floor space in response to higher rents.

The exogenous variation in density driven by the sunlight policy is then used to examine the relationship between density and information transmission. I examine user behavior dynamics on Local Leader Message Board, exploiting differential trajectories of posts on different topics (issues) discovered with an LDA model (Blei et al., 2003). There is clear evidence of responses to shifts in government attention. After the local government increases their reply rate to a certain topic, volume of posts from the city on the same topic gradually increases, exhibiting an S shape over time, which is suggestive of a local diffusion of information about the government's shifting priorities, where people learn these shifts from the replies and increase their postings due to either strategic considerations or psychological encouragements. Consistent with this interpretation, I find that grievance-related topics are significantly more responsive to changes in government reply than non-grievance topics. The observed responses are not caused by a reverse causality where the local government replies more to more frequently raised issues, as increases in postings do not anticipate increases in reply rate.

Consistent with density contributing to more rapid information transmission, posts in southern cities respond systematically faster to changes in government reply than in northern cities. This is true when a host of city-level economic, education and telecommunication infrastructure characteristics are controlled for, and when only using differential responses within provinces. In terms of magnitude, estimates suggest that in the first two quarters after a reply rate increase, a city at the 25th percentile of latitude would experience a cumulative increase in postings that is 2.7 times as large as a city at the 75th percentile of latitude. Survey data further provide direct evidence that similar individuals with similar levels of interest in public issues are more likely to gossip and hear gossip about such issues in the south than in the north, lending support to the conjecture that density is conducive to the spread of word of mouth.

To address alternative explanations of the observed patterns, I present evidence that citizens do not differ systematically across latitudes in their attitudes towards the government, as measured by their beliefs about the role of government in various domains and trust in various government institutions, addressing the concern that different tendencies to use the platform may be driving differential responses across latitudes. There is also no evidence of systematic cross-latitude differences in internet using behavior, alleviating the concern that information may be transmitted online instead of offline.

This paper contributes to the literature on the relationship between density and knowledge spillovers, which has been studied both in the context of agglomeration economies (Rosenthal and Strange, 2004) and human capital externalities (Moretti, 2004). Density is an endogenous outcome for which reasonably exogenous sources of variation are not easy to find. Exceptions are Combes et al. (2010), who use soil quality to instrument for historical settlement patterns, and Rosenthal and Strange (2008), who use geology (landslide hazard, seismic hazard, and the presence of sedimentary rock) to instrument for the prevalence of high buildings. I introduce a new instrument that is shown to have a robust first stage. To my knowledge, Bertaud (2018) and Zhang (2017) are the only works in economics that have discussed the sunlight policy in China, and neither of them uses detailed project-level data to verify that the policy is indeed currently enforced and appears to be a binding constraint on urban density.

The transmission of information and knowledge is hard to observe, prompting scholars to rely on implications of knowledge spillovers including variation in wages and rents across cities with different characteristics (Rauch, 1993; Glaeser and Mare, 2001) or the location choice of firms (Audretsch and Feldman, 1996; Arzaghi and Henderson, 2008) instead of directly measuring them. An important contribution is Jaffe et al. (1993), who study the paper trail of patent citations to identify knowledge spillovers. I measure information diffusion using postings on a public platform, which has the benefit of being directly observable. While this specific type of information is not directly related to productive activities, it has potential implications for government accountability and the quality of governance.[1]

This paper also contributes to the literature on the consequences of residential development regulations (Gyourko and Molloy, 2014; Turner et al., 2014). While development restrictions are usually complicated and varied, necessitating use of indirect measures such as the gap between housing price and marginal construction cost (Glaeser et al., 2005), FAR upper limits readily quantify the degree of restriction on floor space construction. Unlike in the US where zoning ordinances are set at local level, influenced by the interest of local residents and therefore more likely to be plagued by endogeneity, the sunlight policy in China is a national policy that

[1] Campante and Do (2014) shows that in US states where the capital city is situated in a sparsely-populated area, newspapers cover state politics less, voters have less interest and knowledge about state politics, and the level of corrpution is higher. In an international context, Campante et al. (2014) shows that isolated capitals are associated with misgovernance in less democratic countries.

exogenously shifts urban density across a country with the world's largest urban population. In addition to its broad impacts, the centralized nature of the system makes available a dataset that covers a majority of new urban construction projects across the country during the last decade. Drawing on this and other sources, this paper shows that zoning regulations can have measurable impacts on urban density and the speed of information diffusion.

The paper proceeds as follows. Section 2 discusses the motivation behind the sunlight policy, how the policy is enforced, as well as some background on the urban land market. Section 3 describes the main datasets used in this paper. Section 4 derives a relationship between latitude and optimal FAR and empirically looks at the effect of the sunlight policy on building height. The section also shows that regulatory FAR across China can be characterized quite well by a formula linking FAR to latitude through the solar elevation angle. Section 5 shows the dynamic responses of postings to changes in government reply rate, and reports the systematic differences in response speed across latitudes. Section 6 discusses alternative explanations that could give rise to the observed patterns. Finally, Section 7 concludes.

2 Institutional Background

2.1 The Sunlight Policy: Origins and Evolution

Sunshine is an important factor in the culture of *feng shui*,[2] and is valuable to Chinese households for practical purposes including drying clothes[3] and keeping homes warm in winter.[4] Soviet building practices[5] could also have influenced Chinese policymakers' perception about the importance of citizens' right to sunshine. However, there was no law governing building distances in relation to sunlight prior to the 1990s, partly because the pre-1990s urban population was relatively small and land value was low, making sunlight easily guaranteed for residents.

With the unleashing of the commercial real estate market after 1998 and the beginning of large-scale housing construction (Wang, 2011; Fang et al., 2016), the sunlight right of residents started to be taken more seriously. *Urban Residential Planning and Design Ordinance* (GB50180-93) issued by the Department of Construction in 1993 gave this policy a legal status. It requires that the lowest level of any residential building receive at least two hours of sunshine during a specific winter day on the Chinese lunisolar calendar.[6] Having undergone a

[2]A traditional system of beliefs and practice regarding residence location and layout.

[3]Dryers are a very rare sight in contemporary China. A survey (Brockett et al., 2002) conducted in 1999 to 251 households in five Chinese cities across climate zones found that only two households owned a dryer.

[4]Centralized heating is only available to the northern part of the country (Almond et al., 2009).

[5]According to a report by a US delegation to the USSR (Wright, 1971), "Two prime planning concerns for the Soviets are light (three hours of sunlight per room at the March solstice), and breathing room."

[6]This day is Major Cold for most cities, which is the day when the sun is at 300 degrees ecliptic longitude and falls between January 19 and January 21 every year. There are some alternative specifications depending on the climate zone the building is in and whether the building is in a "large" city or a "small or median" city. Buildings in "small or median" cities have to receive three hours of sunshine on Major Cold, unless they are in climate zones IV, V or VI, which correspond roughly to the very south of China and the Tibetan Plateau, in

few revisions, this ordinance is still effective today, wherein the sunlight requirement has not changed since 1993.

Local building code issued by provinces and prefectures respect and reflect this national policy. These documents often contain clauses governing the distance between buildings in relation to sunlight, and in many cases explicitly stipulate a "sunlight distance coefficient", referring to the required ratio of distance between two adjacent buildings to building height. I collect these coefficients from over 200 prefectural building code, and plot them against the latitude of the prefectures in Figure 1.[7] There is a strong positive correlation between the sunlight distance coefficient and latitude - buildings in northern cities are required to be further away from each other. The variation across cities is large. Buildings in Changchun (43.79°N) are required to be 1.95 times as far apart as buildings in Kunming (25.19°N).

Theoretically, the minimum sunlight distance coefficient that guarantees sunlight for the first floor of each building should be equal to the cotangent of solar elevation angle at a location, which depends on the latitude of that location.[8] This theoretical sunlight distance coefficient curve (equal to the cotangent of solar elevation angle) is overlaid in Figure 1, and achieves an R^2 of 0.99 with the coefficients specified in actual local regulations.[9]

Are these local regulations followed in practice? Having a construction plan that satisfies the sunlight requirement for each building is a prerequisite for obtaining a building permit.

which case they have to receive at least one hour of sunshine on the day of Winter Solstice.

[7]Specifically, shown in this figure are sunlight distance coefficients in different cities for residential buildings below seven stories in new urban areas. Buildings above seven stories are regulated via minimum distances (in meters) between buildings. There is typically a separate sunlight distance coefficient for old city centers where the sunlight requirement is slightly relaxed, the spatial extent of which varies from city to city. In the sample, the mean ratio of central city sunlight distance coefficient to new urban area sunlight distance coefficient is 0.86.

[8]Let ϕ denote the latitude (in radians). The formula linking solar elevation angle to latitude is

$$\sin\alpha = \sin\phi\sin\delta + \cos\phi\cos\delta\cos h$$

where α is the solar elevation angle. δ is the declination of the sun, which on the day of Major Cold is roughly -20 degrees ($-\frac{\pi}{9}$). h is the hour angle. In order for a building to receive at least two hours of sunshine on that day, sunlight must reach the building 11 am at the latest. We therefore evaluate α at 11 am and set $h = -\frac{\pi}{12}$ (-15 degrees). Therefore,

$$\alpha(\phi) = \arcsin(\sin(-\frac{\pi}{9})\sin\phi + \cos(-\frac{\pi}{9})\cos(-\frac{\pi}{12})\cos\phi)$$

[9]It is worth noting that cities in southernmost and northernmost China seem to follow the national rule less closely than the rest of the country, at least according to these coefficients specified for buildings below seven stories. Cities in northeast China (north of 40°N) set local building distance coefficients systematically lower than that required by the national sunlight rule, due to the high cost of providing utilities and infrastructure for buildings too far apart, which is particularly challenging as northeastern municipal governments generally face varing degrees of financial distress (See Letter Regarding Compliance of Urban Residence Sunlight Standards in Heilongjiang, an official document issued by the national managing committee of the sunlight policy after their field trip to the northeast in 2007). On the other hand, cities to the south of 25°N typically set their sunlight distance coefficient at 1.0, even though in theory they can further decrease the distance without violating the sunlight right of residents, presumably because other considerations in determining building distances (ventilation, views, green space, parking space, fire safety, etc.) start to kick in when buildings are close enough to each other.

Enforcement of the rule is sophisticated and involves specialized computer software. In many cities, developers are required to use an officially-sanctioned sunlight analysis software to model the changing sunlit conditions of buildings throughout the day, and regulators use this software to evaluate hours of sunlight exposure building by building.[10]

The fact that citizens can take their sunlight-blocking neighbor to court further increases the enforceability of the sunlight policy. The right to sunlight was enshrined in the Property Law enacted in 2007. Since 2007, there have been at least 2,000 court cases across China regarding sunlight rights.[11][12]

In Section 4.3, I present evidence from the regulatory floor area ratio specified for a large fraction of all urban housing projects constructed across China between 2007 and 2016, which supports the notion that these local sunlight requirements are followed quite closely in setting regulatory floor area ratio for individual housing projects. Floor area ratio is defined as the ratio of total floor space on a piece of land to the area of that piece. FAR will be high when buildings are either high or take up a large fraction of land on which they sit. As such, FAR directly measures building density.

2.2 The Urban Land Market

Residential construction in urban China is predominantly channeled through real estate developers instead of individual home buyers. Local governments transfer large land lots to developers, who in turn build apartments that they sell to home buyers.[13] My dataset sheds light on this phenomenon. The median size of a greenfield residential lot in the sample is 1.46 hectares (157,153 square feet), which is much larger than the lot needed for building a typical single-family residence. This number is 3.21 hectares in the 314 prefectures, which are the larger cities. The 20th percentile of the size of a greenfield residential lot is 0.07 hectares (7500 square feet). If we think of 0.07 hectares as about the right size for building a single-family residence, lots larger than 0.07 hectares account for 99.9% of total area transferred. Because the lots are transferred to developers instead of ordinary home buyers, they face financial incentives to build up to the legally allowed FAR upper limit. This justifies studying the regulatory FAR upper limit as a key determinant of urban density in China.

Urban land in China is state-owned. During urban expansion, local governments expropriate rural land surrounding cities from farmers, at which moment they obtain ownership over that

[10] Sunlight Parameter Standards for Buildings (GB/T 50947-2014), Ministry of Housing and Urban-Rural Development, PRC, 2014.

[11] I search for court cases whose cause is "neighboring sunlight dispute" on three of China's leading court case disclosure platforms, and find over 2,000 cases on each of the three platforms. The three platforms are *openlaw.cn*, *China Judgements Online* and *itslaw.com*.

[12] In a headline news story, a man won a case against Zifeng Tower, the tallest skyscraper in Nanjing, for blocking sunlight to his apartment.

[13] Fang et al. (2016) uses the fact that most housing construction is organized as large residential communities to propose a housing price index that measures time variation of price within communities.

land. They then transfer the use right of land to developers or industrial firms for commercial, industrial and residential development. The tenure length of this use right is generally fixed and varies by use type (40 years for commercial, 50 years for industrial, 70 years for residential).

There are three legal procedures in which local governments transfer land to users: allotment, auction and contract. Allotment (22.6% of transactions in the sample) applies to land for public infrastructure and buildings, where the lots are transferred at close to zero cost.[14]

Auction is the legally required means for transferring commercial and residential land. Under the most prevalent auction form (*guapai*), the lot for transfer is advertised beforehand, and individuals or firms who want to bid for the lot post their bids online after passing local government's qualification checks and making a deposit. The auction generally lasts ten days and happens online, during which period bidders can continually submit new bids, and the highest bid is updated online. At the end of the bidding period if there are two or more bidders who offer the same price, an on-site auction is held to determine the winner.

Finally, contract mainly applies to transactions involving changes of land status. Section A.1 contains a detailed discussion of situations where this is applicable.

3 Description of Data

To increase transparency in land transactions, the Ministry of Land and Resources[15] requires local governments to report all land transactions, which are subsequently disclosed on the Ministry's website.[16] Disclosed information includes location, zoned use, price, size, regulatory FAR upper limit (and in some of the industrial land transfers lower limit), identity of receiving party, means of transfer, land grade,[17] date of contract, etc. All the transactions are publicly viewable. I scraped all records from the website in early 2017. The resulting dataset contains over 1.3 million transactions from across China. Since all construction on greenfield urban land has to obtain land from the government and therefore is registered in the dataset, I observe close to the universe of new urban construction in the 2007-2016 period.[18]

Of the land parcels transfered in or after 2007, around 40% of them are zoned for residential uses. These residential parcels have a median price of 318 RMB per square meter and a median FAR upper limit of 2.29. Using Baidu Map API, I geocoded all parcels, translating their addresses into longitudes and latitudes. 98.4% of parcels were successfully located.

[14] 66.5% of allotted lots in the sample involve zero payment, excluding observations where the price is missing, which is more likely than not to be missing because it is zero.

[15] In 2018 the Ministry was consolidated into a new Ministry of Natural Resources.

[16] http://landchina.mlr.gov.cn.

[17] Government-accessed land quality, which reflects, among other things, whether the lot is connected to road and telecommunication networks and serviced by utilities including heating, electricity, water, gas and sewage.

[18] I cross-checked total area of land transfered in my dataset with official numbers published in Land and Resources Yearbooks. My dataset covers 69.8% in area of all land transfered during 2007-2015. Land transactions prior to 2007 are not covered well by the Ministry of Land and Resources website. Therefore I focus my analysis on transactions that happened in or after 2007 in this paper.

Another source of data I utilize is the 2005 1% population census administered by the National Bureau of Statistics. The census uses a three-stage cluster sampling method to ensure representativeness (Cui et al., 2005) and contains rich information at both household and individual levels. Important for my purposes, the census contains information on whether a household lives in a high-rise (buildings with at least seven floors) and construction year of the building.

To examine information diffusion, I make use of postings on the Local Leader Message Board,[19] a national online platform where people register their complaints and suggestions to government officials in their local area. The platform is managed by People's Daily, China's central official medium, and contains sub-boards for all administrative units (provinces, prefectures and counties). Local governments can respond to messages posted to their board, but ordinary users cannot respond to each other's postings. All the postings and government replies for all localities are openly viewable.[20] In an earlier project (Jiang et al., 2018), we scraped the contents of all postings and replies up until early 2016 along with auxiliary information including time of posting and locality posted to.[21] The resulting dataset contains around 900,000 postings.

Other sources of data, including 1-km^2-resolution population density, nighttime lights, terrain ruggedness, and climatic conditions, are summarized in Table A1.

4 Effects of the Policy on Building Height and FAR

Developers want to maximize their profits under the sunlight constraint. In order to guarantee sunlight to the next building, taller buildings have to be accompanied by larger distances between them. Even so, this section shows that increasing building height will increase total floor space available under the sunlight constraint. The marginal benefit of increasing building height will be larger in the south than the north, all else equal. Consequently, developers in the south will optimally choose to build taller than developers in the north. I then test this hypothesis using 2005 census data recording whether households live in a high-rise (with at least seven floors) and construction year of the buildings.

We are ultimately more interested in residential FAR than building height, as FAR determines the amount of residential space available. The analysis of developer behavior also implies that residential FAR will be higher in the south than in the north, and yields a formula that links latitude to residential FAR. Taken to housing projects constructed across China 2007-

[19] http://liuyan.people.com.cn.

[20] Interviews with LLMB staff suggest that censorship on the platform is much less severe than for social media, partly because the platform is managed by a central organization who has little incentive to cover up for local officials, and indeed may have an incentive to elicit information about local situations.

[21] We matched messages posted to provincial boards to cities by finding city names in the text. In cases where multiple cities are mentioned, the most frequently mentioned city is chosen. Only 8.7% of posts cannot be matched to a city.

2016, this formula proves to describe the cross-latitude variation in regulatory FAR set for individual housing projects quite well.

4.1 Developer Behavior under the Sunlight Policy

Developers operate in perfectly competitive markets. At location i, they take housing price there p_i^H as given. Revenue from constructing floor space F_i^r will thus be $p_i^H F_i^r$. Suppose lot size is fixed by the government.

Construction cost $CC(F_i^r, h_i)$ depends on total floor space constructed F_i^r and the height of buildings h_i, and satisfies $\frac{\partial CC}{\partial F} > 0$, $\frac{\partial CC}{\partial h} > 0$. It is increasingly more costly to make buildings taller, so that a low building hosting the same floor space as a tall building will be less costly to build.

The developer's profit maximization problem can be considered in two steps. She chooses a layout of buildings that minimizes construction costs to achieve a given desired FAR (i.e., total floor space), and she then chooses a profit-maximizing FAR.

The configuration that minimizes construction costs to achieve a given FAR under the sunlight policy is to place buildings in ranks and files reminiscent of Le Corbusier's *Radiant City* plan such that there is just enough room for sunlight to reach the first floor of each building.[22] Figure 2 illustrates this configuration. Leaving more space between buildings will have to be compensated with taller buildings to achieve the same amount of floor space, which is more costly as the marginal cost of height is increasing.

We calulate FAR associated with the layout on the left of Figure 3 where the length of a building takes up one full dimension of the lot. Although it may seem odd to construct such elongated buildings, note that a more realistic staggered layout such as that on the right of Figure 3 actually achieves the same FAR.

The relationship between distance d and height h is given by solar elevation angle α,

$$h = d\,\tan\alpha$$

Let l be the length of a parcel and l_0 be the width of a building. l_0 is fixed due to housing unit layout considerations. Then the number of buildings that can be constructed on the parcel is $N = \frac{l}{l_0+d}$. Floor area ratio is equal to the fraction of land taken up by buildings multiplied by number of floors. Let h_f be the height of a floor, which is also fixed. We have

$$\text{FAR} = \frac{Nl_0}{l}\frac{h}{h_f} = \frac{l_0}{h_f}\frac{h}{l_0 + \frac{h}{\tan\alpha}} = \frac{l_0}{h_f}\frac{1}{\frac{l_0}{h} + \frac{1}{\tan\alpha}}$$

.

We see that under this configuration of buildings, it is possible to increase floor space by

[22] We ignore idiosyncrasies in the shape of buildings, e.g., setbacks.

increasing building height while adhering to the sunlight policy. Building taller is however more expensive. Let S be lot size. The marginal benefit of increasing height is given by

$$\text{MB}_h = p^H S \frac{d\text{FAR}}{dh} = \frac{p^H S}{h_f} \left(\frac{l_0}{l_0 + \frac{h}{\tan\alpha}} \right)^2$$

We see that the marginal benefit of height is decreasing in h, i.e., total revenue as a function of height is concave. Moreover, the marginal benefit of height is increasing in solar elevation angle α, which implies larger marginal benefits of height in the south.

The marginal cost of height is given by

$$\text{MC}_h = \frac{dCC(F^r, h)}{dh} = \frac{\partial CC}{\partial F^r} \frac{dF^r}{dh} + \frac{\partial CC}{\partial h} = \frac{\partial CC}{\partial F^r} S \frac{d\text{FAR}}{dh} + \frac{\partial CC}{\partial h}$$

Suppose that holding height fixed, the marginal cost of building one extra square foot is constant, i.e., $\frac{\partial CC}{\partial F^r} = \eta$. The developer's optimal building height under the sunlight policy is a level h^* such that $\text{MB}_h = \text{MC}_h$, i.e.,

$$\frac{(p^H - \eta) S}{h_f} \left(\frac{l_0}{l_0 + \frac{h^*}{\tan\alpha}} \right)^2 = \frac{\partial CC}{\partial h} \tag{1}$$

From this, we see that given the sunlight policy, developers in the south (under larger α) will choose to build taller than their counterparts in the north, under the same construction cost schedule determined by local wage levels and construction technology and the same housing price.

4.2 Evidence from Building Height

In this subsection, I test the hypothesis that developers in the south will build taller under the sunlight policy *ceteris paribus* using census data from 2005, which record whether a household lives in a high-rise (with at least seven floors) and construction year of the building. Specifically, I estimate the following equation:

$$\text{High-Rise}_{hct} = \lambda_{\text{prov}(c)} + \gamma_1 \text{Lat}_c + \gamma_2 \text{Post}_t + \gamma_3 \text{Lat}_c \times \text{Post}_t + \beta_1' X_c + \beta_2' X_h + \epsilon_{hct}$$

High-Rise_{hct} is an indicator for whether household h in city c lives in a high-rise conditional on the building being constructed in year t. $\lambda_{\text{prov}(c)}$ are province fixed effects, which control for province-specific determinants of building height. Post_t is an indicator for whether the building construction year is no earlier than 1993, the year when the sunlight policy was written into law. X_c is a vector of city-level controls in 2005 that captures determinants of housing prices and local wages,[23] including urban population, population density, employment, GDP per capita

[23]Direct data on housing prices and wages are not available for counties.

and share of land with above 15% slopes within a 10-km radius from city center.[24] Cities surrounded by rugged terrain tend to have a less elastic housing supply and higher housing prices (Saiz, 2010).

Ideally we would like to observe all buildings built in all years, but the nature of population census data determines that we can only observe occupied buildings. To address the concern that households in the north and the south sort into buildings of different vintages differentially, the equation includes a vector of household-level controls X_h. These include both residence characteristics and demographics. Residence controls contain use porpose of residence (just for living or for both living and production), log number of rooms, log square footage, whether shared with other households, access to tap water, presence of kitchen, toilet and bath facilities, type of fuel and how the residence was obtained (rent or bought, SOE housing, affordable housing or commercial housing). Demographics contain log number of household members and the following characteristics of household head: gender, age, ethnicity, *hukou* registration place, type of *hukou* (rural or urban), literacy, education, whether worked for more than one hour last week, log income, source of income and marital status.

In other words, we test whether similar households residing in similar residences in similar cities have different probabilities of living in a high-rise depending on latitude of the city, and whether that difference only occurs for buildings built after the sunlight policy was put in place. Table 1 shows regression estimates. Column 1 is a parsimonious specification containing just province fixed effects. We see that households living in buildings constructed after 1993 are over 30% more likely to be in a high-rise, reflecting China's urbanization process. The coefficient on latitude is not significantly different from 0, suggesting that the prevalence of high-rises was not different across latitudes before the sunlight policy. However, the negative coefficient on the interaction of post and latitude suggests that buildings constructed after the policy are more likely to be high-rises in the south than in the north. Magnitude of the coefficient indicates that across the interquartile range of latitudes (moving from 39.2 degrees to 25.4 degrees), households would become 5.5% more likely to live in a high-rise conditional on the building being constructed after 1993.

Column 2 adds 2005 city-level controls. Coefficient on post decreases slightly, suggesting that part of the increase is driven by changes in city composition within different building vintages. However, the interaction remains negative and significant. It is worth noting that surrounding slopes significantly increase building height, as expected from the model where slopes increase housing prices and therefore the marginal benefit of height. Columns 3 and 4 add residence controls and household controls consecutively. The coefficients remain quite stable. Column 5 further includes city fixed effects, which absorb latitude and city-level controls. We see that the coefficients of interest remain stable.

Column 6 includes both province fixed effects and their interactions with the post dummy.

[24]City center is defined as the brightest 1-km cell in terms of 2005 nighttime light. In cases where multiple cells are equally bright, their geometric center is calculated.

This allows for province-specific trends in building height, and identifies the effect of policy solely from within-province differential growth across latitudes. This limited variation induces much less precise estimates. Reassuringly, the point estimates remain almost unchanged. This indicates that even within the same province, a southern city exhibits faster increase of high-rises than a northern city after the policy was in place.

Is the differential growth in building height across latitudes caused by differential economic growth? City-level economic outcomes are not available before 1994, which prevents me from directly including them in the regression. However, we can test whether they exhibit differential growth in the period we do have data for. I estimate the following equation:

$$y_{ct} = \lambda_c + \delta_t + \gamma_t \times \text{Latitude}_c + \epsilon_{ct}$$

where y_{ct} is an economic outcome (log urban population, log employment, log GDP or log GDP per capita) for city c in year t. γ_t captures any differential economic growth across latitudes. Each regression is run on a panel balanced on the outcome, i.e., cities with any missing values in the outcome are excluded from the regression.

Table 2 shows the estimated set of γ_t. From columns 1 to 4, we see that there is no evidence of differential growth across latitudes in urban population, employment or GDP, although GDP per capita does seem to grow faster in the north between 2002 and 2005. Columns 5 to 8 include province by year fixed effects to detect differential economic growth within provinces. There is no evidence for differential growth in any of the four variables within provinces. Taken together, these findings suggest that the differential changes in the prevalence of high-rise buildings are likely due to differential incentives created by the sunlight policy.

Could this differential change in building height have been caused by other policy events than the sunlight policy? Afterall, mid- to late-90s was also the time when the government privatized housing and kick-started the real estate sector. Fang et al. (2016) indicates that housing construction only started to grow rapidly after 1998, when the central government completely abolished the traditional model of providing housing as an in-kind transfer to SOE employees, and started to provide subsidized residential mortgages. I examine the timing of building height change in the data to see whether the change appears to have happened in 1998 or in 1993 when the sunlight policy was introduced. To be specific, I estimate

$$\text{High-Rise}_{hct} = \lambda_c + \delta_t + \gamma_t \text{Latitude}_c + \beta_1' X_c + \beta_2' X_h + \epsilon_{hct}$$

Figure 4 plots the trajectory of $\delta_t + \gamma_t \text{Latitude}_c$ for a northern city (75th percentile of household latitude) and a southern city (25th percentile of household latitude). We see that the two trends started to diverge almost exactly when the sunlight policy was introduced and years before 1998. This lends support to the claim that the sunlight policy drove a differential growth of high-rises in the south.

4.3 Evidence from FAR

So far we have seen the effects of the sunlight policy on building height. We are more interested in the floor area ratio of buildings, as that determines how much floor space is available for households to consume. Panel A of Figure 5 shows a binned scatterplot of log regulatory floor area ratio of individual residential projects over latitude. There appears to be a strong negative relationship between FAR and latitude - lots in the south are allowed to accommodate more floor space. Even though no controls are included in this exercise, the binned conditional means fall around the fitted negative trend quite closely.[25]

Recall that developer's optimal FAR at a given latitude is given by

$$\text{FAR}^* = \frac{l_0}{h_f} \frac{1}{\frac{l_0}{h^*} + \frac{1}{\tan\alpha}}$$

From the land transactions we observe regulatory FAR upper limit set for individual housing projects. The government transfers use right of land parcels through auctions. If the government's goal is to maximize their auction revenue, they will set regulatory FAR at exactly developers' desired level under the sunlight policy, as that will maximize developers' bids. In this subsection I evaluate whether this expression describes the cross-latitude variation in regulatory FAR well.

By Equation 1, the optimal height h^* depends on the construction cost schedule and housing prices, which we do not have sufficient information about. Therefore in this section I ignore potential developer optimization with respect to building height, and make the approximating assumption that all buildings are of the same size. Width of a building is assumed to be 14 meters, and height is assumed to be 21 meters (around 7 floors).[26]

This approximation only works well if the determinants of optimal building height do not differ systematically across latitudes. To ensure this, I include flexible geographical fixed effects and a rich set of controls in the following estimating equation. The identifying assumption is that within relatively small spatial cells, latitude is not correlated with unobserved determinants of floor area ratio conditional on a rich set of controls.

$$\log \text{FAR}_{pct} = -\beta \log\left(\frac{2}{3} + \frac{1}{\tan\alpha_p}\right) + \lambda_{g(\text{lng}_p, \text{lat}_p)} + \gamma_1' X_p + \gamma_2' X_c + \gamma_3' X_t + \epsilon_{pgc} \qquad (2)$$

If local governments set floor area ratios to the maximum level possible that just suffices to guarantee every resident's right to sunlight, we would expect β to be 1. The extent to which β is lower than 1 measures the extent to which local governments deviate from the sunlight

[25]The Appendix contains scatterplots of regulatory FAR over latitude using the raw data of individual lots along with their non-parametric fits (Figure A3), where we can have an idea of the variation at lot level. The figure also breaks down data into eastern, central and western China. In all regions the negative relationship between regulatory FAR and latitude is present.

[26]These approximations are provided by a Chinese architect.

policy. Deviations can take two forms - they can either set an FAR that is lower than that required by the sunlight policy,[27] or they can allow developers to build more floor space than allowed under the sunlight policy. Only the latter is a violation of the national policy, but both would cause the correlation between FAR and our formula to be lower than one.

Turning to the controls in Equation 2, I include fine-grained longitude-latitude cell fixed effects to control for unobservables. In the main specifications, I use 0.1 degrees longitude by 2 degrees latitude fixed effects, as we want to maintain some variation in latitude within these cells. $\lambda_{g(\text{lng}_p,\text{lat}_p)}$ denote these longitude-latitude cell fixed effects.

X_p is a vector of project-level controls, including whether the lot is greenfield (new land) or brown-field, type of residential building (fixed effects for luxury villas, regular housing, median and low price housing, economy housing,[28] affordable rental housing,[29] or public rental housing[30]), whether the lot is transferred through allotment, contract or auction, fixed effects for land grade, log size of lot, log distance to 2005 city center,[31] and 2005 nighttime lights. If these covariates vary systematically across latitudes and influence either developers' demand for FAR or the government's desired FAR level, omitting them will induce bias in the estimation of β.

X_c is a vector of covariates at the city level in 2006[32], a city being defined as either a prefecture jurisdiction area[33] or a county. The controls include urban population, GDP per capita, population density, employment, share of land with slopes above 15% within a 10-km radius from city center, climatic controls including mean temperature in January, mean temperature in July, and average annual precipitation over the period of 1970-2000.

Finally, X_t is a vector of time covariates including year of transaction and month of year

[27] There are many other potential considerations in setting a regulatory FAR upper limit than complying with the sunlight rule (ventilation, views, green space, unit layout, fire safety, etc.). In practice, the sunlight rule may not be binding in sparsely-populated places. The *Urban planning and management technical specifications* of Gansu province exemplifies this well, a paragraph in which reads: "In the enforcement of this regulation, when facing projects of different scales and localities of different characteristics, the relevant targets are not unique, but should be determined within a range in accordance with other requirements in this document... The vast Gobi and oasis areas to the west of Yellow River are suitable for low-density construction, while middle and eastern parts of our province require construction of relatively high density."

[28] *jingji shiyong fang*. These are government-subsidized housing, where the government grants land to developers at very low prices (the median price of such land is 14.29 yuan per square meter in the sample) and waives related municipal fees, in exchange for their charging a lower price to home buyers.

[29] *lianzu fang*, targeted at households eligible for urban minimum living standards assistance (*dibao*) and facing difficulty with housing.

[30] *gongzu fang*, lottery-based rental housing targeted at households of relatively low income but ineligible for affordable rental housing, migrant workers and people who are newly employed.

[31] City center is again defined as the brightest 1-km cell in terms of 2005 nighttime light. In cases where multiple cells are equally bright, their geometric center is used.

[32] I choose to control for these city level characteristics in 2006 instead of later years to make these covariates more closely resemble predetermined variables that ensure consistent estimation, since I use land transactions in or after 2007 in the regressions.

[33] *shixiaqu*, including districts governed by a prefecture, but excluding counties under the prefecture. This is an appropriate unit of analysis, as a prefecture directly controls only land in its districts, while counties have autonomy in managing their land. The prefecture jurisdiction area roughly corresponds to the urban core of a prefecture.

fixed effects. Year fixed effects control for both business cycles in real estate development and secular trends, while month of year fixed effects control for possible seasonality in the demand for floor space.

Before turning to regression estimates, panel B of Figure 5 shows a binned scatterplot where both log FAR and latitude are residualized within these longitude-latitude cells and over the aforementioned controls. Residualizing within the fine cells reduces the policy-induced variation compared to noise, and therefore leads to more dispersion in the data. However, we still see a clear negative relationship between latitude and FAR.

Table 3 presents the regression results, where I call $-\log\left(\frac{2}{3}+\frac{1}{\tan\alpha_p}\right)$ "predicted log FAR". In column 1 where log regulatory FAR is simply regressed on predicted log FAR with no controls, β is estimated at 0.68, and is highly significant. Unsurprisingly, β is estimated to be less than 1, but the magnitude of 0.68 suggests that the model does a reasonably good job at characterizing mean FAR conditional on latitude. The r-squared of this regression is just 0.055, due to dispersion in FAR at a given latitude. Through the lens of the model, this can be rationalized with variation in optimal height h^* due to different demand conditions and construction costs at different locations within a latitude.

Adding longitude-latitude-cell fixed effects, time fixed effects, lot-specific controls and city economic controls sequentially in columns 2 through 4 changes the estimated magnitude of β only slightly. It is interesting to note that larger lots tend to have smaller FAR upper limits, and the elasticity is estimated at around -0.035. This is likely due to the fact that large residential communities are more likely to be designed for relatively well-off households and therefore feature low-density housing.[34] The FAR upper limit also decreases slowly as one moves from city center to the suburbs, with a 50% increase in distance to city center decreasing FAR upper limit by just 1.5% to 2%. This is consistent with the common observation that the suburbs of Chinese cities are often replete with high-rise buildings.[35] In column 5 where the climatic controls are added, precision decreases a lot due to the high correlation between temperature and latitude, which renders the coefficient insignificant.

How large is the estimated effect of latitude on regulatory FAR under the sunlight policy? Our preferred specification in column 4 indicates that moving across the interquartile range of latitudes, regulatory FAR would increase by 15% all else equal. This suggests that on average, the sunlight policy is a binding constraint on urban density across China.

Table 4 examines whether the sunlight policy is *always* a binding constraint for density. Column 1 shows that projects on larger land parcels deviate signifcantly more from the policy, consistent with the earlier conjecture that larger residential communities are more likely to be middle-class or high-income low-density housing. In these cases, the sunlight requirement is not a binding constraint for density, as density is restricted to be lowered than that allowed

[34]Barr and Cohen (2014) shows a similar negative correlation between plot size and FAR for commercial buildings in New York City.

[35]See, for example, `https://www.wired.com/2016/12/aurelien-marechal-block-china-cities/`.

by the sunlight requirement. Housing projects situated on above-median-sized parcels have an estimated β 0.2 lower than housing projects situated on below-median-sized parcels. Column 2 shows that distance to city center does not significantly change the degree to which the sunlight policy shapes density across space. Projects far away from city center also appear to be constrained by the sunlight policy, consistent with the earlier hypothesis that the suburbs are also places of high-density living. Column 3 shows that FAR in larger cities deviate significantly more from what would be predicted from the sunlight policy. However, this effect disappears when controlling for GDP per capita, as column 4 indicates. Instead, richer cities appear to deviate significantly more from the policy, with housing projects in above-median GDP per capita cities having an estimated β 0.25 lower than housing projects in below-median GDP per capita cities. The main effect of GDP per capita on FAR is significantly negative, suggesting that deviation from the sunlight predictions is largely caused by projects in richer cities being subject to heavier development restrictions, rather than them violating the policy by building denser. This could be caused by richer households demanding more open space, or higher car ownership and better public transit infrastructure in richer cities lowering commuting costs and hence making distance less costly. Column 5 interacts the predicted FAR formula with an indicator for prefecture. We see that overall prefectures deviate from the sunlight policy more than counties, with β around 0.31 lower in prefectures than in counties. Column 6 adds interactions between predicted FAR and the four variables previously considered: parcel size, distance to city center, population, GDP per capita. The coefficient on the interaction between prefecture and predicted FAR drops to -0.21, suggesting that these characteristics account for some, but not all, of the differences between prefectures and counties.

One could make the argument that the observed regulatory FAR upper limits may be higher than developers' desired FAR levels (i.e., not binding). Section B shows that there is a strong positive correlation between regulatory FAR and land price within cities, after a host of project-level covariates are controlled for. This is consistent with the notion that regulatory FARs tend to be binding, and these positive correlations reflect the shadow value of lifting the upper limits.

5 Relationship between Density and the Speed of Information Diffusion

This section starts by showing that consistent with higher residential density in the south induced by the sunlight policy, 1-km^2 grid cells with similar population densities in 1990 have diverged in population density by 2010, with cells in the south becoming significantly denser. The policy also induces differences in land prices across latitudes. The relationship between population density and the speed of information transmission is then studied. An analysis of observed dynamics of user behavior on the Local Leader Message Board reveals that people respond to issues that recently receive replies from the local government by increasing postings about the same issue. An increase in local government's reply rate to a petition topic significantly increases subsequent petitions of the same topic. The increase, however, takes time to materialize. Estimates suggest that the number of petitions reaches its peak three quarters out from the initial reply rate increase before starting to decline. This is consistent with a diffusion process of information about shifts in government attention. The response of postings to government reply is significantly faster in southern cities, where population densities are higher.

5.1 Latitude, Density and Land Prices

Does the sunlight policy induce higher population density and land prices in the south? I estimate the following equation:

$$y_{ic} = -\gamma \log\left(\frac{2}{3} + \frac{1}{\tan \alpha_i}\right) + \lambda_{\text{prov}(c)} + \beta_1' X_i + \beta_2' X_c + \epsilon_{ic}$$

where the unit of observation is a 1-km^2 grid cell. $-\log\left(\frac{2}{3} + \frac{1}{\tan \alpha_i}\right)$ is predicted log FAR, as previously, calculated as a function of solar elevation angle and therefore latitude. $\lambda_{\text{prov}(c)}$ is a province fixed effect. X_i is a vector of cell-level controls including log nighttime lights in 1992 and log population density in 1990. These control for initial densities and levels of economic development before the policy was legally in place. Consequently we estimate differential changes in densities across latitudes since the early 1990s. Controlling for initial density is important as the literature has often found high persistence in population density across time, sometimes over millennia (Davis and Weinstein, 2002; Bleakley and Lin, 2012; Barjamovic et al., 2017). X_i also includes log nighttime lights in 2013 to control for potentially differential economic growth across latitudes. X_c is a vector of city-level controls including log city population in 1997, GDP per capita in 1997[36], log city population in 2012, GDP per capita in 2012 and share of land with above 15% slope within a 10-km radius from city center. Again, this controls for both initial conditions and possibly differential trends in GDP per capita and citywide population across latitudes. Alternatively, we can regard the controls in 2012 as a consequence of the sunlight policy and do not include them in the specifications. Estimates of γ turn out to be quite similar when the 2012 controls are not included.

Since the observations here are high-resolution grid cells, it is important to account for potential spatial autocorrelation in errors. I use a clustering algorithm to find contiguous "urban clusters", and cluster my standard errors at the level of these discovered spatial clusters. Section A.2 describes this procedure in detail. Notably, the discovered urban clusters exhibit close fits to Zipf's Law in both population and land area (see Section A.2). The regressions use grid cells that belong to urban clusters spanning at least 5 km^2. This excludes small patches

[36] Year 1997 falls slightly after the policy was introduced, and is the first year for which I have data for all cities.

of urban area, which may not be considered cities in the usual sense.

Table 5 presents the estimates. The outcome in column 1 is average residential FAR in cell, calculated as a weighted average of project FARs with project parcel size as weights. We see that the estimated coefficient is smaller than that in Table 3, since more weight is given to larger lots, which deviate more from the sunlight policy. The coefficient is however still sizable (above 0.5). Column 2 reports results for population density in 2010. We see that similarly dense and developed cells have diverged in population density in 2010, and cells with higher predicted FAR under the sunlight policy (i.e., lower latitudes) have become significantly denser.

A standard urban economics model with homogeneous households would predict higher land prices where population density is higher. If households spend a fixed share of their income on housing, an assumption standard since Helpman (1998), more people on one square kilometer of land implies more spending on housing within that square kilometer, which accrues to land and results in a higher price for residential land. If land is freely allocated to its most profitable use between residential and commercial uses, i.e., zoning does not distort land uses, we should expect commercial land prices to be higher in the south as well. Even if zoning causes residential and commercial land prices to be different in the same location, denser places provide greater market access to consumers, which is valuable to non-tradable goods producers in a city (e.g., restaurants, retail, etc.) since travel within cities is costly. The higher market access will be capitalized into higher prices for commercial land. Finally, to the extent that commercial floor space is elastically supplied, we should also expect to see higher commercial FARs in denser places in response to the higher commercial rent there.

These predictions are borne out in the data. Columns 3 and 4 examine residential and commercial land prices respectively. To isolate variation in land prices solely due to location, project characteristics are partialed out. In particular, I run a hedonic regression of (log) per square meter real price on cell fixed effects, log parcel size, transaction year and month-of-year dummies, land grade dummies, project type dummies and a greenfield dummy. I then use estimated cell fixed effects to represent land price in each cell. We see that both residential and commercial land are more expensive in the south, with residential land prices more responsive to latitude than commercial land prices. It is worth noting that the share of land with above 15% slope within a 10-kilometer radius from city center is positively correlated with land price, consistent with findings in Saiz (2010) for US cities. A standard deviation increase in the share of surrounding steep slopes is associated with a 15% increase in residential land price and an 11% increase in commercial land price. In column 5, we see that denser places indeed have higher commercial FARs as well. While surrounding slopes increase residential FAR, they do not have a significant impact on commercial FAR.

5.2 Response of Postings to Government Replies on LLMB

To study whether people respond, either due to strategic considerations or psychological encouragements, to the type of petitions that seem to have received recent attention from the local government, I use variation in different topics being discussed on a given city's sub-board on LLMB. In an earlier paper (Jiang et al., 2018), we analyzed the contents of these petitions using LDA topic models (Blei et al., 2003). Here I assign each posting to the topic that makes up the largest share of the petition. I then exclude petitions on rural issues from analysis, since we are interested in posts written by urban residents here.

Section E contains highest-probability words for each topic. We see that the petitions address a wide range of issues from house expropriation, wage arrears and conflicts with property-management companies[37] to pollution, teacher compensation and pyramid schemes. Jiang et al. (2018) provides evidence that most of the postings express personal grievances rather than convey policy suggestions. I further divide the topics into grievance topics and non-grievance topics, where non-grievance topics are those that do not directly address a personal issue. These include comments on policies and compliments to government officials. Local government agencies leave publicly viewable responses to a subset of the postings. 58.5% of postings in the sample receive replies.

It is worth noting that many of the aired grievances are not easily resolved through the legal system, either because the issue in question does not involve legal infringements (e.g., poor teacher pay and unpleasant neighborhood environment), petitioners lack the resources to resort to legal procedures (e.g., wage arrears for migrant workers) or that the justice system may not be reliable for protecting the individual (e.g., when the infringers are local government officials in housing expropriations or powerful property-management companies in property conflicts). Consequently, people seek redress for their grievances on this platform by creating publicity and trying to direct the government's attention towards their specific issue. Since the issues raised are wide-ranging, there is uncertainty around which issues the government will consider seriously. As a result, an increase in government reply rate to a topic can convey information that the government is likely to be paying attention to the issue and that the likelihood that this issue will be redressed becomes higher. We should therefore expect postings on the same issue to increase, and especially so if the topic in question is a grievance topic.

To test whether increased government reply rate to a topic increases subsequent postings on the same topic, I estimate the following equation:

$$\text{Posts}_{ict} = \gamma_{ct} + \delta_{it} + \beta \text{Reply}_{ict-1} + \epsilon_{ict}$$

where Posts_{ict} is log number of postings on topic i in city c's LLMB sub-board in quarter t. γ_{ct} are city-by-quarter fixed effects. These fixed effects allow for city-specific trends in number

[37] The prevalence of this type of conflicts is again caused by the fact that the majority of Chinese urban residents live in large residential compounds that house hundreds or even thousands of families.

of petitions over time. Including them leverages variation generated by local government's differential replies to different topics within a city and a period. In other words, we are looking at whether the government's reply to a certain topic relative to other topics drives up subsequent postings on the topic relative to other topics. This guarantees that the estimated β is not confounded by city-specific trends that drive both government replies and public engagement (e.g., popularity of the platform as an engagement tool). δ_{it} are topic by quarter fixed effects controlling for the general salience of different issues over time.

Table 6 presents the estimates. In column 1, city fixed effects, topic fixed effects and quarter fixed effects are included, but city-by-quarter fixed effects are not. These city-by-quarter fixed effects are included in column 2, where the estimated coefficient on last-quarter reply rate drops substantially. This suggests that the estimate in column 1 is confounded by city-specific trends capturing factors such as the platform's popularity. The estimate in column 2 is significant, and suggests that a 10% increase in reply rate to a topic will increase postings on that topic by 0.37% the next quarter. Column 3 adds topic-by-quarter fixed effects, and the coefficient remains largely unchanged.

In addition to last-quarter reply rate to the given petition topic, column 4 adds the average reply rate to all other topics in the city's sub-board last quarter. We see that this coefficient is not significantly different from 0, suggesting that the increase in postings is due to a perceived shift of government attention towards specific issues, rather than a general increase in government responsiveness. Column 5 adds an indicator for grievance topics and its interaction with last-quarter reply rate to test whether grievance topics and non-grievance topics respond differently to government reply. Grievance postings are found to be significantly more responsive to government replies. In fact, non-grievance postings appear to respond negatively to increases in reply rate, possibly due to people changing their communication tactics from policy suggestions to more outright complaints.

5.3 Dynamic Response to Government Replies

What does the time trajectory of response to increased government replies look like? I estimate the following distributed lag model:

$$\text{Posts}_{ict} = \gamma_c + \delta_{it} + \sum_{j=0}^{J} \beta_j \text{Reply}_{ict-j} + \epsilon_{ict} \tag{3}$$

Estimates for varying lag lengths from 4 to 10 are presented in Table 7. We see that after an increase in reply rate, postings on the same topic gradually increase and reach its peak after three quarters before starting to gradually decline.[38] This process of gradual increase likely reflects the diffusion of information about latest shifts in local government attention.

[38] Volumes of postings on the platform are not too high to obscure replies made a few quarters ago. An average prefecture receives around 250 postings per year during 2008-2015.

Could the observed relationship reflect a reverse causality, i.e., local governments are more likely to reply to issues that are raised more frequently? I further include leads of reply rate in Equation 3. The estimated coefficients adjusted to reflect responses to a 10% increase in government reply rate are plotted in Figure 6 for four leads and varying lengths of lags. We see that none of the leads of reply rate have significant coefficients, indicating that higher volumes of posts on a topic do not anticipate more government replies to the topic. In the three quarters after an increase in reply rate, postings gradually increase. They then start to decrease and return to baseline levels after around eight quarters.

Figure 7 shows the cumulative response of log number of posts to a 10% increase in government reply rate obtained from estimating the following equation:

$$\text{Posts}_{ict} = \gamma_c + \delta_{it} + \sum_{j=0}^{9} \sigma_j \Delta\text{Reply}_{ict-j} + \sigma_{10}\text{Reply}_{ict-10} + \epsilon_{ict}$$

Here an S shape of post increase becomes apparent. According to point estimates, in the eight quarters after a 10% reply rate increase, postings on the same topic would increase by 5.1%.

5.4 Latitude and the Differential Speed of Post Increase

Does density contribute to faster information diffusion? If this is the case, we should expect posts in southern cities to respond faster to changes in government reply. To test this hypothesis, I estimate the following equation:

$$\text{Posts}_{ict} = \gamma_c + \delta_{it} + \sum_{j=0}^{J} \beta_j \text{Reply}_{ict-j} + \sum_{j=0}^{J} \eta_j \text{Reply}_{ict-j} \times \widehat{\log\text{FAR}}_c + \sum_{j=0}^{J} \text{Reply}_{ict-j} \times \sigma_j' \text{X}_c + \epsilon_{ict} \tag{4}$$

$\widehat{\log\text{FAR}}_c$ is "predicted log FAR" as previously defined, and is a function of latitude. In particular, $\widehat{\log\text{FAR}}_c = -\log\left(\frac{2}{3} + \frac{1}{\tan\alpha_c}\right)$, where α_c is the solar elevation angle in city c. The set of η_j's captures how latitude affects the dynamic response path of posts to government reply. To address the concern that latitude could be correlated with other characteristics than density, I also interact three sets of city-level controls with each lag of reply rate.

The first set of controls are economic controls including GDP per capita, government expenditure as a share of GDP, government social security expenditure as a share of GDP, and government low-income assistance expenditure as a share of GDP. The government expenditure measures capture the size of local government as well as weights placed on social welfare.

The second set of controls are relative shares of city residents at six education levels,[39]

[39] These are primary school, middle school, high school, 3-year college, 4-year university and master and above. These shares are tabulated from the 2005 1% census.

Controlling for education levels is important, as people with different education and skill levels might interact with each other differently and/or have different attitudes towards the government.

The third set of controls concern the city's communication infrastructure. These include the following variables normalized by city population: number of post offices, landline users, mobile phone users and broadband internet users. Faster information diffusion could be due to more extensive telecommunication networks instead of face-to-face communication facilitated by high density. I also include city's annual passengers transported and annual freight transported normalized by city population as measures of a city's general openness.

In the most saturated specification, I interact province fixed effects with each of the lags in addition to the controls, so that we only exploit differential responses to government reply within provinces.

Table 8 reports estimates of the η_j's. Figure 8 plots cumulative differential increases. The five panels in Figure 8 correspond to the five specifications in Table 8. Plotted coeffcients are appropriately scaled to reflect cumulative differential increases after a 10% increase in reply rate, between a city at the 75th percentile of latitude and a city at the 25th percentile of latitude. Positive values indicate a larger number of posts in the southern city.

Column 1 of Table 8 shows estimates where no controls are included. We see that in the first three quarters after an increase in government reply, cities in the south (with higher predicted log FAR) experience larger increases in posts of the same topic. In terms of magnitude, the estimates suggest that in the first two quarters after a reply rate increase, a city at the 25th percentile of latitude would experience a cumulative increase in postings that is 2.7 times as large as a city at the 75th percentile of latitude. The gap is however closing over time. After the fourth quarter, cities in the north experience larger increases in posts and therefore gradually catch up. Column 2 adds education controls. Column 3 adds GDP per capita and the government expenditure controls, and column 4 further adds the telecommunication and openness controls. We see that the estimates only change marginally. Column 5 adds the full set of interactions between province fixed effects and reply rate lags. Estimates become larger in magnitude, but the overall pattern remains unchanged - southern cities experience larger increases in posts in the first few quarters, consistent with the hypothesis that higher density contributes to faster information diffusion.

To the extent that we worry about latitude being correlated with unobserved determinants of information transmission conditional on the controls in Table 8, Table 9 reports estimates where log predicted FAR is replaced by log observed FAR[40] as a robustness check. These estimates are potentially biased due to the endogeneity of actual FAR. Reassuringly, we see that higher density is again associated with larger responses of petitions to government attention shifts, although unlike in Table 8, less dense places do not appear to catch up after four quarters.

[40] To be precise, observed FAR here is city-level average of residential FAR upper limit.

Section F contains placebo tests testing for differences across longitudes, where regressions in Table 8 are repeated with log predicted FAR replaced by longitude. Since density does not systematically vary over longitudes, we should expect the speed of information transmission not to vary across longitudes. Table A3 suggests that this is indeed the case.

5.5 Direct Survey Evidence

In this subsection, I provide direct evidence on the relationship between latitude and word of mouth drawing on the China Social Governance Survey conducted by Tsinghua University researchers in 2015 (Zheng et al., 2018), which contains a host of questions on political and cultural attitudes. The survey is conducted to a nationally representative sample of respondents. I focus on the 1,362 urban respondents in my analysis, who live in 78 different cities.

Table 10 reports results on whether the respondent heard gossip or gossiped about news concerning the economy, politics or society in the last month. Columns 1 and 3 include only latitude as an explanatory variable, while columns 2 and 4 add individual-level controls. I control for the individual's overall interest in public issues measured with the individual's frequency of consuming news about politics,[41] whether the respondent is interested in politics[42] and whether the respondent often talks about politics with family and friends.[43] The controls also include gender, education level, *hukou* status (rural or urban), age, and size of city living in.[44]

Overall, 41% of respondents indicate that they heard gossip about public issues in the previous month. A smaller share of 26% indicate that they gossiped about public issues in the previous month. Latitude is negatively associated with the probability of gossip. A five-degree increase in latitude is associated with a 3 percentage points decrease in the probability of hearing gossip and a 3.5 percentage points decrease in gossiping for similar individuals with similar levels of interest in public issues. This lends further support to the notion that higher density in the south is conducive to the spread of word of mouth.

6 Threats to Validity

While the results in Subsection 5.4 suggest that higher densities in southern cities facilitate information diffusion about local government's changes in reply priorities, there could be other reasons why postings in southern cities systematically respond faster to local government's reply rate changes, even within provinces. In this section, I discuss some of the other possible

[41]These are fixed effects for six answers: a few times every day, once every day, a few times every week, once or twice every week, less than once every week, never.

[42]Fixed effects for four answers: completely uninterested, relatively uninterested, relatively interested, very interested.

[43]Fixed effects for three answers: often, occasionally, never.

[44]Fixed effects for population over one million, over 100,000, or under 100,000.

explanations, and conclude that they are not sufficiently convincing.

6.1 Different Attitudes towards the Government

The first alternative explanation for faster responses of postings to changes in government reply rate in southern cities within a province is that southern and northern residents may hold different attitudes towards the government, and will behave differently given the same information about shifts in government attention. To test whether this explanation is plausible, I again draw on the China Social Governance Survey.

Table 11 examines whether within provinces, northern and southern residents hold systematically different views about the government. Panel A concerns whether the respondent thinks the government is "solely responsible" or "primarily responsible" in guaranteeing people education, pension, jobs, healthcare and housing, as opposed to individuals securing these on their own. Specifications in odd-numbered columns include just province fixed effects, while specifications in even-numbered columns add the following individual-level controls: gender, education level, *hukou* status, age and size of city living in. We see that there is no systematic differences across latitudes in people's beliefs about government's role in any of the five areas, whether or not individual characteristics are controlled for. This is not caused by a lack of differences in opinions - among the five areas, as few as 28% of respondents think the government is responsible for providing housing, and as many as 63% of respondents think the government is responsible for providing healthcare.

Panel B shows people's trust in various entities. The outcome is whether the respondent has "complete trust" or "substantial trust" in, respectively, courts, the central government, National People's Congress, government officials, the army, police, the local government, newspapers, radio and TV programs, and social organizations. All columns include province fixed effects and individual-level controls. We see that there is no systematic difference across latitudes within provinces in terms of trust in the government institutions, although northern residents do seem slightly less likely to trust the army, and more likely to trust social organizations. Again, the lack of significance is not due to uniform responses. The share of people who report trust ranges from 18% for government officials to 69% for the army.

6.2 Online Information Diffusion

Another alternative explanation for faster responses in the south to changing government priorities is that the information may be transmitted online instead of through face-to-face communication. While it seems hard to believe that people would rely solely on sending this information online without talking about it face to face, and that the regressions in Table 8 have included a vector of local telecommunication controls, I present evidence from the China Social Governance Survey that there are no systematic differences in internet use behavior across latitudes, so that to the extent that information travels faster in southern cities, it is likely due to offline interactions.

Table 12 reports the within-province relationship between individual responses to internet-related questions and the latitude of their cities. Individual characteristics are controlled for in all specifications. Column 1 concerns whether the individual can access internet through cell phone. Column 2 concerns whether the individual has other means of access to internet (broadband, optical fiber, dial-up, etc.). Column 3 is on whether the respondent uses internet for at least 30 minutes per day, and column 4 is on whether the respondent obtains political information online at least once every week. Column 5 looks at whether the respondent ever expressed political opinions online, and column 6 looks at whether the respondent uses social media. In all cases, there is no significant difference across latitudes.

7 Conclusion

Does urban density facilitate information and knowledge diffusion, a conjecture dating back to Marshall (1890)? Using plausibly exogenous variation induced by the sunlight policy in China, I present evidence indicating that density does play a role in accelerating the transmission of information as evidenced by the response speed of online postings to local government's shifting attention. On a practical level, I find that a policy exogenously shifts urban density across a country with the world's largest urban population. The relationship between latitude and residential floor area ratio is clear and robust, and can be described quite well by a formula derived from developer profit maximization. This provides a useful source of variation that can be leveraged in exploring a wide range of questions regarding housing affordability, commuting costs, and urban sprawl. As the developing world undergoes rapid urbanization and city building, the way streets and buildings are laid out today could affect urban density and productivity for decades or more to come. It is therefore important to understand how urban planning practices could affect market outcomes and economic efficiency.

This paper also raises interesting questions about the governance of cities. It is not clear *a priori* whether more desirable social outcomes can be achieved if citizens are able to immediately react to policy changes. Word of mouth facilitated by urban density can either open up opportunities for designing effective policies or undermine the effectiveness of policies that are not strategy-proof. The information set of citizens mediated by density should therefore be considered in the policy-making process. On the other hand, the flow of information aided by urban density could also increase government accountability, as is suggested in Campante and Do (2014). More work is required to understand whether better knowledge of government behavior facilitated by urban density leads to increased accountability and improved governance.

If information does flow faster in denser cities, a natural question is what types of information are more sensitive to distance, and therefore benefit more from density. It is conceivable that information about rapidly changing situations - "knowledge of circumstances of the fleeting moment" in the words of Hayek (1945), or information that is tacit and hard to codify[45] will be particularly hard to transmit over distance. Looking into these finer distinctions will enrich our understanding of the productivity benefits of cities and inform better policies. I leave this exciting topic to future studies.

[45] Juhász and Steinwender (2018) shows how the trade volume of cotton textile products with varying degrees of specification codifiability reacted differently to being connected to the global telegraph network in the 19th century.

References

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015): "The economics of density: Evidence from the Berlin Wall," *Econometrica*, 83, 2127–2189.

Almond, D., Y. Chen, M. Greenstone, and H. Li (2009): "Winter heating or clean air? Unintended impacts of China's Huai River policy," *American Economic Review: Papers & Proceedings*, 99, 184–190.

Arzaghi, M. and J. V. Henderson (2008): "Networking off Madison Avenue," *Quarterly Journal of Economics*, 75, 1011–1038.

Asami, Y., X. Gao, and W. Katsumata (2006): "Evaluating land-use restrictions concerning the floor area ratio of lots," *Environment and Planning C: Government and Policy*, 24, 515–532.

Audretsch, D. and M. Feldman (1996): "R&D spillovers and the geography of innovation and production," *American Economic Review*, 86, 630–640.

Barjamovic, G., T. Chaney, K. Coşar, and A. Hortaçsu (2017): "Trade, Merchants, and the Lost Cities of the Bronze Age," Tech. rep.

Barr, J. and J. P. Cohen (2014): "The floor area ratio gradient: New York City, 1890-2009," *Regional Science and Urban Economics*, 48, 110 – 119.

Bertaud, A. (2018): *Order Without Design: How Markets Shape Cities*, MIT Press.

Bertaud, A. and J. K. Brueckner (2004): "Analyzing building height restrictions - predicted impacts, welfare costs, and a case study of Bangalore, India," Tech. rep., The World Bank.

Bleakley, H. and J. Lin (2012): "Portage and Path Dependence ∗,″ *The Quarterly Journal of Economics*, 127, 587 – –644.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003): "Latent Dirichlet Allocation," *Journal of machine Learning research*, 3, 993–1022.

Brockett, D., D. Fridley, J. Lin, and J. Lin (2002): "A tale of five cities: the China residential energy consumption survey," *ACEEE Summer Study on Building Energy Efficiency*, 29–40.

Brueckner, J. K., S. Fu, Y. Gu, and J. Zhang (2017): "Measuring the stringency of land use regulation: The case of China's building height limits," *Review of Economics and Statistics*, 99, 663–677.

Brueckner, J. K. and S. V. Lall (2015): "Cities in developing countries: fueled by rural–urban migration, lacking in tenure security, and short of affordable housing," in *Handbook of regional and urban economics*, Elsevier, vol. 5, 1399–1455.

Campante, F. R. and Q.-A. Do (2014): "Isolated capital cities, accountability, and corruption: Evidence from US states," *American Economic Review*, 104, 2456–2481.

Campante, F. R., Q.-A. Do, and B. Guimaraes (2014): "Capital cities, conflict, and misgovernance: Theory and evidence," Sciences Po Economics Discussion Papers 2014-13, Sciences Po Departement of Economics.

Chauvin, J. P., E. Glaeser, Y. Ma, and K. Tobio (2017): "What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States," *Journal of Urban Economics*, 98, 17–49.

Chen, T. and J.-S. Kung (2016): "Do land revenue windfalls create a political resource curse? Evidence from China," *Journal of Development Economics*, 123, 86 – 106.

Combes, P.-P., G. Duranton, L. Gobillon, and S. Roux (2010): "Estimating agglomeration economies with history, geology, and worker effects," in *Agglomeration Economies*, ed. by E. L. Glaeser, Chicago University Press, Chicago.

Cui, H., X. Li, and W. Zhang (2005): "China's inter-census survey in 2005," *Beijing: National Bureau of Statistics of China*.

Davis, D. R. and D. E. Weinstein (2002): "Bones, bombs, and break points: the geography of economic activity," *American Economic Review*, 92, 1269–1289.

Dingel, J. I., N. A. Miscio, and D. R. Davis (2018): "Cities, lights, and skills in developing economies," .

Donaldson, D. and R. Hornbeck (2016): "Railroads and American economic growth: A 'market access' approach," *The Quarterly Journal of Economics*, 131, 799–858.

Ester, M., H.-P. Kriegel, J. Sander, and X. Xu (1996): "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, Portland, OR, AAAI Press, vol. 96, 226–231.

Fang, H., Q. Gu, W. Xiong, and L.-A. Zhou (2016): "Demystifying the Chinese housing boom," *NBER macroeconomics annual*, 30, 105–166.

Fick, S. E. and R. J. Hijmans (2017): "WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas," *International Journal of Climatology*, 37, 4302–4315.

Fischer, G., F. Nachtergaele, S. Prieler, H. van Velthuizen, L. Verelst, and D. Wiberg (2008): "Global agro-ecological zones assessment for agriculture (GAEZ 2008)," *IIASA, Laxenburg, Austria and FAO, Rome, Italy.*

Gabaix, X. and R. Ibragimov (2011): "Rank- 1/2: a simple way to improve the OLS estimation of tail exponents," *Journal of Business & Economic Statistics*, 29, 24–39.

Gaughan, A. E., F. R. Stevens, Z. Huang, J. J. Nieves, A. Sorichetta, S. Lai, X. Ye, C. Linard, G. M. Hornby, S. I. Hay, H. Yu, and A. J. Tatem (2016): "Spatiotemporal patterns of population in mainland China, 1990 to 2010," *Scientific Data*, 3, 160005.

Glaeser, E. L. and J. D. Gottlieb (2009): "The wealth of cities: Agglomeration economies and spatial equilibrium in the United States," *Journal of Economic Literature*, 47, 983–1028.

Glaeser, E. L., J. Gyourko, and R. Saks (2005): "Why Is Manhattan so expensive? Regulation and the rise in housing prices," *The Journal of Law and Economics*, 48, 331–369.

Glaeser, E. L. and D. C. Mare (2001): "Cities and skills," *Journal of labor economics*, 19, 316–342.

Gyourko, J. and R. Molloy (2014): "Regulation and housing supply," *NBER working paper.*

Harari, M. (2017): "Cities in bad shape: Urban geometry in India," *mimeo.*

Hayek, F. A. (1945): "The se of Knowledge in Society," *American Economic Review*, 35, 519–530.

Helpman, E. (1998): *Topics in Public Economics: Theoretical and Applied Analysis*, Cambridge, UK: Cambridge Univ. Press, chap. The size of regions, 33–54.

Iyer, L., X. Meng, N. Qian, and X. Zhao (2013): "Economic transition and private-sector labor demand: Evidence from urban China," *NBER working paper.*

Jacobs, J. (1969): "The economy of cities," *New York: Vintage.*

Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993): "Geographic localization of knowledge spillovers as evidenced by patent citations," *the Quarterly journal of Economics*, 108, 577–598.

Jiang, J., T. Meng, and Q. Zhang (2018): "From internet to social safety net: The policy consequences of online participation in China," .

Juhász, R. and C. Steinwender (2018): "Spinning the web: Codifiability, information frictions and trade," *NBER working paper*.

Liu, X. and Y. Long (2016): "Automated identification and characterization of parcels with OpenStreetMap and points of interest," *Environment and Planning B: Planning and Design*, 43, 341–360.

Long, Y. (2016): "Redefining Chinese city system with emerging new data," *Applied Geography*, 75, 36–48.

Marshall, A. (1890): *Principles of Economics*, London: Macmillan.

Moretti, E. (2004): "Human capital externalities in cities," in *Handbook of regional and urban economics*, Elsevier.

Rauch, J. (1993): "Productivity gains from geographic concentration of human capital: Evidence from the cities," *Journal of Urban Economics*, 34, 380–400.

Rosenthal, S. S. and W. C. Strange (2004): "Evidence on the nature and sources of agglomeration economies," in *Handbook of regional and urban economics*, Elsevier, vol. 4, 2119–2171.

——— (2008): "The attenuation of human capital spillovers," *Journal of Urban Economics*.

Rozenfeld, H. D., D. Rybski, X. Gabaix, and H. A. Makse (2011): "The area and population of cities: New insights from a different perspective on cities," *American Economic Review*, 101, 2205–25.

Saiz, A. (2010): "The geographic determinants of housing supply," *The Quarterly Journal of Economics*, 125, 1253–1296.

Turner, M. A., A. Haughwout, and W. Van Der Klaauw (2014): "Land use regulation and welfare," *Econometrica*, 82, 1341–1403.

Vogel, K. B., R. Goldblatt, G. H. Hanson, and A. K. Khandelwal (2018): "Detecting urban markets with satellite imagery: An application to India," *NBER working paper*.

Wang, S.-Y. (2011): "State misallocation and housing prices: Theory and evidence from China," *American Economic Review*, 101, 2081–2107.

Wright, J., ed. (1971): *Industrialized building in the Soviet Union*, National Bureau of Standards Special Publication 334.

Xu, M., C. He, Z. Liu, and Y. Dou (2016): "How did urban land expand in China between 1992 and 2015? A multi-scale landscape analysis," *PloS one*, 11, e0154839.

Zhang, H. (2017): "Sunlight distance restriction, population density and the growth of Chinese cities (in Chinese)," *China Economic Quarterly*, 17, 333–354.

Zheng, Z., Y. Su, and Y. Zhang (2018): "The social origins of public support in China and its changes: evidence from the 2015 China social governance survey (in Chinese)," *Political Science Research*, 03, 51–61.

Zhou, Y. and Y. Long (2016): "SinoGrids: A practice for open urban data in China," *Cartography and Geographic Information Science*, 43, 379–392.

Table 1: Effects of the Sunlight Policy on Building Height

	(1) High-rise	(2) High-rise	(3) High-rise	(4) High-rise	(5) High-rise	(6) High-rise
Latitude	-0.001	0.001	-0.003	-0.003		-0.004
	(0.005)	(0.004)	(0.004)	(0.004)		(0.004)
Post	0.331***	0.275***	0.271***	0.262***	0.266***	0.266
	(0.058)	(0.050)	(0.045)	(0.046)	(0.045)	(0.251)
Latitude × Post	-0.004**	-0.003*	-0.005***	-0.004***	-0.005***	-0.004
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)
Slope within 10-km city radius		0.237***	0.201***	0.200***		0.197***
		(0.035)	(0.032)	(0.032)		(0.032)
Province FEs	Yes	Yes	Yes	Yes	Yes	Yes
City controls	No	Yes	Yes	Yes	No	Yes
Residence controls	No	No	Yes	Yes	Yes	Yes
Household controls	No	No	No	Yes	Yes	Yes
City FEs	No	No	No	No	Yes	No
Province trends	No	No	No	No	No	Yes
Observations	248365	240405	240091	239818	247753	239818
R-Squared	0.135	0.166	0.276	0.287	0.332	0.298
RMSE	0.371	0.364	0.339	0.337	0.326	0.334

Notes: Dependent variable is indicator that the household lives in a building with at least seven floors. Post indicates that the building was constructed in or after 1993, when the sunlight policy was written into law. City controls include 2005 levels of urban population, population density, employment, GDP per capita and share of land with above 15% slopes within a 10-km radius from city center. Residence controls include use porpose of residence (just for living or for both living and production), log number of rooms, log square footage, whether shared with other households, access to tap water, presence of kitchen, toilet and bath facilities, type of fuel and how the residence was obtained (rent or bought, SOE housing, affordable housing or commercial housing). Household controls contain log number of household members and the following characteristics of household head: gender, age, ethnicity, *hukou* registration place, type of *hukou* (rural or urban), literacy, education, whether worked for more than one hour last week, log income, source of income and marital status. Data source: 2005 census random sample. Standard errors are clustered at city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Parallel Economic Growth across Latitudes

	(1) Urban Population	(2) Employment	(3) GDP	(4) GDP pc	(5) Urban Population	(6) Employment	(7) GDP	(8) GDP pc
Year=1998 × Latitude	0.002		0.002**	0.001	0.001		-0.004	-0.005
	(0.002)		(0.001)	(0.001)	(0.003)		(0.003)	(0.003)
Year=1999 × Latitude	0.002		0.001	0.001	-0.001		-0.008	-0.006
	(0.003)		(0.002)	(0.002)	(0.005)		(0.008)	(0.007)
Year=2000 × Latitude	0.000		0.002	0.002	-0.000		-0.007	-0.005
	(0.002)		(0.002)	(0.002)	(0.007)		(0.009)	(0.009)
Year=2001 × Latitude	0.001	-0.002	0.002	0.003	0.005	-0.003	-0.008	-0.007
	(0.002)	(0.004)	(0.002)	(0.002)	(0.007)	(0.004)	(0.011)	(0.010)
Year=2002 × Latitude	-0.001	-0.003	0.004	0.006**	-0.000	0.000	-0.007	-0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.007)	(0.005)	(0.011)	(0.011)
Year=2003 × Latitude	0.004	-0.003	0.006	0.008**	-0.008	-0.003	-0.011	-0.010
	(0.005)	(0.003)	(0.004)	(0.004)	(0.009)	(0.006)	(0.013)	(0.013)
Year=2004 × Latitude	0.004	-0.003	0.008*	0.010**	-0.009	-0.004	-0.008	-0.007
	(0.005)	(0.003)	(0.004)	(0.004)	(0.009)	(0.008)	(0.015)	(0.016)
Year=2005 × Latitude	0.003	-0.006	0.009	0.012**	-0.013	-0.003	-0.005	-0.004
	(0.005)	(0.004)	(0.006)	(0.006)	(0.010)	(0.007)	(0.016)	(0.016)
City FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	No	No	No	No
Province-Year FEs	No	No	No	No	Yes	Yes	Yes	Yes
Observations	17289	12504	17964	17244	17289	12504	17964	17244
R-Squared	0.957	0.951	0.988	0.971	0.961	0.953	0.991	0.977
RMSE	0.218	0.210	0.146	0.140	0.208	0.204	0.131	0.125

Notes: City fixed effects are included in all specifications. Columns 1 to 4 include year fixed effects, while columns 5 to 8 include province by year fixed effects. Outcome variables are in logs. Each regression uses a panel that is balanced on the outcome variable. For urban population, GDP and GDP per capita, the omitted year is 1997. For employment, the omitted year is 2000. Employment data before 2000 are not available. Standard errors are clustered at province level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Predictive Power of the Sunlight Policy

	DV: log FAR for Residential Project				
	(1)	(2)	(3)	(4)	(5)
Predicted log FAR	0.679***	0.628*	0.682**	0.682**	0.486
	(0.055)	(0.329)	(0.320)	(0.328)	(0.406)
Log size		-0.031***	-0.035***	-0.037***	-0.037***
		(0.003)	(0.003)	(0.003)	(0.003)
Log distance to center		-0.029***	-0.026***	-0.039***	-0.039***
		(0.004)	(0.004)	(0.004)	(0.004)
Lng-lat-cell FEs	No	Yes	Yes	Yes	Yes
Time FEs	No	No	Yes	Yes	Yes
Lot-specific controls	No	Yes	Yes	Yes	Yes
City controls	No	No	No	Yes	Yes
Climatic controls	No	No	No	No	Yes
Observations	228428	197183	197183	188402	188402
R-Squared	0.055	0.346	0.378	0.378	0.378
RMSE	0.445	0.369	0.360	0.360	0.360

Notes: These regressions concern residential development projects across China for which the land parcel is transfered between 2007 and 2016. Predicted log Floor Area Ratio is calculated as a function of latitude using implications of a simple model. See text for details. Lng-lat cells are 0.1 degrees longitude by 2 degrees latitude spatial units. Time FEs include year and month-of-year fixed effects. Lot-specific controls include whether the lot is greenfield or brown-field, type of residential building, whether the lot is transferred through allotment, contract or auction, fixed effects for land grade, log lot size, log distance to 2005 city center and 2005 nighttime lights. City controls include 2006 levels of urban population, GDP per capita, population density, employment, share of land with slopes above 15% within a 10-km radius from city center. Climatic controls include mean temperature in January, mean temperature in July, and average annual precipitation over the period of 1970-2000. Standard errors are clustered at city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Is the Sunlight Policy Always a Binding Constraint on Density?

	DV: log FAR for Residential Project					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted log FAR	0.886***	0.850**	0.933***	0.997***	0.848***	0.929***
	(0.329)	(0.331)	(0.334)	(0.330)	(0.318)	(0.325)
Parcel size, above median	-0.229***	-0.228***	-0.224***	-0.216***	-0.080***	-0.210***
	(0.035)	(0.035)	(0.034)	(0.034)	(0.006)	(0.033)
Distance to city center, above median	-0.039***	-0.004	-0.002	-0.003	-0.052***	-0.002
	(0.007)	(0.031)	(0.031)	(0.031)	(0.007)	(0.031)
Population, above median	0.037***	0.037***	-0.066	-0.020	0.015	0.025
	(0.013)	(0.013)	(0.057)	(0.059)	(0.014)	(0.064)
GDP pc, above median	-0.020	-0.020	-0.020	-0.204***	-0.026*	-0.163**
	(0.014)	(0.014)	(0.014)	(0.065)	(0.014)	(0.065)
Predicted log FAR × Parcel size, above median	-0.201***	-0.200***	-0.195***	-0.185***		-0.175***
	(0.042)	(0.042)	(0.041)	(0.041)		(0.040)
Predicted log FAR × Distance to city center, above median		0.049	0.053	0.052		0.071*
		(0.040)	(0.040)	(0.039)		(0.039)
Predicted log FAR × Population, above median			-0.141**	-0.076		0.014
			(0.072)	(0.074)		(0.080)
Predicted log FAR × GDP pc, above median				-0.254***		-0.189**
				(0.077)		(0.077)
Prefecture					-0.014	0.075
					(0.097)	(0.101)
Predicted log FAR × Prefecture					-0.309***	-0.206**
					(0.089)	(0.095)
Lng-lat-cell FEs	Yes	Yes	Yes	Yes	Yes	Yes
Time FEs	Yes	Yes	Yes	Yes	Yes	Yes
Lot-specific controls	Yes	Yes	Yes	Yes	Yes	Yes
City controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	188402	188402	188402	188402	188402	188402
R-Squared	0.368	0.368	0.368	0.368	0.368	0.369
RMSE	0.363	0.363	0.363	0.363	0.363	0.363

Notes: These regressions concern residential development projects across China for which the land parcel is transfered between 2007 and 2016. Predicted log Floor Area Ratio is calculated as a function of latitude using implications of a simple model. See text for details. Lng-lat cells are 0.1 degrees longitude by 2 degrees latitude spatial units. Time FEs include year and month-of-year fixed effects. Lot-specific controls include whether the lot is greenfield or brown-field, type of residential building, whether the lot is transferred through allotment, contract or auction, fixed effects for land grade, and 2005 nighttime lights. City controls include 2006 levels of population density and employment and share of land with slopes above 15% within a 10-km radius from city center. Standard errors are clustered at city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Latitude, Density and Prices

	(1) Resi. FAR	(2) Population	(3) Resi. Land Price	(4) Comm. Land Price	(5) Comm. FAR
Predicted log FAR	0.545***	0.772***	2.029***	1.745***	0.578**
	(0.163)	(0.202)	(0.467)	(0.470)	(0.272)
Log nighttime lights 1992	0.003	-0.069***	-0.057***	-0.004	0.021**
	(0.007)	(0.014)	(0.018)	(0.018)	(0.010)
Log nighttime lights 2013	0.072***	0.259***	0.476***	0.431***	0.174***
	(0.008)	(0.008)	(0.025)	(0.024)	(0.015)
Log population density 1990	0.093***	0.795***	0.188***	0.161***	0.101***
	(0.007)	(0.025)	(0.027)	(0.027)	(0.014)
Log city population 1997	0.033***	0.006	0.111***	0.073*	0.016
	(0.012)	(0.023)	(0.039)	(0.040)	(0.023)
Log city population 2012	-0.014	0.050***	0.062*	0.122***	-0.012
	(0.009)	(0.018)	(0.032)	(0.030)	(0.017)
GDP pc 1997	-0.021	0.152***	0.107**	0.052	-0.001
	(0.016)	(0.049)	(0.045)	(0.042)	(0.027)
GDP pc 2012	-0.046***	-0.011	0.020	0.040	-0.052**
	(0.012)	(0.025)	(0.036)	(0.032)	(0.021)
Slope within 10-km city radius	0.154***	-0.078	0.499***	0.352***	-0.043
	(0.035)	(0.049)	(0.077)	(0.089)	(0.051)
Province FEs	Yes	Yes	Yes	Yes	Yes
Observations	25439	234713	14104	11756	17810
R-Squared	0.225	0.790	0.326	0.354	0.089
RMSE	0.333	0.407	0.837	0.876	0.627

Notes: Units of observation in these regressions are 1 km^2 grid cells. Dependent variable in column 1 is log average residential FAR in the cell, calculated as a weighted average of residential project FARs with project parcel size as weights. Dependent variable in column 2 is log 2010 population density in the cell, obtained from Worldpop. Dependent variable in column 3 is residualized log residential land price in the cell, which is an estimated cell fixed effect from a hedonic regression of log per square meter real price on cell fixed effects, log parcel size, transaction year and month-of-year dummies, land grade dummies, project type dummies and a greenfield dummy. Dependent variable in column 4 is residualized log commercial land price in the cell obtained from a hedonic regression following the same procedure. Dependent variable in column 5 is log average commercial FAR in the cell, calculated as a weighted average of commercial project FARs with project parcel size as weights. Predicted log Floor Area Ratio is calculated as a function of latitude using implications of a simple model (see text for details). Standard errors are clustered at city-cluster level (see Appendix for details). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Reaction of Petition Posting to Government Replies

	Dependent Variable: Log Number of Posts				
	(1)	(2)	(3)	(4)	(5)
Reply $rate_{t-1}$	0.118***	0.037***	0.040***	0.042***	-0.030**
	(0.018)	(0.010)	(0.010)	(0.012)	(0.015)
Reply rate, other $topics_{t-1}$				0.022	
				(0.060)	
Grievance					-0.123***
					(0.041)
Grievance × Reply $rate_{t-1}$					0.092***
					(0.018)
City FEs	Yes	No	No	No	No
Topic FEs	Yes	Yes	No	No	Yes
Quarter FEs	Yes	No	No	No	No
Topic-Quarter FEs	No	No	Yes	Yes	No
City-Quarter FEs	No	Yes	Yes	Yes	Yes
Observations	85469	84875	84875	84875	84875
R-Squared	0.543	0.630	0.654	0.654	0.630
RMSE	0.631	0.595	0.578	0.578	0.595

Notes: Units of observation in these regressions are city-quarter-topic tuples. Topics are obtained from a 30-topic LDA model. Each post is assigned to its highest topic. Dependent variable is log number of posts about a given topic posted in a given city's Local Leader Message Board in a given quarter. Reply $rate_{t-1}$ is local government's reply rate to posts of the same topic in the previous quarter. Reply rate, other $topics_{t-1}$ is local government's average reply rate to posts of the other topics in the previous quarter. Grievance topics are described in the Appendix. Standard errors are clustered at city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Dynamic Reaction of Petition Posting to Government Replies

	Dependent Variable: Log Number of Posts			
	(1)	(2)	(3)	(4)
Reply rate	0.032	0.044	0.035	-0.003
	(0.025)	(0.029)	(0.037)	(0.047)
Reply rate$_{t-1}$	0.055***	0.053**	0.065**	0.062*
	(0.020)	(0.024)	(0.027)	(0.034)
Reply rate$_{t-2}$	0.082***	0.059**	0.063**	0.064*
	(0.020)	(0.025)	(0.030)	(0.038)
Reply rate$_{t-3}$	0.094***	0.083***	0.076***	0.096***
	(0.021)	(0.020)	(0.024)	(0.024)
Reply rate$_{t-4}$	0.086***	0.071***	0.065***	0.085***
	(0.024)	(0.023)	(0.023)	(0.027)
Reply rate$_{t-5}$		0.077***	0.074***	0.065***
		(0.026)	(0.025)	(0.025)
Reply rate$_{t-6}$		0.073***	0.069***	0.072***
		(0.025)	(0.022)	(0.026)
Reply rate$_{t-7}$			0.051	0.056*
			(0.032)	(0.032)
Reply rate$_{t-8}$			0.026	0.009
			(0.030)	(0.028)
Reply rate$_{t-9}$				0.027
				(0.033)
Reply rate$_{t-10}$				-0.009
				(0.032)
City FEs	Yes	Yes	Yes	Yes
Topic-Quarter FEs	Yes	Yes	Yes	Yes
Observations	26782	19579	14858	11505
R-Squared	0.598	0.622	0.637	0.651
RMSE	0.565	0.564	0.563	0.560

Notes: Units of observation in these regressions are city-quarter-topic tuples. Topics are obtained from a 30-topic LDA model. Each post is assigned to its highest topic. Dependent variable is log per capita number of posts about a given topic posted in a given city's Local Leader Message Board in a given quarter. Standard errors are clustered at city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Differential Speed of Petition Reaction to Government Replies

	Dependent Variable: Log Number of Posts				
	(1)	(2)	(3)	(4)	(5)
Reply rate × Predicted log FAR	0.201	0.208	0.117	0.218	2.129***
	(0.153)	(0.167)	(0.184)	(0.222)	(0.719)
Reply rate$_{t-1}$ × Predicted log FAR	0.355***	0.376**	0.303*	0.515***	0.879
	(0.131)	(0.150)	(0.159)	(0.186)	(0.678)
Reply rate$_{t-2}$ × Predicted log FAR	0.148	0.057	0.086	0.153	0.844
	(0.135)	(0.154)	(0.165)	(0.193)	(0.717)
Reply rate$_{t-3}$ × Predicted log FAR	0.057	0.123	0.166	0.272	0.710
	(0.131)	(0.154)	(0.163)	(0.189)	(0.670)
Reply rate$_{t-4}$ × Predicted log FAR	-0.305**	-0.416***	-0.277*	-0.361*	0.713
	(0.130)	(0.148)	(0.161)	(0.185)	(0.671)
Reply rate$_{t-5}$ × Predicted log FAR	-0.231*	-0.287*	-0.286*	-0.232	1.121*
	(0.130)	(0.148)	(0.157)	(0.178)	(0.634)
Reply rate$_{t-6}$ × Predicted log FAR	-0.304**	-0.434***	-0.359**	-0.203	-0.995
	(0.134)	(0.154)	(0.165)	(0.185)	(0.663)
City FEs	Yes	Yes	Yes	Yes	Yes
Topic-Quarter FEs	Yes	Yes	Yes	Yes	Yes
Education interactions	No	Yes	Yes	Yes	Yes
Economic interactions	No	No	Yes	Yes	Yes
Communication interactions	No	No	No	Yes	Yes
Province FE interactions	No	No	No	No	Yes
Observations	19317	19252	18538	18176	18176
R-Squared	0.623	0.626	0.613	0.618	0.627
RMSE	0.565	0.564	0.566	0.566	0.562

Notes: Units of observation in these regressions are city-quarter-topic tuples. Topics are obtained from a 30-topic LDA model. Each post is assigned to its highest topic. Dependent variable is log number of posts about a given topic posted in a given city's Local Leader Message Board in a given quarter. Predicted log Floor Area Ratio is calculated as a function of latitude using implications of a simple model (see text for details). Coefficients of reply rate and its lags are omitted. Education interactions are reply rate and its lags interacted with shares of six education levels of city residents. Economic interactions are reply rate and its lags interacted with city's GDP per capita, fiscal expenditure to GDP ratio, social security spending to GDP ratio and social assistance to GDP ratio. Communication interactions are reply rate and its lags interacted with the following city-level variables normalized by population: annual passengers transported, annual freight transported, number of post offices, landline users, mobile phone users, broadband internet users. Province FE interactions are reply rate and its lags interacted with province fixed effects. Standard errors are clustered at city-topic level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Differential Speed of Petition Reaction to Government Replies Using Observed FAR

	Dependent Variable: Log Number of Posts				
	(1)	(2)	(3)	(4)	(5)
Reply rate × log FAR	0.126**	0.145***	0.171***	0.204***	0.224**
	(0.054)	(0.054)	(0.059)	(0.066)	(0.102)
Reply rate$_{t-1}$ × log FAR	0.013	0.022	0.052	0.077	0.090
	(0.047)	(0.048)	(0.054)	(0.058)	(0.089)
Reply rate$_{t-2}$ × log FAR	0.063	0.075	0.052	0.088	0.206**
	(0.046)	(0.048)	(0.055)	(0.061)	(0.087)
Reply rate$_{t-3}$ × log FAR	0.071	0.127***	0.103**	0.171***	0.190**
	(0.044)	(0.046)	(0.052)	(0.058)	(0.088)
Reply rate$_{t-4}$ × log FAR	0.015	0.064	0.023	0.043	0.064
	(0.048)	(0.050)	(0.054)	(0.058)	(0.085)
Reply rate$_{t-5}$ × log FAR	-0.043	-0.019	-0.042	-0.001	-0.046
	(0.049)	(0.049)	(0.053)	(0.059)	(0.081)
Reply rate$_{t-6}$ × log FAR	0.060	0.087*	0.040	0.121**	0.091
	(0.047)	(0.047)	(0.050)	(0.056)	(0.080)
City FEs	Yes	Yes	Yes	Yes	Yes
Topic-Quarter FEs	Yes	Yes	Yes	Yes	Yes
Education interactions	No	Yes	Yes	Yes	Yes
Economic interactions	No	No	Yes	Yes	Yes
Communication interactions	No	No	No	Yes	Yes
Province FE interactions	No	No	No	No	Yes
Observations	19211	19146	18432	18070	18070
R-Squared	0.622	0.625	0.613	0.618	0.626
RMSE	0.565	0.564	0.566	0.566	0.562

Notes: Units of observation in these regressions are city-quarter-topic tuples. Topics are obtained from a 30-topic LDA model. Each post is assigned to its highest topic. Dependent variable is log number of posts about a given topic posted in a given city's Local Leader Message Board in a given quarter. Log Floor Area Ratio is log of city-level average residential FAR. Coefficients of reply rate and its lags are omitted. Education interactions are reply rate and its lags interacted with shares of six education levels of city residents. Economic interactions are reply rate and its lags interacted with city's GDP per capita, fiscal expenditure to GDP ratio, social security spending to GDP ratio and social assistance to GDP ratio. Communication interactions are reply rate and its lags interacted with the following city-level variables normalized by population: annual passengers transported, annual freight transported, number of post offices, landline users, mobile phone users, broadband internet users. Province FE interactions are reply rate and its lags interacted with province fixed effects. Standard errors are clustered at city-topic level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Survey Evidence on Word of Mouth

	Experience in the last month			
	(1) Heard Gossip	(2) Heard Gossip	(3) Gossiped	(4) Gossiped
Latitude	-0.005	-0.006**	-0.006**	-0.007***
	(0.003)	(0.002)	(0.003)	(0.002)
Individual controls	No	Yes	No	Yes
Mean of DV	0.410	0.411	0.263	0.266
Observations	1362	1324	1362	1324
R-Squared	0.003	0.187	0.005	0.191
RMSE	0.491	0.448	0.439	0.401

Notes: Dependent variable is whether the respondent heard gossip or gossiped about news concerning the economy, politics or society in the last month. Individual controls include respondent's frequency of consuming news about politics, whether the respondent is interested in politics, whether the respondent often talks about politics with family and friends, gender, education level, *hukou* status, age, and size of city living in. Data are for urban respondents in the 2015 China Social Governance Survey. Standard errors are clustered at city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: Attitudes towards Government not Systematically Different Across Latitudes

	Government should be responsible for									
	(1) Education	(2) Education	(3) Pension	(4) Pension	(5) Jobs	(6) Jobs	(7) Healthcare	(8) Healthcare	(9) Housing	(10) Housing
Latitude	0.009	0.009	-0.001	0.002	-0.002	0.005	-0.021	-0.022	-0.017	-0.013
	(0.016)	(0.017)	(0.018)	(0.017)	(0.020)	(0.019)	(0.017)	(0.017)	(0.016)	(0.016)
Province FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean of DV	0.618	0.623	0.541	0.545	0.373	0.374	0.629	0.631	0.280	0.278
Observations	1362	1324	1362	1324	1362	1324	1362	1324	1362	1324
R-Squared	0.047	0.056	0.045	0.056	0.037	0.057	0.036	0.061	0.044	0.063
RMSE	0.479	0.477	0.492	0.491	0.479	0.476	0.479	0.474	0.444	0.440
	Have Trust in									
	(1) Courts	(2) Central Govt	(3) NPC	(4) Officials	(5) Army	(6) Police	(7) Local Govt	(8) Newspapers	(9) Radio and TV	(10) Social Org.
Latitude	0.014	-0.002	-0.000	0.002	-0.020*	-0.001	-0.000	0.001	0.016	0.019**
	(0.018)	(0.015)	(0.016)	(0.013)	(0.012)	(0.016)	(0.013)	(0.018)	(0.013)	(0.009)
Province FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of DV	0.440	0.674	0.602	0.184	0.687	0.440	0.282	0.222	0.278	0.103
Observations	1324	1324	1324	1324	1324	1324	1324	1324	1324	1324
R-Squared	0.048	0.032	0.051	0.049	0.054	0.041	0.033	0.030	0.044	0.051
RMSE	0.491	0.467	0.484	0.383	0.457	0.493	0.448	0.415	0.444	0.300

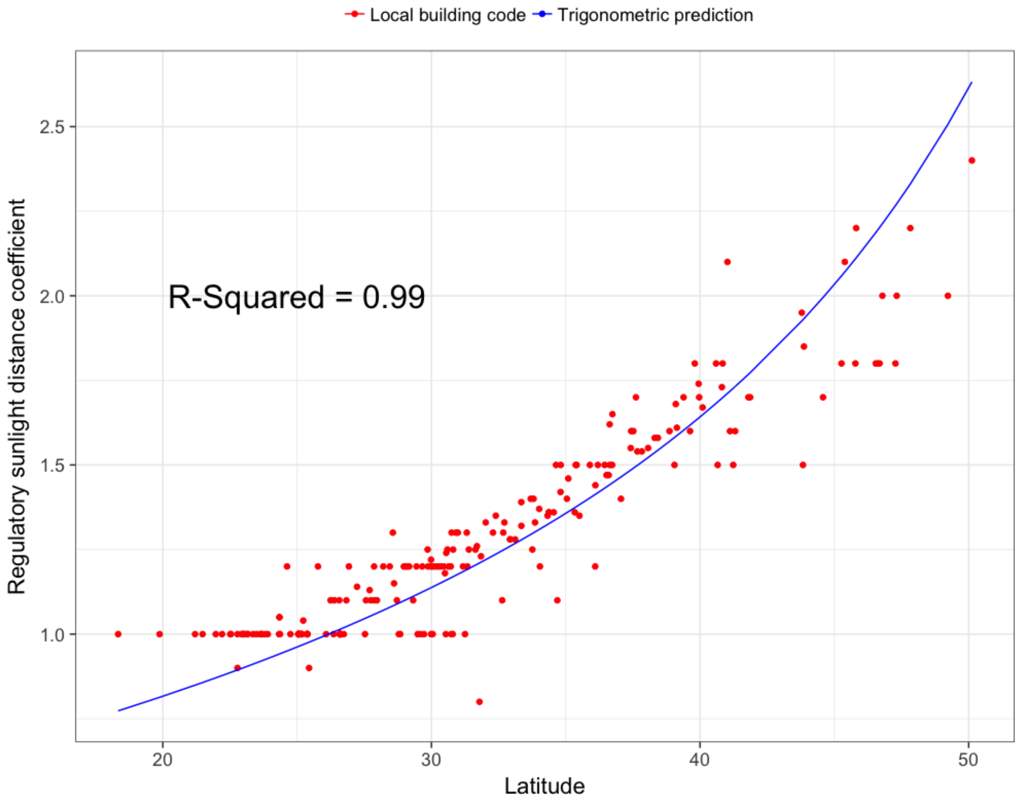
Notes: Dependent variable in Panel A is whether the respondent believes that the government should be "solely responsible" or "primarily responsible" in the listed areas. Dependent variable in Panel B is whether the respondent has "complete trust" or "substantial trust" in the listed entities. Individual controls include gender, education level, *hukou* status, age, and size of city living in. Data are for urban respondents in the 2015 China Social Governance Survey. Standard errors are clustered at city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 12: Internet Use not Systematically Different Across Latitudes

	(1) Mobile Access	(2) Other Access	(3) Frequent Use	(4) Get Info	(5) Air Opinions	(6) Social Media
Latitude	-0.004	-0.010	0.001	0.011	0.003	0.016
	(0.008)	(0.017)	(0.013)	(0.014)	(0.009)	(0.009)
Province FEs	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean of DV	0.625	0.644	0.483	0.393	0.122	0.572
Observations	1324	1324	1324	1324	1324	1324
R-Squared	0.511	0.298	0.429	0.330	0.095	0.530
RMSE	0.343	0.407	0.383	0.405	0.315	0.344

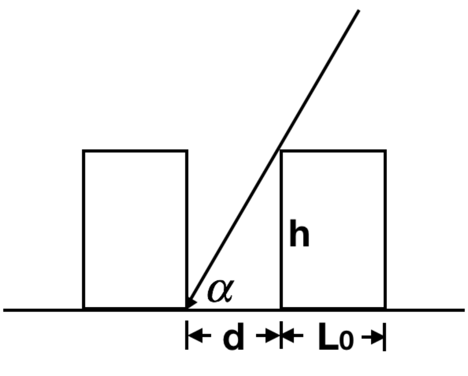
Notes: The dependent variable is whether the respondent has internet access from cell phone in column 1 , through other means (broadband, optical fiber, dial-up, etc.) in column 2, whether the respondent uses internet for at least 30 minutes per day in column 3, whether the respondent obtains political information online at least once every week in column 4, whether the respondent ever expressed political opinions online in column 5, and whether the respondent uses social media in column 6. Individual controls include gender, education level, *hukou* status, age, and size of city living in. Data are for urban respondents in the 2015 China Social Governance Survey. Standard errors are clustered at city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 1: Sunlight Distance Coefficients and Latitude



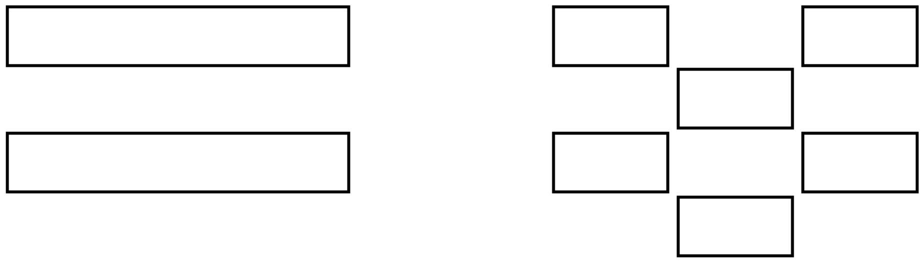
Notes: Y-axis is the sunlight distance coefficient specified in a prefectural building ordinance for residential buildings below seven stories in new urban areas. X-axis is latitude of the prefectures. The blue curve shows the theoretical minimum distance required between buildings to satisfy the sunlight policy at each latitude, equal to cotangent of the solar elevation angle at each latitude at 11 am on January 20, the day specified by the policy.

Figure 2: Solar Elevation Angle and Building Height



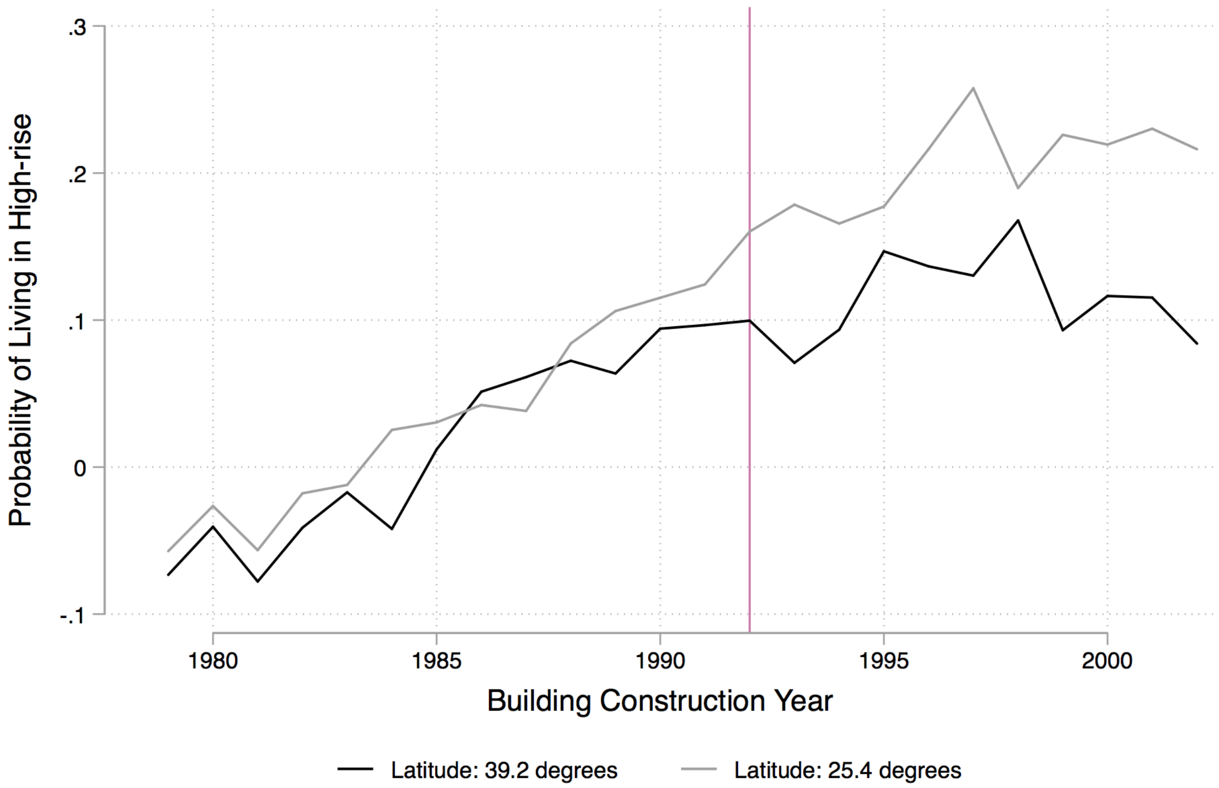
Notes: This figure shows the trade-off between building height and distance between buildings under a given solar elevation angle α.

Figure 3: Illustrations of Building Layout



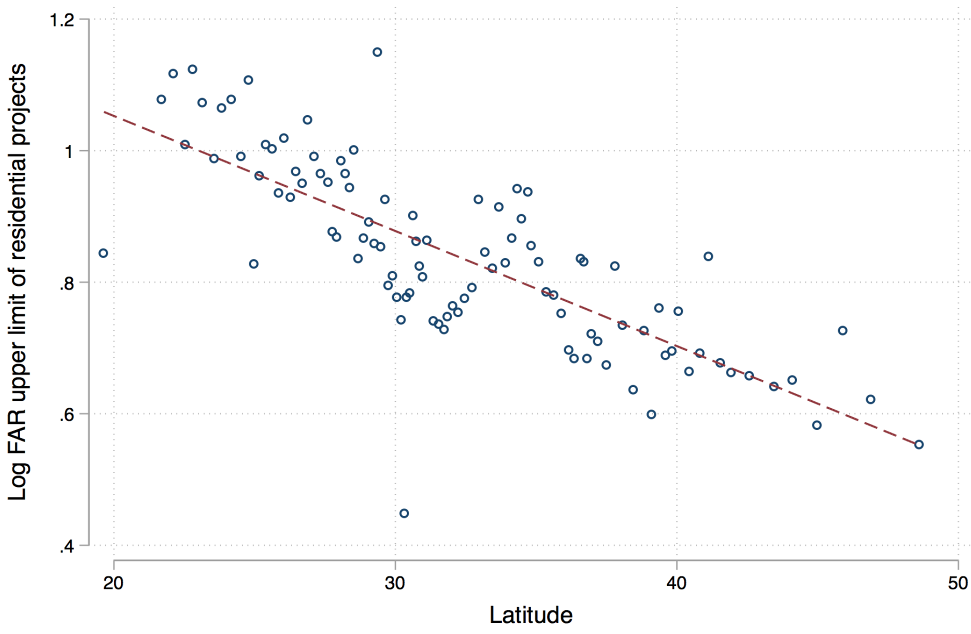
Notes: This figure shows that the two layouts of buildings achieve the same amount of floor space.

Figure 4: High-rises in the North and the South, by Construction Year

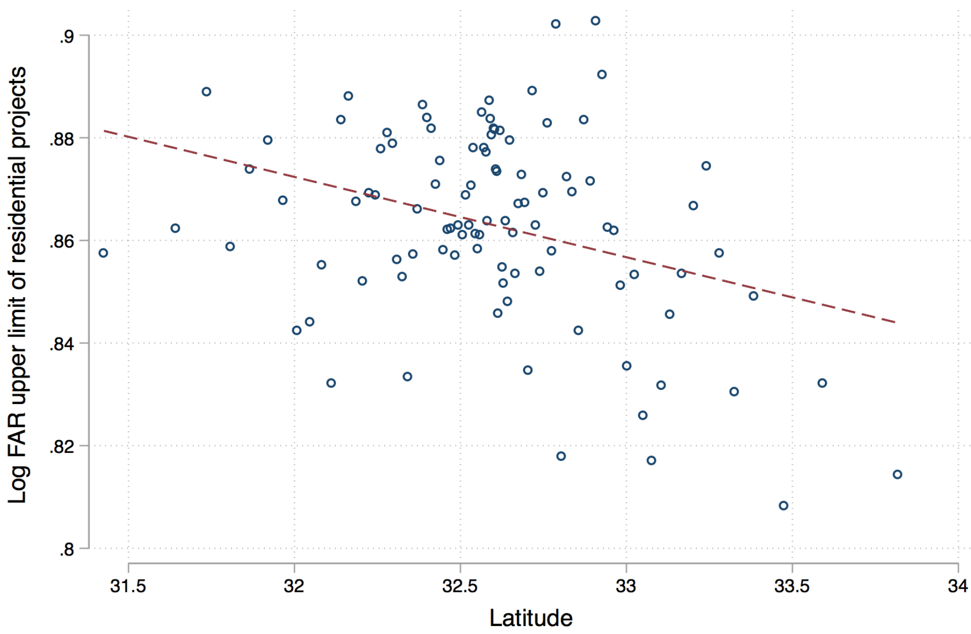


Notes: This figure shows estimated probability of a household living in a high-rise conditional on construction year of the building in a northern city (75th percentile of latitude) and a southern city (25th percentile of latitude). These estimates are obtained by combining year fixed effects and the interactions of year fixed effects and latitude in a household-level regression using 2005 census data where city fixed effects, residence controls and household controls are included. See text for details of the regression and the controls. Red vertical line shows the year preceding the sunlight policy.

Figure 5: Binned Scatterplot of Regulatory Floor Area Ratio and Latitude



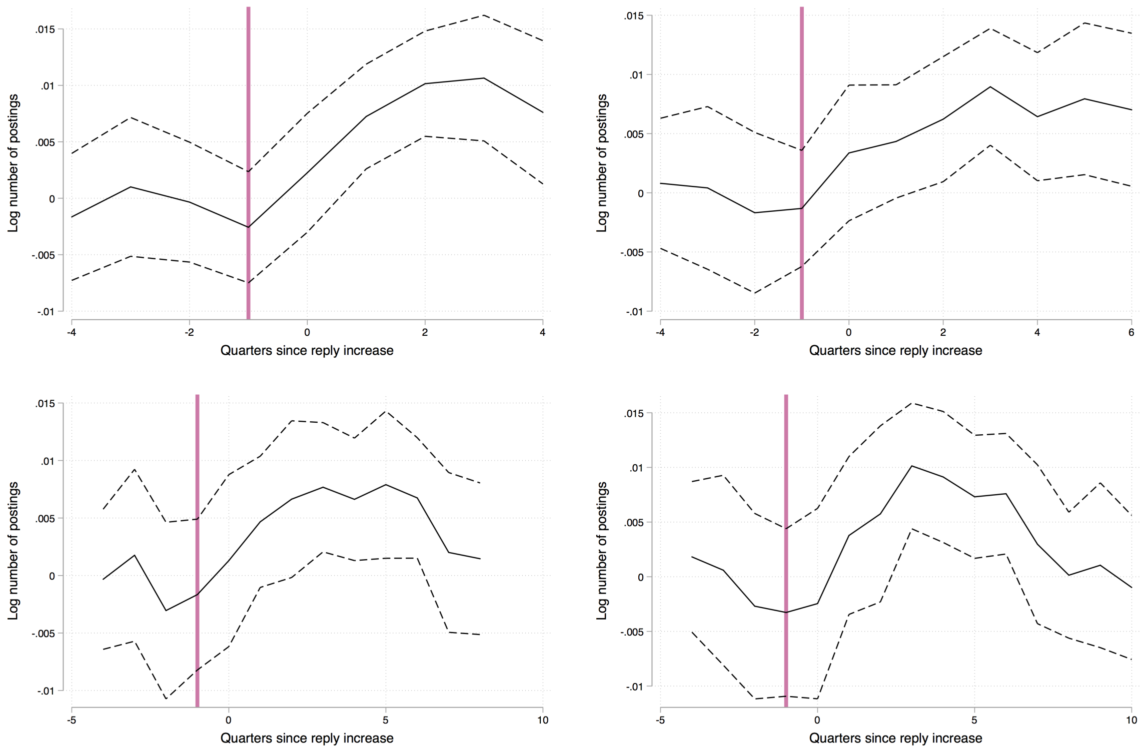
(a) No controls



(b) Within longitude-latitude cells and full controls

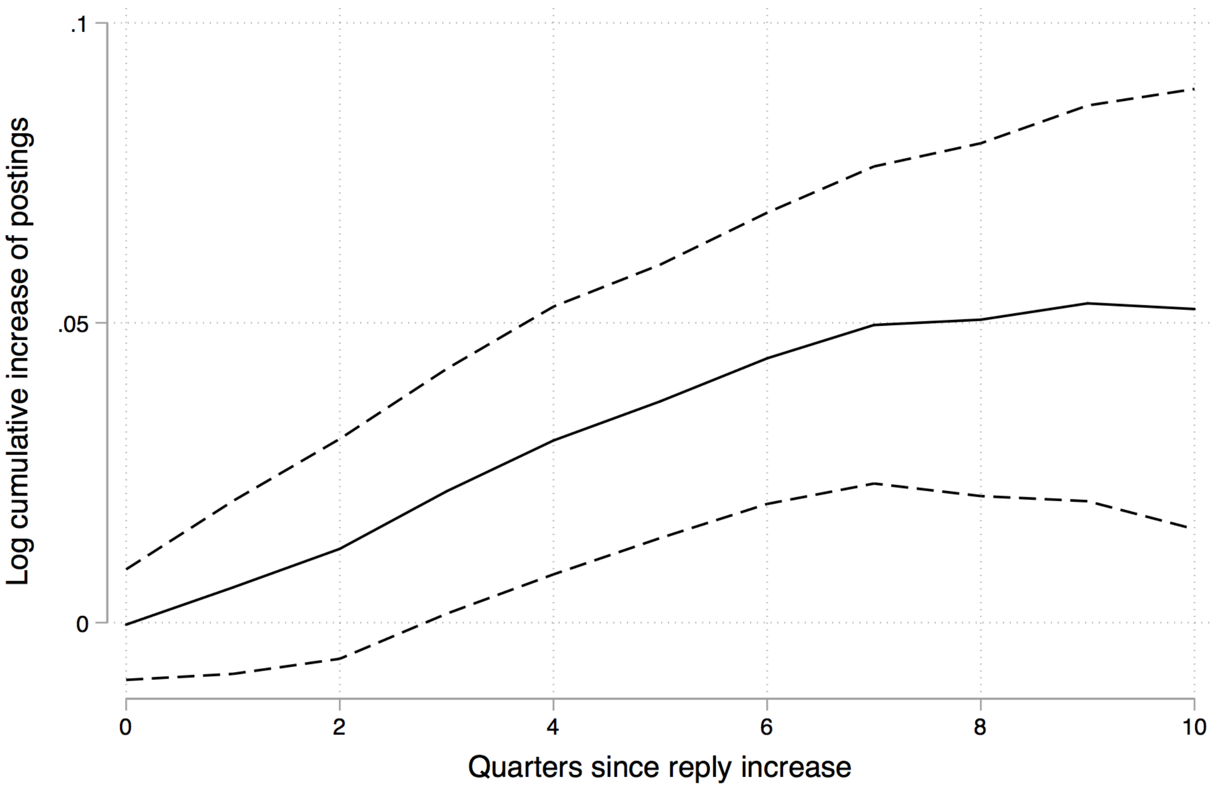
Notes: These figures are binned scatterplots of the latitude and log regulatory FAR of individual residential lots from across China. No controls are included in Panel A. Panel B shows latitude and regulatory FAR residualized within 2 degrees latitude by 0.1 degrees longitude cells and over year fixed effects, month-of-year fixed effects, nighttime light in 2005, a host of lot-level controls and a host of city-level economic controls. See details in text.

Figure 6: Dynamic Response of Postings to a 10% Increase in Government Reply Rate



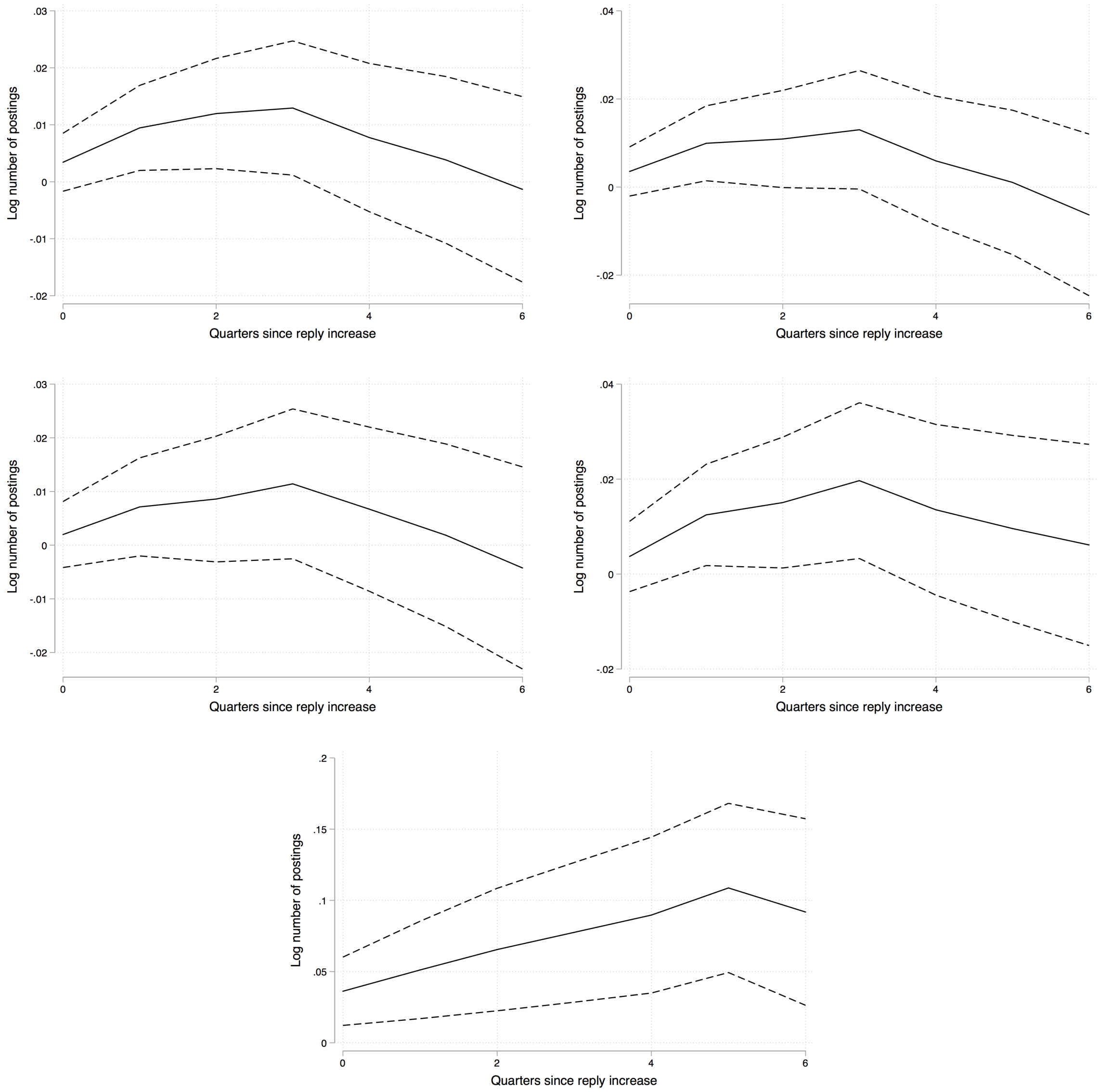
Notes: These figures show coefficients and their 95% confidence intervals estimated from distributed lag models of the dynamic response of number of posts on a topic to a 10% increase in local government reply rate to the same topic. City and topic-by-quarter fixed effects are included in all regressions. The four panels show estimates using 4, 6, 8 and 10 lags of reply rate respectively. Four leads of reply rate are also included in each model. Red vertical lines mark quarter preceding a reply rate increase.

Figure 7: Cumulative Increase in Postings after a 10% Increase in Government Reply Rate



Notes: This figure shows the cumulative increase in log number of postings on a topic after a 10% increase in local government reply rate to the same topic. City fixed effects, topic-by-quarter fixed effects and four leads of reply rate are included in the regression. Dashed lines show 95% confidence intervals.

Figure 8: Cumulative Differential Responses of Postings to a 10% Increase in Reply Rate



Notes: These figures show the cumulative difference in log number of postings between the 75th percentile and the 25th percentile of latitude each quarter after a 10% increase in government reply rate. Positive values indicate larger numbers of posts in southern cities. The five panels correspond to estimates of the five specifications in Table 8. City and topic-by-quarter fixed effects are included in all regressions. Dashed lines show 95% confidence intervals.

Appendices

A Details of Dataset Construction

A.1 Land Transactions

A substantial share of the residential land transaction records mark legal procedures at local land bureaus for converting land to "transferable", as can be seen from the description of transactions. In the planning era, housing was distributed to workers as a benefit, and the housing market did not exist. In the early days of urban housing reforms when the state entitled urban households with property rights over their dwellings to establish a housing market, households were made to pay for obtaining these property rights (Iyer et al., 2013). It was not easy to persuade everyone to pay up, and those who did not pay a sufficiently high price did not enjoy full property rights - they could not sell, rent out or use as collateral their living quarters, unless they pay later to convert their land into a transferable status. To the extent that we are interested in the expansion of urban areas and construction of new residential units, we want to exclude these transactions from the sample. There is no attribute that definitively marks such transactions. As a result, I classify a transaction as a procedural conversion to transferable land if the land is brown-field and the recorded means of transfer is "by contract". This is reasonable because new residential developments are legally required to acquire land through auctions. As a validation of this classification, transactions identified as "conversion to transferable land" have a size of 0.096 hectares (10,333 square feet) at its 90th percentile, much smaller than the size of lots in the remaining sample, whose 25th percentile is 0.11 hectares (11,840 square feet). I exclude these "conversion to transferable" transactions from analysis.

A.2 Identifying Urban Clusters

Administrative boundaries often provide poor characterizations of the boundaries of urban economies. Rozenfeld et al. (2011) propose a City Clustering Algorithm (CCA) that defines cities as maximally connected clusters of populated sites. Dingel et al. (2018) construct cities as contiguous grid cells with nighttime lights above a threshold. Vogel et al. (2018) instead use built-up landcover inferred from daytime satellite imagery to delineate cites.

I define cities as the built environment. I start with all 1 km^2 grid cells that contain at least one road intersection. There are 456,749 such cells, accounting for around 4.8% of China's land territory. I then use DBSCAN, a widely used clustering algorithm (Ester et al., 1996), to group these cells into urban clusters. The algorithm works similarly to the CCA by maximally growing clusters through combining nearby cells within a distance threshold. In a similar effort, Long (2016) defines urban areas by clustering road intersections using a different algorithm.

Figure A1 shows urban clusters discovered within Ordos, Inner Mongolia and Shanghai respectively. In the case of Ordos, we see an isolated urban area (depicted in pink) to the southwest of Ordos' main urban area (depicted in green). Kangbashi, this outlying urban area, has received international media exposure for its vast unoccupied buildings and been dubbed a "ghost town". In the case of Shanghai, we see two separate urban areas located on the outlying islands.

I use the simple test proposed by Gabaix and Ibragimov (2011) to evaluate the conformity of cluster size distribution to Zipf's Law. Figure A2 shows that these urban clusters conform almost perfectly to Zipf's Law in both land area and population. Panel A shows land area distribution across these urban clusters, where clusters smaller than 5 km^2 are excluded. Estimated coefficient of log rank (minus 0.5) on log area is -1.07, with an R^2 of 0.99. Including clusters smaller than 5 km^2 changes goodness of fit with Zipf's Law only marginally. Panel B presents population distribution across urban clusters, where population within a cluster is aggregated up from 2010 1-km^2-cell level population data from Worldpop. Clusters containing fewer than 100,000 people are excluded following Chauvin et al. (2017). Estimated coefficient of log rank (minus 0.5) on log population is -1.04, with an R^2 of 0.98. Including the less populated clusters significantly weakens goodness of fit with Zipf's Law, as is expected. Taken together, this exercise suggests that the urban clusters I identify are meaningful units of analysis.

B Value of Lifting Regulatory FAR

We saw in the text that the sunlight policy describes the cross-latitude variation in residential projects' regulatory FAR quite well. For the policy to impact urban density, these regulatory upper limits have to be actually binding. If developers never want to build as much floor space as these upper limits let them, the policy has no bite.

This can be studied by looking at how land prices respond to FAR upper limit. Brueckner et al. (2017) employs this idea to around 50,000 land transactions from China, and reports an average elasticity of price to FAR upper limit of 0.75 for residential land. Their sample is a fraction of my sample, has a different time horizon (2002-2011), and tends to concentrate in the larger cities.

I estimate the following equation:

$$\log \text{Price}_{pct} = \beta \log \text{FAR}_p + \lambda_c + \gamma_1' X_p + \gamma_2' X_t + \epsilon_{pct}$$

where the dependent variable is log of real price per square meter. λ_c is a city fixed effect. X_p and X_t are the same project-level and time-level controls as in the text. We only include lots transferred through auctions, as auctions better reflect the market value of land. This excludes the vast majority of public housing projects.

Results are presented in Table A2. Our preferred specification is in column 3 with the full set of controls, where the elasticity of price with respect to FAR upper limit is estimated at

0.60. If the upper limit is not binding at all, we would expect the elasticity to be 0. 0.60 implies a substantial degree of constraint on average. Columns 4 through 7 look at heterogeneity in the restrictiveness of regulatory FAR. We see that housing projects far from the city center are not significantly less constrained by regulatory FAR. If anything, they appear to be more constrained. This is again consistent with the perception that there is often strong demand at the edges of Chinese cities. There is not much heterogeneity with respect to city population, and cities with higher GDP per capita appear to be slightly less constrained. In column 7, we see that cities with higher population densities are also not significantly more constrained by air right regulations.

C Relationship between Regulatory FAR and Latitude, Raw Data

Figure A3 plots the raw data of regulatory FAR upper limit of individual residential lots against their latitude, and fits generalized additive models to estimate conditional means of FAR given latitude without controlling for any covariates. Panel A does this for all residential land, and there appears to be a negative relationship between latitude and FAR - lots in the south accommodate more floor space. The curve dips on the left end reaching into Hainan Island (the gap above 20 degrees shows the strait between the mainland and the island). To increase confidence that this negative relationship is not induced by specific regions, panels B-D separate China into three slices by longitude, and repeat the exercise. The pattern is present in all cases.

D Density in Beijing, Shanghai and Shenzhen

Urbanization rate in China grew from 17.9% in 1978 to 57.9% in 2017. The area of urban land grew by 498% between 1992 and 2015 (Xu et al., 2016). Despite this rapid expansion of urban areas across China, the urban fabric varies greatly across cities. Figure A4 shows the distribution of population, points of interest (businesses and public facilities) and road intersection density in three of China' s four largest cities: Beijing, Shanghai and Shenzhen. Though comparable in overall population and GDP per capita,[46] Shanghai is much denser than Beijing, and Shenzhen much denser than Shanghai. The median 1-km urban grid cell in Shenzhen has 0.64 log points (64%) more population than the median urban cell in Shanghai, which in turn has 1.14 log points (114%) more population than the median urban cell in Beijing.

[46] In 2010, a census year, population of Beijing is 18.8 million, Shanghai 22.3 million, Shenzhen 10.4 million. In 2016, GDP per capita of Beijing is 118 thousand yuan, Shanghai 117 thousand yuan, Shenzhen 167 thousand yuan. These data all correspond to districts directly governed by these cities, excluding counties within their jurisdiction boundaries.

This difference in density is even greater in points of interest. The median 1-km urban cell in Shenzhen has 1.23 log points more POIs than the median urban cell in Shanghai, which in turn has 0.86 log points more POIs than its counterpart in Beijing. Turning to the density of road junctions, which characterizes block size and constitutes a most visible aspect of the streetscape, the median urban cell in Shenzhen contains 0.69 log points more road junctions than that in Shanghai. The difference between Shanghai and Beijing is smaller but still substantial, with the median urban cell in Shanghai containing 0.25 log points more road junctions than the median urban cell in Beijing. As these density plots could be sensitive to definitions of what constitute a city, I further restrict the sample to city centers, defined as 1-km grid cells with a 2013 nighttime light luminosity above 50 on a scale of 0 to 63, and plot the same kernel densities for the three cities in Figure A5. We see that the patterns largely remain, though the magnitude of differences becomes somewhat smaller, and the ranking between Shanghai and Beijing in terms of road junction density is reversed.

Lying on the south coast of China and bordering Hong Kong, Shenzhen is 8.3 degrees south of Shanghai in latitude, which in turn is 9 degrees south of Beijing. Although the differences in density across these three cities are certainly not entirely due to the sunlight policy, these patterns are indeed consistent with the policy playing a role.

E Petition Topics

These topics are obtained from a 30-topic LDA model trained on the full set of postings on LLMB up until mid-2016 (for details of the methodology, see Jiang et al. (2018)). The topics generally have very clear meanings. I group them into urban topics and rural topics, and focus on urban issues in this paper. I further classify the urban topics into grievance topics and non-grievance topics, which is quite evident from their contents. Below I list words with highest probability weights in each topic.

E.1 Urban Topics, Grievances

Topic 1: residential compound, property owner, real estate property, have not, residents, problem, elevator, occupants, property management company, inside, management, property management fee, garden, agency, community

Topic 2: vehicle, road, traffic, severe, safety, passenger, crossroads, streetlight, do not have, agency, segment of road, travel, cause, influence, hope

Topic 4: garbage, pollution, severe, residents, environment, life, waste water, influence, one, nearby, hope, leader, health, emission, production

Topic 5: bus, public transport, driver, taxi, vehicle, passenger, convenience, time, have not, travel, hours, hope, traffic, car, train station

Topic 6: police station, one, pyramid selling scam, police, personnel, happen, police department, at that time, call the police, tour guide, touring, have not, hope, friends, police

Topic 7: hospitals, father, reimbursement, doctors, life, mother, treatment, family, medical insurance, children, cannot, handicapped, elderly, in-patient, expenses

Topic 8: demolishing, home, settlement, upgrade, government, house, have not, compensation, leader, my home, construction, planning, secretary, shanty-town, relocated households

Topic 11: household registration, children, handle, have not, policy, certificate, work, cannot, police branch, one, parents, leader, need, please, residence

Topic 12: house, residents, severe, safety, construction, department, remove, in the process of construction, occupants, house, problem, influence, residential compounds, have not, cause

Topic 14: operate, market, department, one, influence, severe, urban management officer, hope, manage, environment, Internet cafe, leader, secretary, gambling, commercial tenant

Topic 15: school, students, children, teacher, primary school, parents, kindergarten, education, middle school, make-up class, attend school, education bureau, leader, learn, one

Topic 17: developer, property owner, real estate transaction, apartment, have not, now, residential compound, property ownership certificate, handle, contract, government, develop, leader, already, purchase

Topic 18: test, work, college students, teacher, driving school, civil servants, have not, not yet, graduation, participate, professional, village officer, employment, one, graduating student, test taker

Topic 21: charge fees, charge, fee, regulation, standard, price, state, whether, reasonable, unreasonable charges, please, document, natural gas, request

Topic 22: company, salary, employee, firm, have not, leader, migrant worker, limited-liability company, worker, secretary, unit, arrears, projects, labor, ten thousand yuan

Topic 24: telephone, have not, handle, staff, information, Internet, company, cannot, one, broadband, make a call, complain, bank, mobile, phone

Topic 26: salary, work, teacher, have not, unit, personnel, retirement, employee, compensation package, life, secretary, policy, leader, state, now

Topic 28: residents, noise, affect, severe, residential compounds, at night, life, disruptive, department, everyday, in the process of construction, normal, environment, one

Topic 30: residential compounds, heating, company, solve, residents, have not, heat, water supply, heat supply, occupants, leader, life, now, water supply shutdown

E.2 Urban Topics, Non-grievances

Topic 3: secretary, problem, solve, leader, hope, hello, not yet, respect, reflect, mayor, thank you, take time from a busy schedule, attention, now, ask

Topic 9: have not, now, one, know, leader, folks, hope, really, secretary, government, once, why not, cannot, location, see

Topic 10: planning, construction, railway, residents, convenience, wide road, traffic, highway, have not, public transport, residential compound, connect, please, nearby

Topic 19: regulation, relevant, department, undertake, require, law, illegal, condition, government, court, report, have not, state, behavior, unit

Topic 20: leader, local, People's Daily website, message board, source, secretary, hello, once, respect, hope, have not, hi, now, thank you, pay attention to

Topic 25: government, problem, mass, work, people, society, leader, folks, cadre, department, one, real, hope, proceed, should

Topic 27: develop, construction, city, economy, one, suggestion, travel, culture, hope, rural areas, people, whole country, hometown, environment, secretary

E.3 Rural Topics

Topic 13: villager, farmer, reservoir, severe, cultivation, cause, damage, production, government, extraction, river course, now, land, secretary, farmland

Topic 16: villager, road, path, leader, secretary, now, have not, build a road, travel, cement road, suddenly, road surface, bumpy, raining day

Topic 23: village, farmer, have not, secretary, leader, policy, state, compensation, subsidy, low-income allowance, this year, loan, one, finance, poverty alleviation

Topic 29: villager, land, have not, secretary, farmland, compensation, land taking, village committee, village cadres, occupy, my home, rural residential land, forcefully, compensation package

F Longitude as a Placebo

Since density does not systematically vary over longitudes, we should expect the response speed of petitions to government reply rate not to vary across longitudes. Regressions reported in Table A3 repeat those in Table 8, with predicted log FAR being replaced by longitude. We see that almost all of the coefficients are insignificant, suggesting that the speed of information transmission does not vary across longitudes.

Table A1: Data Sources

Data	Description and Citations	Time coverage	Source
Land transactions	Over 1.34 million land transaction records containing information on geographical location, price, zoned use, regulatory floor area ratio, size, etc.	2000-2016, covers 69.8% in area of all urban land transferred through the government post 2006	Ministry of Land and Resources
Road junctions	Number of road junctions within 1-km grid cells for all China, Zhou and Long (2016), Liu and Long (2016)	2011	Beijing City Lab
Population density	Population within 1-km grid cells for all China, Gaughan et al. (2016)	1990, 2000, 2005, 2010, 2015	WorldPop and Gridded Population of the World
City-level economic characteristics	GDP per capita, population, population density, share of population attending primary school, middle school, etc.	1994-2014 for prefectures, 1997-2013 for counties	China Data Center, University of Michigan
Nighttime light	Satellite-detected nighttime light at 1-km-grid-cell level for all China	1992-2013	DMSP-OLS
Ruggedness	Share of land with different slopes within 1-km grid cells, Fischer et al. (2008)		FAO Harmonized World Soil Database v1.2
Climate	Average January temperature, average July temperature, and average precipitation within 1-km grid cells, Fick and Hijmans (2017)		WorldClim v2
Population mini-census	A 1% population survey carried out by National Bureau of Statistics. Contains rich household and individual level information	2005	
Online petitions	Around 900,000 postings on a national petition platform that hosts sub-boards for all administrative units	2008-2016	Local Leader Message Board, People's Daily

Table A2: Value of Loosening Regulatory FAR Upper Limit

	Dependent Variable: Log Real Price per Square Meter						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log FAR upper limit	0.843***	0.743***	0.601***	0.573***	0.602***	0.638***	0.624***
	(0.042)	(0.020)	(0.017)	(0.022)	(0.033)	(0.032)	(0.033)
Distance to city center, above median			-0.269***	-0.319***	-0.265***	-0.263***	-0.269***
			(0.017)	(0.031)	(0.018)	(0.018)	(0.017)
Log FAR upper limit × Distance to city center, above median				0.057*			
				(0.029)			
Log FAR upper limit × Population, above median					-0.008		
					(0.039)		
Log FAR upper limit × GDP pc, above median						-0.071*	
						(0.038)	
Log FAR upper limit × Population density, above median							-0.036
							(0.039)
City FEs	No	Yes	Yes	Yes	Yes	Yes	Yes
Time FEs	No	Yes	Yes	Yes	Yes	Yes	Yes
Lot-specific controls	No	No	Yes	Yes	Yes	Yes	Yes
Observations	164346	164346	144241	144241	138502	138257	144241
R-Squared	0.128	0.503	0.558	0.558	0.561	0.561	0.558
RMSE	0.979	0.744	0.704	0.704	0.706	0.706	0.704

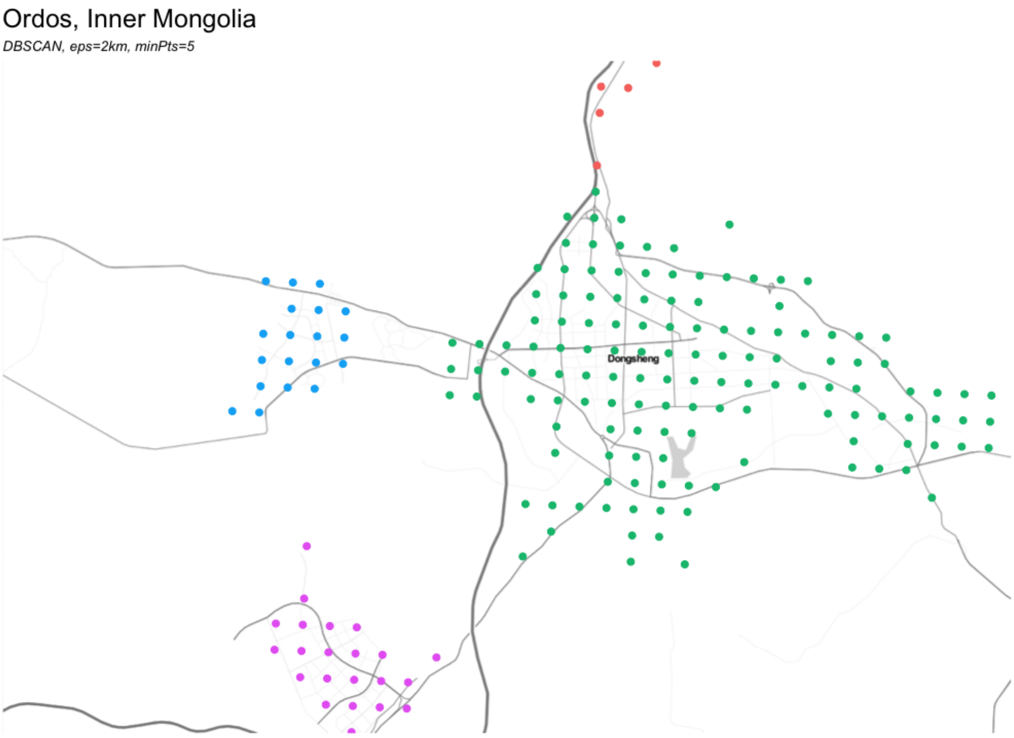
Notes: These regressions concern residential development projects across China for which the land parcel is auctioned off between 2007 and 2016. Dependent variable is log real price per square meter. Time FEs include year and month-of-year fixed effects. Lot-specific controls include whether the lot is greenfield or brown-field, type of residential building, fixed effects for land grade, whether parcel size is above median, whether distance to 2005 city center is above median and 2005 nighttime lights. City-level population and GDP per capita are 2006 levels obtained from statistical yearbooks. City-level population density is 2010 level calculated from Worldpop data. Standard errors are clustered at city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Petition Reaction to Government Replies, Longitude Placebo

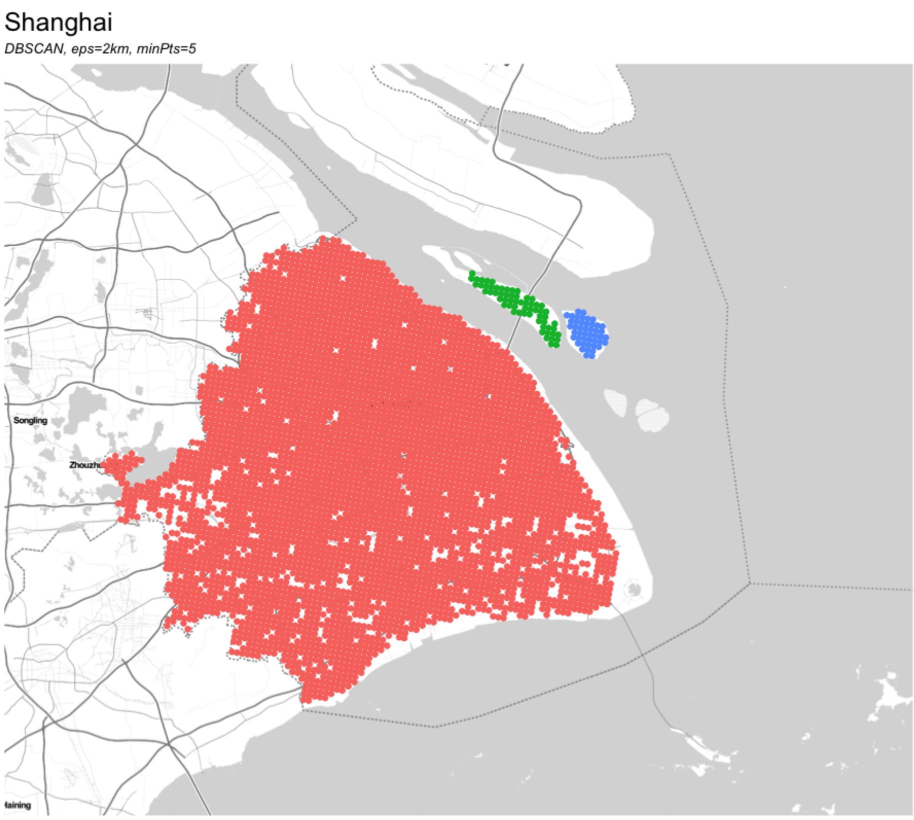
	Dependent Variable: Log Number of Posts				
	(1)	(2)	(3)	(4)	(5)
Reply rate × Longitude	-0.002	-0.002	-0.000	-0.001	-0.039**
	(0.003)	(0.003)	(0.003)	(0.004)	(0.015)
Reply rate$_{t-1}$ × Longitude	0.001	-0.002	-0.002	-0.005	-0.016
	(0.002)	(0.003)	(0.003)	(0.003)	(0.013)
Reply rate$_{t-2}$ × Longitude	0.001	-0.002	-0.001	-0.003	0.020
	(0.002)	(0.003)	(0.003)	(0.003)	(0.013)
Reply rate$_{t-3}$ × Longitude	0.004*	0.001	0.002	0.002	-0.023*
	(0.002)	(0.003)	(0.003)	(0.003)	(0.013)
Reply rate$_{t-4}$ × Longitude	0.005*	0.003	0.003	0.003	-0.020
	(0.003)	(0.003)	(0.003)	(0.003)	(0.014)
Reply rate$_{t-5}$ × Longitude	0.001	0.000	0.002	0.001	-0.002
	(0.002)	(0.003)	(0.003)	(0.003)	(0.012)
Reply rate$_{t-6}$ × Longitude	0.000	-0.004	-0.003	-0.006**	-0.020
	(0.002)	(0.003)	(0.003)	(0.003)	(0.013)
City FEs	Yes	Yes	Yes	Yes	Yes
Topic-Quarter FEs	Yes	Yes	Yes	Yes	Yes
Education interactions	No	Yes	Yes	Yes	Yes
Economic interactions	No	No	Yes	Yes	Yes
Communication interactions	No	No	No	Yes	Yes
Province FE interactions	No	No	No	No	Yes
Observations	19317	19252	18538	18176	18176
R-Squared	0.622	0.625	0.613	0.618	0.626
RMSE	0.565	0.564	0.566	0.566	0.562

Notes: Units of observation in these regressions are city-quarter-topic tuples. Topics are obtained from a 30-topic LDA model. Each post is assigned to its highest topic. Dependent variable is log number of posts about a given topic posted in a given city's Local Leader Message Board in a given quarter. Coefficients of reply rate and its lags are omitted. Education interactions are reply rate and its lags interacted with shares of six education levels of city residents. Economic interactions are reply rate and its lags interacted with city's GDP per capita, fiscal expenditure to GDP ratio, social security spending to GDP ratio and social assistance to GDP ratio. Communication interactions are reply rate and its lags interacted with the following city-level variables normalized by population: annual passengers transported, annual freight transported, number of post offices, landline users, mobile phone users, broadband internet users. Province FE interactions are reply rate and its lags interacted with province fixed effects. Standard errors are clustered at city-topic level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure A1: Identifying Urban Clusters



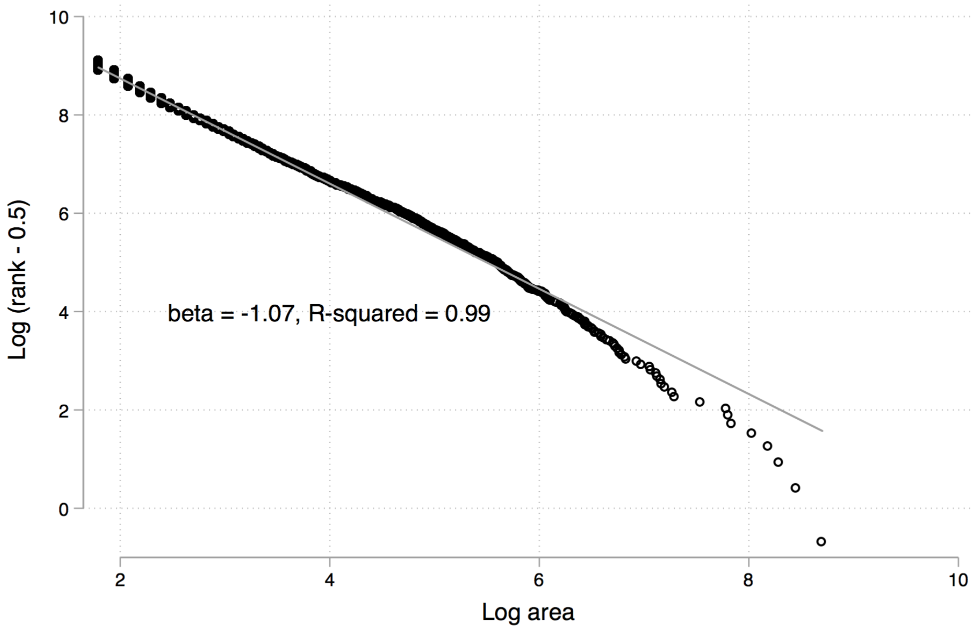
(a) Ordos



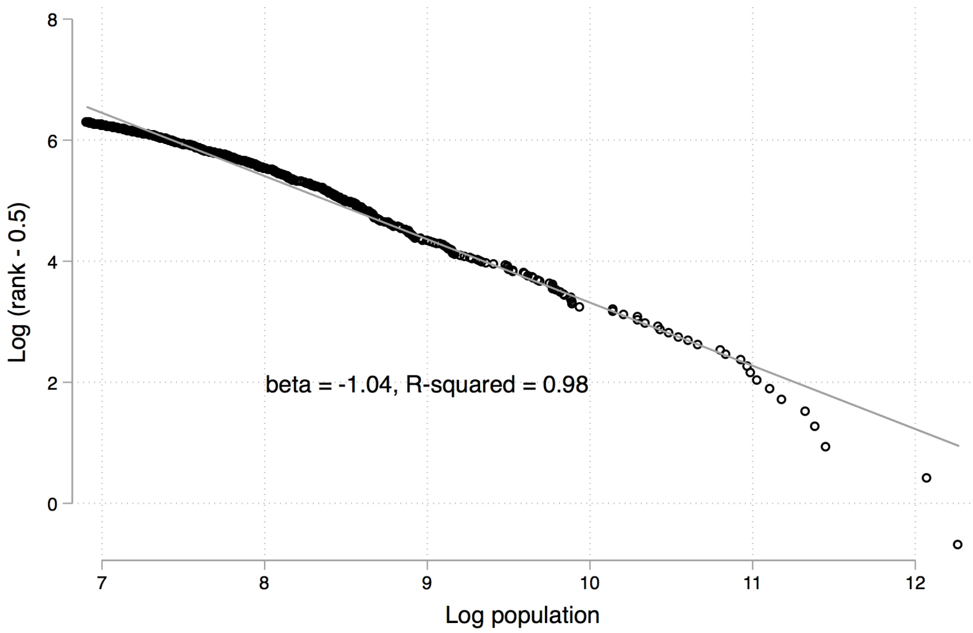
(b) Shanghai

Notes: These figures show urban clusters discovered with DBSCAN within jurisdicition boundaries of Ordos (Panel A) and Shanghai (Panel B). Each dot is a 1 km^2 grid cell containing at least one road intersection. Distinct colors indicate separate clusters. See details of the clustering method in text.

Figure A2: Zipf's Law for Urban Clusters



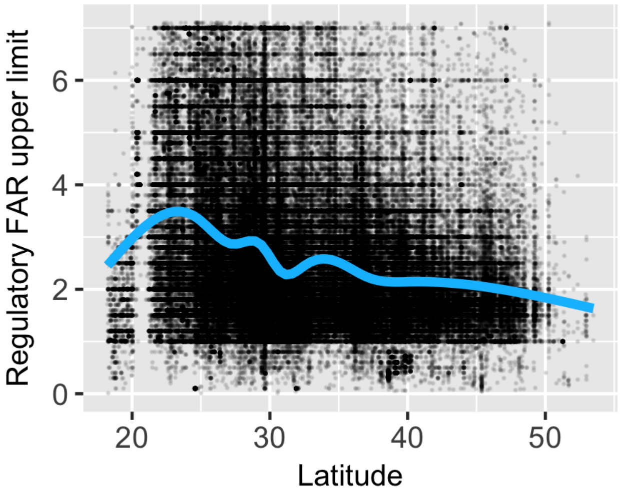
(a) Area



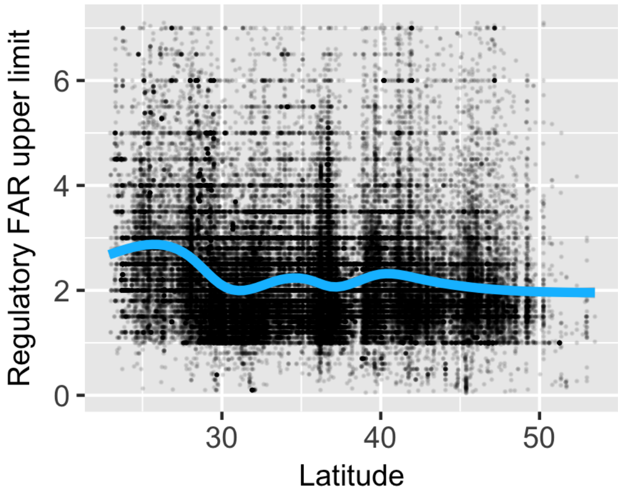
(b) Population

Notes: Panel A shows for identified urban clusters log of rank (minus 0.5) against log of cluster area, following the simple test of Zipf's Law in Gabaix and Ibragimov (2011). Urban clusters smaller than 5 km^2 are excluded from Panel A, but including them does not alter the tight Zipf fit. Panel B shows for identified urban clusters log of rank (minus 0.5) against log of cluster population in 2010, which is aggregated up from Worldpop 1 km^2 population counts. Urban clusters containing fewer than 100,000 people are excluded following the literature. Including these small towns significantly weakens the overall fit with Zipf's Law.

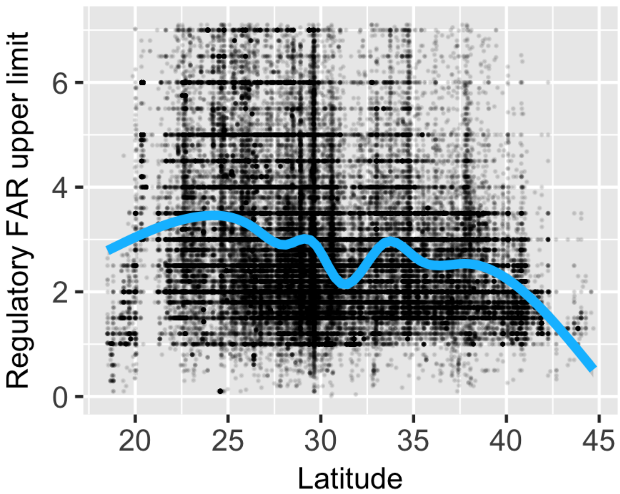
Figure A3: Relationship Between Latitude and Regulatory FAR Upper Limit



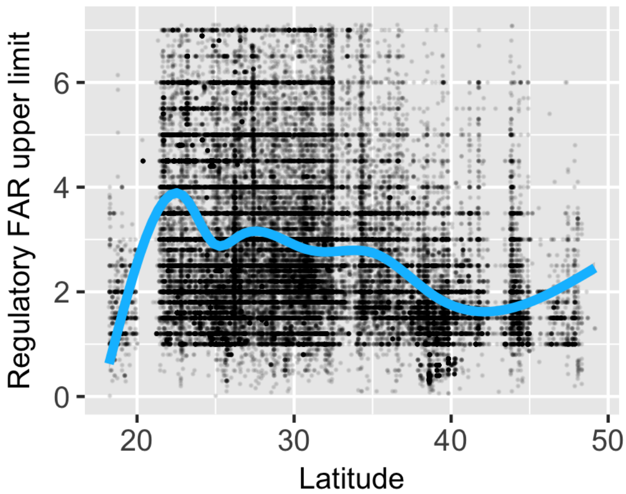
(a) All China



(b) Eastern China (East of 116°E)



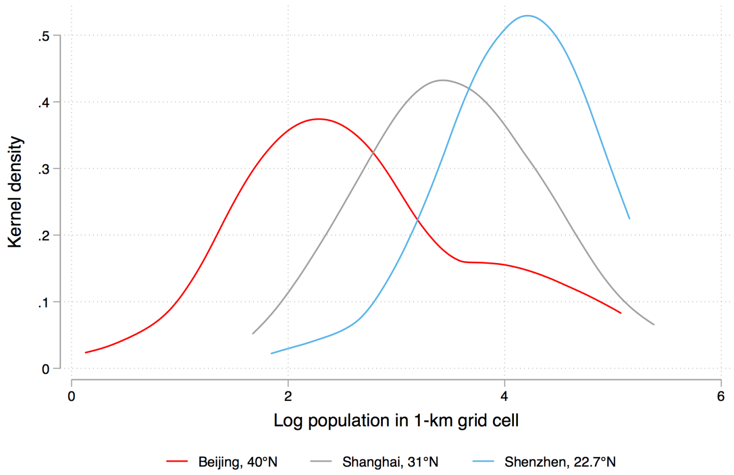
(c) Central China (110°E-116°E)



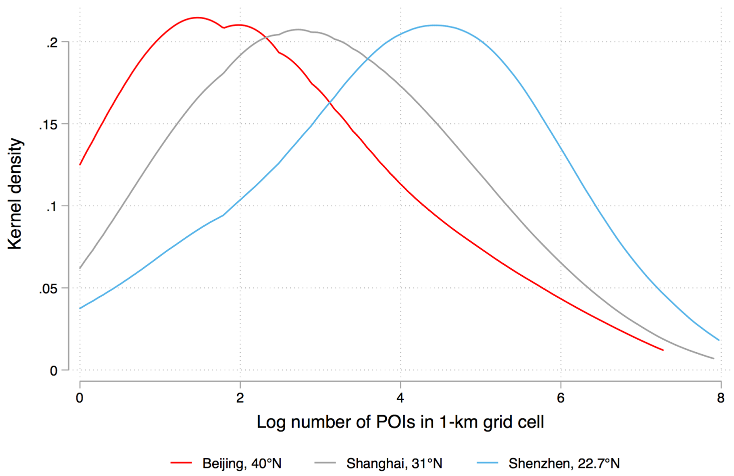
(d) Western China (West of 110°E)

Notes: These figures are scatterplots of the latitude and regulatory FAR of individual residential lots from across China. Blue curves are generalized additive model non-parametric fits.

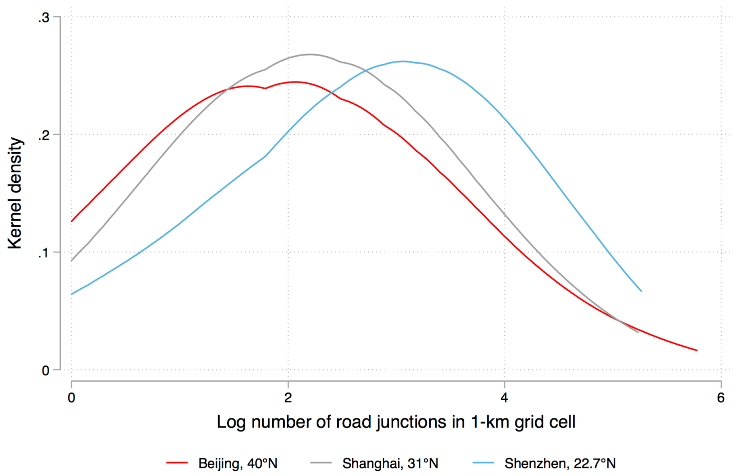
Figure A4: Urban Density in Beijing, Shanghai and Shenzhen



(a) Population



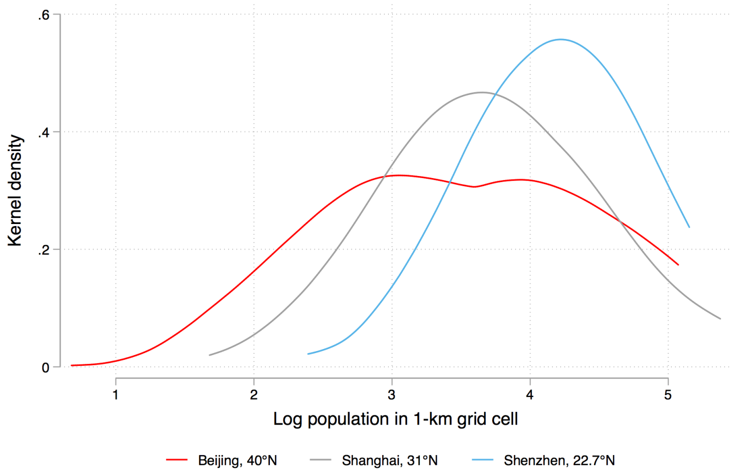
(b) Points of interest



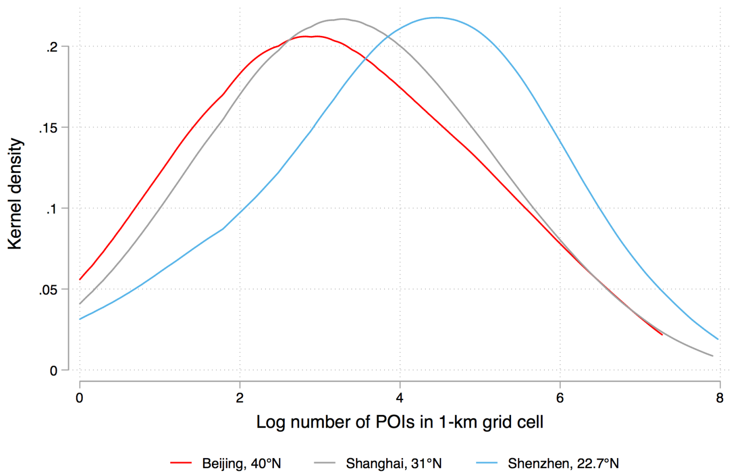
(c) Road junctions

Notes: These are kernel density plots of population, the number of points of interest (businesses and public facilities) and road junctions in 1-km urban grid cells in Beijing, Shanghai and Shenzhen, where "urban" grid cells are defined as those containing at least one road junction. Only districts directly governed by these cities are included. Counties within the jurisdiction boundaries of these cities are excluded.

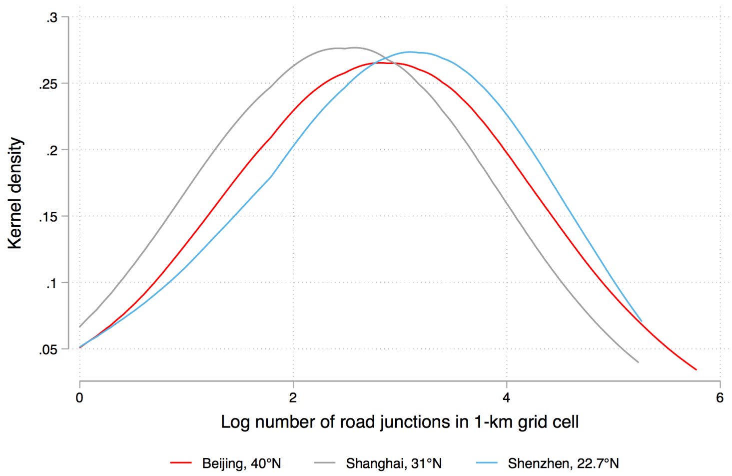
Figure A5: Urban Density in Beijing, Shanghai and Shenzhen, City Centers



(a) Population



(b) Points of interest



(c) Road junctions

Notes: These are kernel density plots of population, the number of points of interest (businesses and public facilities) and road junctions in 1-km grid cells that contain at least one road junction and have a 2013 nighttime light luminosity of at least 50 (on a scale of 0 to 63) in Beijing, Shanghai and Shenzhen. Only districts directly governed by these cities are included. Counties within the jurisdiction boundaries of these cities are excluded.