Forthcoming in *Journal of Urban Economics*.

Walled Cities in Late Imperial China

Yannis M. Ioannides and Junfu Zhang*

Abstract

For thousands of years, the Chinese and many other nations around the world built defensive walls around their cities. This phenomenon is not well understood from an economic perspective. We rationalize the existence of city walls by developing a simple model that relates the dimensions and quality of city walls to a set of economic and geographic variables. We report an empirical analysis using hand-collected and previously unutilized data on city walls in the Ming (1368-1644) and Qing (1644-1911) Dynasties. We find that the circumference of a city wall is correlated with local economic and geographic conditions, that wall size is positively correlated with population size in the jurisdiction, and that frontier cities subject to a higher probability of attack tend to have stronger city walls. We examine the physical size distribution of walled cities in late imperial China and show that city sizes above a certain cutoff follow a Pareto law, although the Pareto coefficient decreases algebraically with the cutoff point. This result complements findings in the existing literature that focuses almost exclusively on the population size distribution of cities. We also find that cities with walls in the past have higher employment and population densities at the present time.

Keywords: City walls, Zipf's law, power law, Pareto distribution, persistence, China.

JEL Classification: R12.

---

*Ioannides: Department of Economics, Tufts University, Medford, MA 02155; Tel: 617-627-3294; Email: yannis.ioannides@tufts.edu. Zhang: Department of Economics, Clark University, Worcester, MA 01610; Tel: 508-793-7247; Email: juzhang@clarku.edu. Mark Henderson provided background information about the collection of the data on city walls in Qing Dynasty used in this study. We thank Rui Du for his superb assistance. This paper has also benefited from comments by Anna Hardman, Yueran Ma, Alexander Moradi, Esteban Rossi-Hansberg, Xiaocong Xu, and seminar participants at the Boston Fed, Fudan University, the 9th Urban Economics Association Meeting in Bethesda, MD, the 2015 NBER Urban Economics Summer Institute in Cambridge, MA, the 2015 World Economic History Congress in Kyoto, Japan, and the 15th Conference on Research on Economic Theory and Econometrics in Tinos, Greece. We are truly grateful to the editor, Gilles Duranton, and two anonymous referees, whose comments and suggestions helped improve the paper immensely. Zhang would like to thank the late professor Liang Sicheng, whose effort to preserve Beijing's city walls has long been an inspiration.

> There is no real city in Northern China without a surrounding wall, a condition which, indeed, is expressed by the fact that the Chinese use the same word *ch'eng* for a city and a city-wall: for there is no such thing as a city without a wall. It is just as inconceivable as a house without a roof. Osvald Sirén (1924, pp.1-2)

1 Introduction

China's continuous history over thousands of years as a unitary state provides a useful context for studying a large urban system from a historical perspective. Many important questions arise in this context: What factors determine urban development in history? How does the distribution of city sizes look like in history? Do any historical features of cities have persistent effects on cities today? Although questions like these are clearly of great scientific interest, urban economists have not systematically studied them in the context of China primarily due to lack of data. In this paper, we aim at studying systematically the urban system in late imperial China. Our effort is made possible by the observation that almost all of Chinese cities had defensive walls in history, and data on such walls are consistently available. We establish that dimensions of city walls are related to economic fundamentals and thus they can be used to proxy for economic variables that are normally unavailable. Our empirical analysis takes advantage of rich data on walled cities, and demonstrates an approach that will hopefully prove useful for researching other urban issues in history.

Archaeological evidence reveals that as early as over 4,000 years ago, human settlements in China were often surrounded by walls. Throughout the recorded history of China, major cities always had defensive walls. In the imperial period, the great majority of urban residents lived in walled cities (Chang, 1977). It is a surrounding wall that most Chinese people used to essentially distinguish a proper city from towns and villages.[1] City walls represented a most salient feature of Chinese cities throughout history until the mid-twentieth century, when the government sought to demolish city walls all over the country in the name of shaking off "the shackles of the past." Today, complete city walls have been preserved for only a few Chinese cities, including Jingzhou, Pingyao, Xi'an, and Xingcheng. In most other cities, one can hardly see a trace of a city wall.

City walls in China were built primarily for defensive purposes. Typical city walls were thick enough to allow soldiers, horses, or even chariots to march on the top. They were usually fortified by adding battlements, towers, and barbican gates (see Figure 1). Earlier

---

[1] City walls were also common in other civilizations. According to Homer's *Iliad*, a story which took place about 3,200 years ago, the city of Troy had strong walls with high towers and great gates. At the archaeological site of Troy in Hisarlik, Turkey, excavations revealed that a stone-walled human settlement existed more than 4,000 years ago. According to the Bible, when Moses led the Israelites out of Egypt, which probably occurred some 3,400 years ago, many cities in the Middle East were fortified by city walls. The walls of Jerusalem and Damascus are mentioned repeatedly in the Bible. In some of these cities, such as Jerusalem, medieval city walls have survived and remained a tourist attraction today. Elsewhere, as in Paris, their survival till the modern era allowed city planners to build modern amenities. For example, the site of the Thiers Wall, the last remaining of the city walls of Paris, provided vacant land for the Boulevard Peripherique, the ring road that runs roughly along Paris' municipal boundaries.

Figure 1: City wall of Beijing in late 19th century

Source: http://www.photographium.com/south-gate-beijing-china-1874, with authors' modifications.

city walls were generally made of rammed earth only. Starting in the Ming Dynasty (1368-1644), it became a common practice to have city walls faced with bricks. Most Chinese cities had moats surrounding their city walls.[2]

One of our intentions in this study is to demonstrate that the evidence on city walls in late imperial China, which we discuss at length below, functions as an important window to understanding China's urban development. Defense considerations are closely related to a city's geographic location. Cities surrounded by a rich hinterland would naturally be preyed upon by bandits or disgruntled nobles. Coastal locations may be more vulnerable to incursions by foreigners, but their wealth may have generated envy by non-foreigners too. Cities in remote locations were critically important in the Chinese state's ability to fend off invaders, and had to be able to provide for their sustenance during sustained sieges. City walls defined city life in many instances.[3] The well-kept records of city walls has allowed researchers to use their physical size as a proxy for their populations in case where historical data on populations are unavailable (Skinner 1977b, 1977c).

Despite the long history of city walls, modern urban economics has paid little attention to it. The classic monocentric city model puts the city on a featureless plain. The balance between agglomeration economies and diseconomies determines the physical structure of the city. Such models have no place for a city wall. In fact, to the best of our knowledge, no model of city walls exists.

In the present study, we rationalize the existence of city walls using a simple mono-

---

[2] Moats and city walls were usually built at the same time: The earth used for the city wall was dug out of the ground right outside the wall, resulting in a ditch that was then filled with water to serve as a moat.

[3] The early English literature on the life of the Chinese people almost surely would refer to city walls. See for example Buck (1931) and Waln (1933), both of whom lived in China for many years and wrote extensively about its people during the early 20th century. We thank Anna Hardman for the latter reference.

centric city model. In our model, city walls are a man-made amenity built to protect residents, property, and valuable belongings from enemies and bandits. The model relates the dimensions of a city wall to key economic variables.

We test the predictions of the model by using two unique and hitherto unutilized data sources. The first data set was constructed by digitizing hand-collected information from a monumental work, the 130-chapter *Important Notes on Reading the Geography Treatises in the Histories* (*Du Shi Fang Yu Ji Yao*), written by Gu Zuyu (1631-1692), an early Qing Dynasty (1644-1911) scholar. In his book, Gu sought to cover the history and geography of all places in China in the late Ming Dynasty. We coded data from Gu Zuyu's work on the circumference of city wall and population of the associated jurisdiction for 1,182 cities. These data are used to confirm the positive correlation between the size of city wall and population, as suggested by our model.

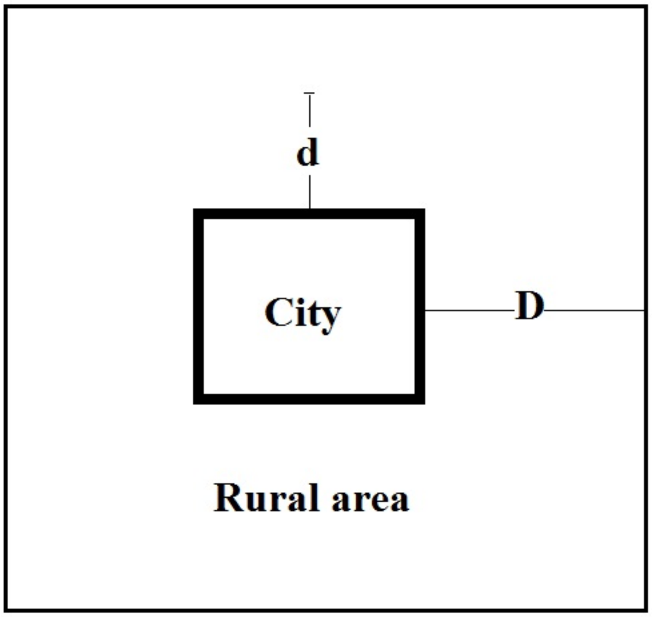
The second data set has been assembled by a group of researchers led by the anthropologist G. William Skinner (1925-2008). They hand-collected data on city walls for the late Qing Dynasty from more than 900 published gazetteers. Their data contain information on various dimensions of city walls for more than 1,600 cities. Using these data, we show that cities facing higher probabilities of being attacked tended to have stronger walls, as predicted by our model. We also report estimation results for city wall sizes in the Qing Dynasty in relation to such geographic fundamentals as quality and properties of the soil, terrain ruggedness, distance from Beijing and from the Silk Road, and broad geographic characteristics of their sites.

Since city walls were built to protect urban residents and properties from outside attacks, they served as physical and to some extent economic boundaries of cities. Therefore, the land area inside a city wall is a natural proxy for the size of the city. Using data from both Qing and Ming Dynasties, we examine whether city size distribution follows a power law as suggested by our model. For both periods, we find evidence that above a certain size cutoff the physical size of walled cities indeed follows a power law. This finding, combined with recent contributions by Dittmar (2011) and Desmet and Rappaport (2013), provides a deep historical dimension to the empirical regularities of city size fundamentals and the city size distribution.[4] It not only helps us better understand walled cities, but also highlights important implications for a fuller understanding of the historical dimensions of cities.

---

[4] Economists have recently devoted a considerable amount of research effort to this topic on city size distribution. There have been some serious attempts on the theory side to provide a microeconomic foundation for the size distribution of cities. See, e.g., Krugman (1996), Gabaix (1999), Eeckhout (2004), Duranton (2006), Córdoba (2008), Hsu (2012), and Lee and Li (2013). On the empirical side, economists have assembled a lot of data to characterize the distribution of city sizes. See, e.g., Dobkins and Ioannides (2000), Black and Henderson (2003), Ioannides and Overman (2003), Eeckhout (2004), Soo (2005), Rossi-Hansberg and Wright (2007), Holmes and Lee (2009), Rozenfeld *et al.* (2011), and Ioannides and Skouras (2013). Gabaix and Ioannides (2004) provide an exhaustive survey of the earlier empirical and theoretical literature, and Ioannides (2013, Chapter 8) of the more recent one. Among the empirical studies by economists, to the best of our knowledge, Rozenfeld *et al.* (2011) is the only one that examines the size distribution of urban land areas (in addition to city population). They find that for both the U.K. and the U.S., the distribution of city areas follows Zipf's (1949) law.

Figure 2: A square-shaped city and surrounding rural area



Finally, we report empirical results using contemporary data for 288 prefectural level cities for which presence or not of walls in the Qing Dynasty is known. We show that a higher density of economic activity, measured by population and employment density in 1984 and 2013, is associated with existence of walls in the past. These results complement the recent literature on the persistent effects on cities of man-made amenities even after they have been removed.[5]

2 Model

Consider a square-shaped city surrounded by rural area (Figure 2).

Urban production

Inside the city, N workers live with a density of 1 and are employed in the production of a homogenous manufactured good (say, clothing) according to the following production function:

$$Y_c = RN^{\alpha}, \quad 0 < \alpha < 1,$$

where Y_c is the total output of clothing and R is a production amenity, such as river size. It is assumed that R follows a power law, which is supported by evidence from both the U.S. and China.[6] Land is not used in manufacturing production, but each worker demands inelastically one unit of residential land in the city. One possible interpretation is that

---

[5] See, e.g., Bleakley and Lin (2012), Brooks and Lutz (1999), Hornbeck and Keniston (2014), and Siodla (2015).

[6] Krugman (1996) plots the log flow size of the 25 largest rivers in the United States against their log rank and finds a strong linear relationship, suggesting a power law distribution. We conduct a similar analysis

production occurs at the household level. We will use clothing as the numeraire good: $P_c = 1$, inside the city. The cost and freight price of clothing varies by location outside the city, which is to be discussed in more detail below.

Workers are paid according to the value of their marginal product, so the urban wage rate is given by

$$W_u = P_c \alpha R N^{\alpha-1} = \alpha R N^{\alpha-1}.$$

Total "profit" in the urban sector accrues to the local government. It is equal to:

$$P_c Y_c - W_u N = R N^{\alpha} - \alpha R N^{\alpha-1} N = (1-\alpha) R N^{\alpha}.$$

Rural production

Farmers live outside of the city in the rural area. They are uniformly distributed with each farmer working with λ units of land. We assume that $\lambda \gg 1$, i.e., population density is much lower in the rural area than in the urban area. Using λ units of land, each farmer can produce x units of food with a fixed proportions technology. Through a share-cropping contract, the landlord will pay the farmer θx and keep $(1-\theta)x$, where $0 < \theta \leq 1$.[7] As we will see shortly, for spatial equilibrium θ must be a function of the distance to the city. That is, the landlord only needs to pay a farmer sufficiently so that he would be indifferent between staying in the rural sector and moving to the urban sector.

Utility

Individuals, whether workers or farmers, have identical utility functions, given by $U = AF^{\beta}C^{1-\beta}$, where F is the quantity of food; C is the quantity of clothing; $0 < \beta < 1$ is a fixed parameter; and $A \equiv \beta^{-\beta}(1-\beta)^{-(1-\beta)}$ is a scaling constant.[8] This implies that a person with wage W has indirect utility

$$V = \beta^{-\beta}(1-\beta)^{-(1-\beta)} \left(\frac{\beta W}{P_F}\right)^{\beta} \left(\frac{(1-\beta)W}{P_C}\right)^{1-\beta} = W P_F^{-\beta} P_C^{-(1-\beta)}.$$

Transport costs

Some of the food produced in the rural area is shipped to the city in exchange for clothing produced in the city. There are no shipping costs within the city boundary. Outside the city, there is an iceberg shipping cost if goods are moved perpendicularly to the city edge;

---

using the 25 largest rivers in China, for which the log-rank-log-size regression gives

$$\underset{(t=15.39)}{\text{log\_rank} = 10.72 - 1.07 \cdot \text{log\_size},} \qquad R^2 = 0.91$$

also suggesting a power law distribution.

[7] To simplify the analysis, we assume that rural land is owned by an absent landlord who lives in the city or by the local government in the city.

[8] For simplicity, we ignore housing by assuming that in both urban and rural areas, one unit of housing of the same quality is provided to each individual by the local government. Since λ is assumed to be much greater than 1, this does not affect rural production.

there is no shipping cost if goods are moved parallel to the city edge.[9] In particular, if a good is sold for price P at the location of production, to offset the shipping cost its price will be $Pe^{\tau d}$ if it is moved over distance d perpendicularly to the city edge. Here $\tau > 0$ is a fixed parameter for both goods.

Spatial equilibrium

Let P_f be the food price inside the city, and recall $P_c = 1$, the price of clothing (the numeraire) inside the city. Then in the city a worker's indirect utility is

$$V_u = W_u P_f^{-\beta} = \alpha R N^{\alpha-1} P_f^{-\beta}. \tag{1}$$

Outside the city, the further away from the city edge, the more a landlord has to pay a farmer so that the farmer can attain the same level of utility as a worker in the city. Let D be the maximum distance from the city edge where it is feasible for farmers to trade with workers in the city (see Figure 2). At this distance, the landlord will have to pay the farmer all he has produced, x; that is, land rent is zero at the outer edge of the rural area. Food price at distance D is $P_f e^{-\tau D}$, so a farmer's income is $xP_f e^{-\tau D}$. Clothing price at distance D is $P_c e^{\tau D} = e^{\tau D}$. A farmer's utility at distance D is

$$V(D) = xP_f e^{-\tau D} \left(P_f e^{-\tau D}\right)^{-\beta} \left(e^{\tau D}\right)^{-(1-\beta)} = xP_f^{1-\beta} e^{-2(1-\beta)\tau D}.$$

Spatial equilibrium requires that $V(D) = V_u$: $xP_f^{1-\beta} e^{-2(1-\beta)\tau D} = \alpha R N^{\alpha-1} P_f^{-\beta}$. Thus, the price of food inside the city is:

$$P_f = \alpha R N^{\alpha-1} e^{2(1-\beta)\tau D} x^{-1}. \tag{2}$$

Similarly, suppose the landlord pays $\theta(d)x$ to the farmer at distance d, then the farmer's utility is

$$V(d) = \left[\theta(d) x P_f e^{-\tau d}\right] \left(P_f e^{-\tau d}\right)^{-\beta} \left(e^{\tau d}\right)^{-(1-\beta)} = \theta(d) x P_f^{1-\beta} e^{-2(1-\beta)\tau d}.$$

Spatial equilibrium requires that $V(d) = V(D)$: $\theta(d) x P_f^{1-\beta} e^{-2(1-\beta)\tau d} = x P_f^{1-\beta} e^{-2(1-\beta)\tau D}$, which implies that

$$\theta(d) = e^{-2\tau(1-\beta)(D-d)}.$$

Market equilibrium

At distance d, a farmer's income is $\theta(d) x P_f e^{-\tau d}$. At price $P_f e^{-\tau d}$, this farmer's demand for food is

$$\frac{\beta \theta(d) x P_f e^{-\tau d}}{P_f e^{-\tau d}} = \beta \theta(d) x.$$

---

[9] This assumption simplifies the calculation with a square-shaped city. The same assumption is typically made in circular city models, where it is assumed that shipping along an arc is costless, and similarly for linear but "thick" cities.

The farmer consumes a fraction β of the food output he receives as compensation, his real income in terms of food, and sells the remainder $(1-\beta)\theta(d)x$ on the market in exchange for clothing. This quantity of food from a single farmer at distance d will be sold for $(1-\beta)\theta(d)xP_f e^{-\tau d} = (1-\beta)xP_f e^{(2\beta-2)\tau D+(1-2\beta)\tau d}$. Combined with equation (2), this revenue from selling food is $(1-\beta)x\alpha RN^{\alpha-1}e^{2(1-\beta)\tau D}x^{-1}e^{(2\beta-2)\tau D+(1-2\beta)\tau d} = (1-\beta)\alpha RN^{\alpha-1}e^{(1-2\beta)\tau d}$.

Note that the revenue a farmer receives from selling food depends on distance d. Since the density of farmers at distance d is $\frac{4(\sqrt{N}+2d)}{\lambda}$, the total food revenue for all farmers is

$$\int_0^D \frac{4(\sqrt{N}+2t)}{\lambda}(1-\beta)\alpha RN^{\alpha-1}e^{(1-2\beta)\tau t}dt = \frac{4(1-\beta)\alpha RN^{\alpha-1}}{\lambda}\int_0^D (\sqrt{N}+2t)e^{(1-2\beta)\tau t}dt.$$

At equilibrium, this should equal the total expenditure on food by all the workers in the city. Remember that each worker's income is $W_u = \alpha RN^{\alpha-1}$. Utility maximization requires that a fraction β of the income be spent on food. Thus the total expenditure on food by N workers is $N\beta\alpha RN^{\alpha-1} = \beta\alpha RN^{\alpha}$. Food market equilibrium requires that farmers' revenue from food equals workers' expenditure on food, which can be simplified as

$$\frac{4(1-\beta)}{\beta\lambda N}\int_0^D (\sqrt{N}+2t)e^{(1-2\beta)\tau t}dt = 1. \tag{3}$$

Evaluating the integral gives the food market equilibrium condition as

$$\frac{4(1-\beta)}{\beta\lambda N}\left\{\frac{\sqrt{N}}{(1-2\beta)\tau}\left[e^{(1-2\beta)\tau D}-1\right] + \frac{2e^{(1-2\beta)\tau D}}{(1-2\beta)^2\tau^2}\left[(1-2\beta)\tau D-1\right] + \frac{2}{(1-2\beta)^2\tau^2}\right\} = 1. \tag{4}$$

This equilibrium condition defines a maximum distance, within which farmers trade with the city, as a function of city population and other parameters as the unique root of equation (4): $D(N,\lambda,\tau,\beta)$. We can confirm by total differentiation of (4) that $D_N(\cdot) > 0$ and $D_\lambda(\cdot) > 0$.

Equilibrium city population

Now assume that in the rural hinterland (far away from the outer edge of the rural area), there is an infinite supply of population who live in subsistence with reservation utility $\underline{V}$.[10] These people will move to the city as long as a worker's utility is higher than $\underline{V}$.[11] Thus in equilibrium, a worker's utility will be exactly $\underline{V}$. From equations (1) and (2), this implies that

$$\underline{V} = \alpha RN^{\alpha-1}\left[\alpha RN^{\alpha-1}e^{2(1-\beta)\tau D}x^{-1}\right]^{-\beta} = \left(\alpha RN^{\alpha-1}\right)^{1-\beta}x^{\beta}e^{-2(1-\beta)\beta\tau D}.$$

---

[10] According to Skinner's estimate, in the late Qing Dynasty, only 5.3-6.6% of the Chinese population lived in urban areas (Skinner 1977b, p.225).

[11] Throughout history, the Chinese people migrated across regions and between rural and urban areas. See a six-volume, systematic study on the history of internal migration in China by Ge et al. (1997). Government regulations related to the mobility of people varied substantially over different periods of time, ranging from facilitating internal migration in the Song Dynasty to restricting it in the early Ming Dynasty.

We rewrite this equation in log form and replace D with $D(N, \lambda, \tau, \beta)$ to get:

$$(1-\alpha)\ln N + 2\beta\tau D(N, \lambda, \tau, \beta) = \ln\alpha + \ln R + \frac{\beta}{1-\beta}\ln x - \frac{1}{1-\beta}\ln \underline{V}. \tag{5}$$

This closes the model by determining N and thus total urban and rural population, and allows us to solve implicitly for equilibrium city population $N^* = N\left(R, \alpha, x, \lambda, \tau, \underline{V}\right)$. Since $D_N(\cdot) > 0$ and $D_\lambda(\cdot) > 0$, it follows that $N_R > 0$, $N_\alpha > 0$, $N_x > 0$, $N_\lambda < 0$, and $N_{\underline{V}} < 0$. That is, population in the urban sector increases with both worker's and farmer's productivity; it decreases with the land-farmer ratio (i.e., land productivity) in food production and the reservation utility of potential urban-sector workers. All of these results make intuitive sense.

From equation (5), it readily follows that a power law distribution of R would lead to a power law distribution of N^* if $D(N, \lambda, \tau, \beta)$ were a function of $\ln N$. To explore this possibility, we consider a special case for which we can explicitly solve for $D(N, \lambda, \tau, \beta)$. In particular, let $\beta = \frac{1}{2}$. Then evaluating the integral in equation (3) gives

$$4D^2 + 4D\sqrt{N} - \lambda N = 0.$$

The positive root of this equation is:

$$D = \frac{-4\sqrt{N} + \sqrt{16N + 16\lambda N}}{8} = \frac{\sqrt{N}\left(\sqrt{1+\lambda} - 1\right)}{2}.$$

Thus equation (5) becomes

$$(1-\alpha)\ln N + \tau\sqrt{N}\left(\sqrt{1+\lambda} - 1\right)/2 = \ln\alpha + \ln R + \ln x - 2\ln\underline{V}. \tag{6}$$

In this special case, a power law distribution of R leads to a near-power-law distribution of N^* because $\sqrt{N}$ can be closely approximated by a linear function of $\ln N$ (see Technical Note, Appendix A). We have viewed R as river size, a natural production amenity. Actually, R can be anything that affects urban productivity. In particular, R can be interpreted as the accumulation of all past productivity shocks. If productivity growth rate is always a random draw from the same distribution, then R converges to a lognormal distribution, whose upper tail is hardly distinguishable from a power law.

Total surplus

We have assumed above that the total surplus in the urban sector, $(1-\alpha)RN^\alpha$, accrues to the local government in the city. We now examine the market value of the landlord's share of the food output. Both the surplus in urban and rural sectors are taxed away and thus do not enter the local market. We evaluate the surplus using the local market price. Although not explicitly modeled here, we assume that this total surplus is used by the central and local governments to support the public sector, provide public goods (e.g., construction of roads and city walls), and protect public safety.

Recall that at distance d, the landlord's share of the output is

$$1 - \theta(d) = 1 - e^{-2(1-\beta)\tau(D-d)}.$$

From each farmer at distance d, the landlord gets $x\left[1 - e^{-2(1-\beta)\tau(D-d)}\right]$. Its market value, measured using the local price $P_f e^{-\tau d}$, is $xP_f\left[e^{-\tau d} - e^{(2\beta-2)\tau D+(1-2\beta)\tau d}\right]$. Thus the value of the total food surplus is

$$\begin{aligned}
&\int_0^D xP_f\left[e^{-\tau t} - e^{(2\beta-2)\tau D+(1-2\beta)\tau t}\right]\frac{4\left(\sqrt{N}+2t\right)}{\lambda}dt \\
= \quad & \frac{4\alpha RN^{\alpha-1}}{\lambda}\int_0^D\left[e^{2\tau D(1-\beta)-\tau t} - e^{(1-2\beta)\tau t}\right]\left(\sqrt{N}+2t\right)dt.
\end{aligned}$$

Total surplus in the two sectors is

$$\Pi(N) = (1-\alpha)RN^{\alpha} + \frac{4\alpha RN^{\alpha-1}}{\lambda}\int_0^D\left[e^{2\tau D(1-\beta)-\tau t} - e^{(1-2\beta)\tau t}\right]\left(\sqrt{N}+2t\right)dt.$$

Since D is an increasing function of N, it is clear that $\Pi'(N) > 0$: the surplus increases with city size.

City wall

The presence of walls suggests threats. A city may be attacked (in a war, or by bandits) with probability γ. We assume that an attack causes a loss only to the city, because the city is more densely populated and all the surplus is stored in the city. A city wall will reduce the loss should an attack occur.

The circumference of the wall (or equivalently, the area inside the wall) is determined by N^*, the equilibrium city population defined in equation (5) above. Thus a power law distribution of N^* implies a power law distribution of the (circumference of or the area inside) city wall.

The quality of city wall, h (which we may think of as the height), and the size of city population N^* affect the loss when the city gets attacked. Specifically, we assume that defense technology is such that a city will only retain a fraction ϕ of its total surplus if an attack happens. We assume $0 < \phi(N^*, h) < 1$, $\phi_h > 0$, and $\phi_{hh} < 0$; that is, improving the quality of the city wall will make it more protective, but at a decreasing rate. Similarly, we assume that a larger city may be easier to protect: $\phi_{N^*} > 0$ and $\phi_{N^*N^*} < 0$.[12] Both the size and the quality of the city wall are costly. Specifically, we assume that the cost of maintaining a city wall is $c(N^*, h)$, where $c_h > 0$ and $c_{hh} > 0$, i.e., the marginal cost of quality is positive and increasing. Similarly, the marginal cost of protecting a larger city is also positive and increasing.

---

[12] This assumption is not inconsistent with the requirement that a minimum wall size is necessary to protect even one person, with defence being a classic public good even in this city context.

A social planner (who aptly, in the Chinese case, could be a government official, or the emperor) chooses the optimal quality of city wall to maximize the expected surplus:

$$\max_{h} \; \gamma\phi(N^*, h)\Pi(N^*) + (1-\gamma)\Pi(N^*) - c(N^*, h).$$

The first order condition, $\gamma\phi_h(N^*, h^*)\Pi(N^*) - c_h(N^*, h^*) = 0$, gives the optimal quality of city wall as a function of city size and the probability of being attacked:

$$h^* = h(\gamma, N^*).$$

It is straightforward to show that $h_\gamma = \frac{\phi_h \Pi}{c_{hh} - \gamma\phi_{hh}\Pi} > 0$. That is, conditional on urban population size (or area inside the city wall), the quality of city wall should be increasing in the probability of getting attacked.

Lee and Li (2013) propose a model in which equilibrium city size is determined by the product of a series of random factors including, for example, natural amenities and industry composition. They use a more general version of the central limit theorem to prove that equilibrium city size converges to a lognormal distribution. In a sense, their model is a cross-sectional counterpart of the random growth model; it simply allows all the random factors to influence city size contemporaneously. Imagine that when a local government official decided to build a new city wall in imperial China, he would likely consider all kinds of factors including the current population and its expected future growth, local food and water supply, political responsibilities of the local government, trade with other cities, etc. Lee and Li (2013) suggest that as long as these factors are random across cities and only weakly correlated, equilibrium city-size distribution would be asymptotically lognormal even if they are all built at exactly the same time.

On grounds of intuition, however, we claim that defense considerations provide an argument in favor of a lower bound to what would have been a lognormal distribution of city sizes. The cost of defending a single resident and thus maintaining a city wall contains a *fixed* component, therefore very small cities would be indefensible and thus infeasible. Such an assumption does not affect the result that $h_\gamma > 0$. For realizations of the random factors invoked by Lee and Li (2013) that would push city size downwards, this consideration acts as a “reflective barrier,” preventing cities from becoming too small. As Duranton and Puga (2014, p. 836) argue in the context of random urban growth, this has a dramatic effect in that it transforms a lognormal to a Pareto distribution, with a mode at the lower tail and an upper tail that would be fatter than that of a lognormal. That is, the lower bound eliminates the lognormal distribution’s thin lower tail and replaces it with a mode at the lower bound. Preventing cities from becoming too small requires that the upper tail accommodate more cities, thus fattening the upper tail. We appeal to intuition that the cross-sectional arguments employed by Lee and Li (2013) could also be suitably modified to accommodate a lower bound in a cross-sectional context. This, in turn, provides a *novel* justification for a Pareto distribution of city sizes, which is an apt explanation for sizes of walled Chinese cities.

Long Run effects of city walls

We next wonder whether existence of medieval city walls might have persistent effects on the structure of Chinese cities at the present time. A number of studies establish path dependence due to a variety of prior urban structures or land use, or possibly obsolete man-made amenities. Notably, they include: cities having grown out of portage sites (Bleakley and Lin 2012); long-abandoned street car routes now affecting land use in Los Angeles via their impact on zoning (Brooks and Lutz 2014); the effects of the 1906 San Francisco Fire facilitating greater density in burned relative to unburned areas (Siodla 2015); and, the effect of the 1872 Boston Fire creating opportunity for widespread reconstruction, possibly due to positive externalities from neighbors' reconstruction (Hornbeck and Keniston 2014).

The main model above imposes a population density of 1, within the city wall, and $\frac{1}{\lambda}$, in the rural hinterland. These assumptions are of course arbitrary, and it is straightforward to modify the model to allow for endogenous density. We briefly describe below how this can be done and discuss its implications.

In order to endogenize density, we need to introduce the consumption of land (housing) as an argument in the utility function. For full generality, land consumption should vary with location within the city. For spatial equilibrium, an individual's spending on land and transportation should be constant across all urban locations, leading to the so-called Mills–Muth condition: as distance from the CBD increases, the price of land must decline to compensate for increased transportation costs. The inverse of the land consumption function is just population density. In the absence of spatially dispersed amenities (*c.f.* Ioannides 2013, Ch. 5), the Mills-Muth condition implies that the land gradient should decrease with distance from the CBD, land consumption increase and population density decrease. Using the Mills-Muth condition along with demand for density and integrating over all urban locations yields an equation for total population as a function of the city's physical size, and therefore of the wall surrounding it. Depending on the city's geometry, that is circular or square perimeter, this translates into an equation for population as a function of wall size, N. In general, we would expect that population is increasing in wall size, which is indeed the case for standard parameter restrictions. By working accordingly, we may endogenize density and obtain a general relationship between urban population and wall size. In general, it would not just be a proportional one, unlike in the model above.

The canonical Mills-Muth model evolves around a CBD, with distance from which determining the land rental gradient. However, land use in medieval cities of China was subject to zoning rules: areas inside the city wall were usually divided into government quarters and residential neighborhoods, with commercial districts and open markets scattered in different locations (Zhang 2003). One can hardly identify a CBD in a typical walled city. We also know from studies of land use in such places as Moscow following the collapse of communist rule (Ioannides, *ibid.*) that ill-defined property rights together with administered land use would have implied incentives to the actors in medieval cities which could lead to land use outcomes that are very different than those in cities with well-defined land markets. Therefore, intuition suggests that a general relationship would likely hold between city wall sizes and populations across cities, as of the time when city walls were planned. Yet, additions and alterations over the years, plus natural "wear-and-tear" (which was particularly serious in China where city walls were typically built with rammed earth and bricks rather than the more durable rocks and stones used in many European countries), cause an attenuation of this relationship. However, it is reasonable to consider that city wall size defines a city's "capacity" to accommodate population, which is likely to work much like the man-made amenities discussed by the literature referred to in the beginning of this section. Thus, regardless of the specific institutions ruling land use at different historical epochs, the existence of walls may have a confining effect on individuals' location, which would likely interfere and possibly prevent urban sprawl. This when combined with modern notions of urban externalities would likely imply a positive effect of city wall on density of economic activity long after the walls have vanished.

Summary

We have presented a simple model to explain the existence of city walls. It has five implications that can be explored empirically:

1. *The size of the city, i.e., land area inside the city wall, is positively correlated with the size of (urban and rural) population.* The city wall is built to protect urban population, so its size is increasing with urban population. Given $D_N(N, \lambda, \tau, \beta) > 0$ and $D_\lambda(N, \lambda, \tau, \beta) > 0$, where λ is land per farmer, the land intensity of rural production, urban population increases with rural population. Thus the sum of urban and rural population is increasing with the area inside the city wall.

2. *The quality of a city wall is increasing in the probability of being attacked.* This follows directly from the model.

3. *A power law distribution of the productivity parameter R implies a distribution of city size close to a power law.* As pointed out above, this depends on how closely the function $D(N, \lambda, \tau, \beta)$ can be approximated by a linear function of $\ln N$. We have demonstrated that when $\beta = \frac{1}{2}$, a power law distribution of R gives a near-power-law distribution of city size.

4. *A fixed component of the cost of maintaining a city wall implies a lower bound in the distribution of city size. This in turn implies that a contemporaneous random factors theory of lognormal city sizes leads to a Pareto distribution for city sizes.*

5. *Presence of a city wall in history implies a higher density of economic activity in the present time.*

The remainder of the paper explores these implications empirically.

3 Data

3.1 Data on walled cities in Qing Dynasty (1644-1911)

The first data set used in this study comes from a long-term research project led by the late G. William Skinner.[13] Skinner was best known for his spatial approach to Chinese history. He collaborated with a network of researchers to create a large public database of historical Chinese social, economic, and political data at the county level.[14] One of their datasets, dubbed "ChinaW" where "W" refers to "walls", contains more than 150 variables measuring attributes for all cities, county seats, and *yamen*-level units recorded in China Proper (with Tibet and Outer Mongolia excluded) during the 19th century. It has detailed information on city walls, which we use for this study.

As Yue, Skinner, and Henderson (2007) explain, they first use two publications in the late Qing Dynasty to identify every administrative *yamen* at the prefectural and county levels and every territorial unit at the county level, which results in 2,402 units of observations. Some cities host more than one *yamen* at different levels of administration, which reduces the number of relevant observations to 1,869 for city-wall variables. Skinner's research team then use local gazetteers to find information on city walls in these places. See Appendix B for details.

In Table 1, we present some descriptive statistics on city walls. The key variable, area inside city wall, is calculated by Skinner's research team from the length of city wall assuming a typical square-shaped city. It has an average of 0.78 square kilometers. A few other variables are shown to help us envision the physical structure of a typical walled city in the late Qing Dynasty. The average city wall is 7.47 meters high and 7.26 meters thick at its base. It has 4 gates and 8 towers. 96 percent of cities also had moats surrounding their city walls. It is clear that a city wall of these features would forcefully impose a boundary that defines the physical size of the city.

Based on the estimated area inside city wall, the ten largest cities in the late Qing Dynasty were Nanjing, Suzhou, Beijing, Xi'an, Hangzhou, Yulin, Quanzhou, Hefei, Dingzhou, and Taiyuan. Four of the ten cities (Beijing, Nanjing, Xi'an, and Hangzhou) were capital cities during different dynasties. Others were well-known in the Chinese history for their economic or military significance.

---

[13] Prior to his death in 2008, Skinner was widely considered "the most eminent anthropological sinologist in the United States." For a biographical memoir of Skinner, see Hammel (2009).

[14] This database, including the dataset used here and many others, are all available for free download from the G. W. Skinner Data Archive website maintained at Harvard University: http://dvn.iq.harvard.edu/dvn/dv/hrs.

Table 1: Descriptive statistics for city walls

Variable name	Mean	Std. Dev.	No. of Obs.
Estimated area inside city wall (km^2)	0.78	3.93	1,623
Circumference of city wall (km)	2.76	2.23	1,623
Height of city wall (m)	7.47	2.88	1,467
Thickness of city wall (m)			
At base	7.26	4.03	309
At top	4.18	2.35	274
Unspecified base or top	5.30	3.01	820
Number of gates	4.24	1.34	1,599
Number of towers	8.19	9.64	1,125
Presence of moat	0.96	0.19	1,337
Distance to Beijing (10 km)	104.4	51.52	1622
Distance to Silk Road (10 km)	19.24	20.47	1622

Local gazetteers describe the dimensions of city walls using two traditional Chinese units of length, *li* and *zhang*. The Skinner research group recorded the data using these traditional units and then created separate variables to convert them into the metric units: 1 *li* = 0.5 kilometers and 1 *zhang* = 3.33 meters. Here we report the statistics using the metric units. Area inside city wall is calculated from the length of city wall using the formula $area = \left(\frac{length\ of\ city\ wall}{4}\right)^2$, i.e., assuming a square-shaped city.

3.2 Data on walled cities in Ming Dynasty (1368-1644)

The second data set contains information on the circumferences of city walls and jurisdiction population sizes. We hand-collected these data from a 130-chapter publication titled *Important Notes on Reading the Geography Treatises in the Histories* (*Du Shi Fang Yu Ji Yao*), written by the historical-geography scholar Gu Zuyu (1631-1692). See Appendix C for more details.

For our empirical analysis, we only collect data on jurisdiction population and the circumference of city wall.[15] The unit of population is *li*. The Ming dynasty organized households into different *li*'s. Each *li* had 110 households, which were divided into ten groups; each group had one household as the group leader and ten households as group members. The government created this community-level administrative system for collecting taxes, mobilizing service labor, and providing services such as education. In Ming dynasty, the average household had 5-7 people.[16] Thus one *li* had about 700 people. The unit of the circumference of city wall is *li* (the same Chinese character, but with a different meaning), which is about half of a kilometer. In Gu's book, this circumference is almost always rounded to a whole number: "over twelve *li*" or "close to four *li*." In these cases, we recorded

---

[15] We used the online version of Gu's book available here: http://www.guoxue123.com/biji/qing/dsfjy/. On a few occasions when we noticed possible typos in the online version, we double checked the text using the published version (Gu 2005).

[16] Calculations are based on different government publications in the Ming Dynasty. See Liang (2008, pp. 272-273).

Table 2: Circumference of city walls and population sizes in Ming Dynasty

	Circumference of city wall, *li*			Jurisdiction population, *li*		
	Mean	Std Dev	Obs.	Mean	Std Dev	Obs.
Empire capitals	82.0	19.8	2	882.0	256.0	2
Prefectural-level cities	10.2	7.05	167	350.1	507.5	189
Subprefectural-level cities	6.65	3.79	169	63.7	80.6	204
County-level cities	4.44	2.18	844	44.6	76.0	1,112
Total	5.70	5.17	1,182	86.6	219.7	1,507

As the unit of length, a *li* is half a kilometer. As the unit of population, a *li* has 110 households, or about 700 persons. There were two capital cities because the empire moved its capital from Nanjing (1368–1421) to Beijing (1421–1644). Jurisdiction population includes those living both inside and outside the city wall.

the whole number but indicated in our data file whether the number is rounded up or down.

Table 2 shows the average circumference of city wall and average jurisdiction population for cities by administrative level. Cities at higher levels tend to have longer city walls. Similarly, cities at higher levels tend to have larger jurisdiction population. This second fact is not surprising because by construction the population of a lower jurisdiction is included in the population of the higher jurisdiction. The two capital cities are outliers in terms of city-wall circumference, obviously because the emperors could use resources from the whole empire to build them, not only for protective purposes but also to symbolize the grandeur of the empire.

For 1,178 cities, we have both wall length and jurisdiction population. For those cities with missing city-wall data, there are four different types: (1) For 181 of them, Gu's book simply did not mention whether there was a city wall. (2) For 106 of them, the book clearly indicated that the city had no city wall (or had a wall before but it had collapsed). (3) For 24 of them, the government of the lower jurisdiction was located in the capital of a higher level government and did not have its own capital city. (4) For the rest, only 14 of them, the book indicates that the city did have a wall but did not provide any information on the circumference of the wall. Overall, we think that when the data on the circumference of city wall are missing, it is most likely that the city had no wall. In other words, Gu's book and thus our data seem to capture almost all the walled cities in the Ming Dynasty.

4 Results

4.1 The size of city wall and jurisdiction population

We first use the data from the Ming Dynasty to check whether the size of city wall is positively correlated with jurisdiction population. The results are in Table 3. Regressions in the first three columns use the full sample, except that the two capital cities are excluded as outliers. We try different specifications to allow for different possible nonlinearities. In

Table 3: City walls and population sizes in Ming Dynasty

	(1) DV: Wall length	(2) DV: Area inside wall	(3) DV: Log wall length	(4) DV: Log wall length
Jurisdiction population	0.0057***	0.0121***		
	(0.0012)	(0.0031)		
Log jurisdiction population			0.1505***	0.1398***
			(0.0228)	(0.0275)
Administrative level dummies	Yes	Yes	Yes	No
Province dummies	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes
Adjusted R^2	0.4442	0.2780	0.4506	0.2496
No. of observations	1,176	1,176	1,176	843

Regressions in the first three columns use the full sample but exclude two capital cities as outliers; the last column uses the sample of county level cities only. Area inside city wall is estimated as $\left(\frac{length\ of\ city\ wall}{4}\right)^2$, assuming a square-shaped city. Standard errors in parentheses are clustered by province. ***: $p < 0.01$.

the first column, we regress the length of city wall on jurisdiction population. We also regress the area inside city wall on jurisdiction population, which is in column 2. In column 3, we regress log length of city wall on log jurisdiction population. Since the population of a higher level jurisdiction is aggregated from population of lower level jurisdictions within its boundaries, one may be concerned with the regressions that treat jurisdictions at different levels as independent observations. Thus, in column 4 we also present the results from a log-log regression using county-level cities only. In all regressions, we control for province dummies. When using the full sample in columns 1-3, we also control for administrative level dummies.

Across different specifications, the size of city wall is always positively correlated with jurisdiction population. That is, when a larger population pays taxes to the local government, the local government tends to be located in a city with a longer city wall and thus it tends to have a larger urban area. This is consistent with the prediction of our model. Note that with fixed urban density, the coefficient should be 0.5 if we used city population in the log-log regressions. But we don't observe city population; instead, we are using jurisdiction population, the sum of urban and rural population. Given this, we are encouraged by the result that the estimated elasticities in columns 3 and 4 are of the same order of magnitude as 0.5, the elasticity of wall size with respect to city population size.

4.2 Wall size in Qing Dynasty and geographic fundamentals

Recall the predictions emanating from equation (5), namely that population in the urban sector increases with both worker's and farmer's productivity, that it decreases with land

Table 4: Wall size in Qing Dynasty and geographic fundamentals

	(1) DV: Wall length	(2) DV: Wall length	(3) DV: 1000*Log (wall length)	(4) DV: 1000*Log (wall length)
Length of nearby rivers and canals[a]	4.495*		1.768**	
	(2.651)		(0.753)	
Area of nearby inland water bodies[a]		0.212***		0.020
		(0.043)		(0.012)
Distance to Beijing	-3.925**	-3.702***	-0.910**	-0.888**
	(1.275)	(1.268)	(0.362)	(0.363)
Distance to Silk Road	-2.069	-2.416	-1.086	-1.087
	(2.728)	(2.711)	(0.775)	(0.776)
Wetland rice suitability index[b]	0.238***	0.174***	0.098***	0.094***
	(0.059)	(0.060)	(0.017)	(0.017)
Wheat suitability index[b]	0.004	0.004	0.0265*	0.0253*
	(0.051)	(0.051)	(0.015)	(0.015)
Terrain ruggedness index[b]	-123.5***	-121.3***	-37.92***	-40.00***
	(41.44)	(40.98)	(11.77)	(11.72)
South of Yangtze River	-562.5***	-544.0***	-306.2***	-310.3***
	(135.3)	(134.2)	(38.42)	(38.38)
Coastal city	331.7*	397.1**	149.5***	142.4**
	(200.8)	(199.0)	(57.00)	(56.93)
Administrative level dummies	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes
Adjusted R^2	0.287	0.296	0.271	0.269
No. of observations	1,622	1,622	1,622	1,622

Standard errors are in parentheses. *: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$. [a] Rivers, canals, and lakes within 10 km of the city's centroid are counted. [b] Average index within a 50 km circle around the city's centroid.

Data sources: Shape files for lakes, rivers, and canals are downloaded from the Harvard Library (http://dx.doi.org/10.7910/DVN/M7WEFY). Shape files for the Silk Road are downloaded from https://www.arcgis.com/home/item.html. The geodesic distance and river length variables are generated using ArcMap. Rice and wheat suitability indexes are from the Food and Agriculture Organization of the UN (http://www.fao.org/nr/gaez/). The terrain ruggedness index is from Diego Puga's data archive (http://diegopuga.org/data/rugged/).

productivity in food production and in the reservation utility of potential urban-sector workers. Results reported in Table 4 constitute broad tests of those predictions and of the effects of other determinants of wall length using the Qing Dynasty data. Specifically, we regress wall length and log wall length on a set of locational, geographic attributes of cities for which data are available and that we expect to be potential determinants in the light of the above discussion. The legend of Table 4 also reports the definitions of these variables.

Given that we interpret the worker productivity variable in our model as river size, we begin with two proxy measures for river size. The first one is the total length (based on polyline length) of all rivers and canals that pass within 10 kilometers of the walled city's centroid. The second is the total water area (based on polygon area) of all rivers, canals, and lakes that pass through or lie within 10 kilometers of the walled city's centroid.[17] Since these two measures are highly correlated, we include them in our regressions as alternative proxies. As shown in Table 4, coefficients of the river size measures are always positive and are significant in all but one specifications.[18]

Coefficients on other variables also turn out as expected. The distance to Beijing is negatively correlated with wall length. This might suggest a need to better protect urban areas closer to the capital, but more likely it reflects the fact that Beijing is on the edge of a large plain where the flat and fertile land could sustain bigger cities. Greater distance from Beijing connotes lower labor productivity. The city's distance to the Silk Road, a major medieval trade route that could proxy for a trade-related advantage, is negatively correlated with wall length but the estimates are not statistically significant. The wetland rice suitability index has a highly significant, positive coefficient in all regressions; the wheat suitability index also has a positive coefficient although it is (marginally) significant only in the log regressions; the terrain ruggedness index, due to Nunn and Puga (2012), quantifies topographic heterogeneity in wildlife habitats providing concealment for preys and lookout posts, has highly significant, negative coefficients in all regressions. These results suggest that larger cities tend to be situated in areas where the land is more fertile and easier to cultivate, which accords with intuition. The south of Yangtze River dummy has negative and highly significant coefficients, which reflects the fact that the southern part of China is more mountainous. With more geographic constraints, cities in the south of the Yangtze River were more difficult to grow, more costly to build, but easier to protect, and thus should need shorter defensive walls. Coastal cities tend to have longer walls, consistent with both higher productivity and greater need for protection.

---

[17] We first draw a 10 km radius circle around the walled city's centroid. If a river or a canal passes through this circle, its whole length is counted in the length measure and its whole surface area is counted in the area measure. Similarly, even if only part of a lake lies within the circle, its whole surface area is counted in the area measure.

[18] One might argue that the correlation between wall length and river size should be unconditional. Indeed, if we regress wall length on either of these river size measures without any controls, the positive coefficient is much more significant.

4.3 Quality of city walls

We next investigate the qualities of city wall using the Skinner data for the late Qing Dynasty. Our model suggests that controlling for city size, cities with a higher probability of being attacked by enemies would build walls of higher quality. We use three alternative quality measures: height of the wall, number of towers on the wall, and thickness of the wall. Whereas a higher and thicker city wall is stronger against attacks, towers provide a better view of enemies outside of the wall and make it easier to defend. The first two measures are straightforward and directly available from the Skinner data.

The thickness measure is more complicated. Since the average city wall in our sample is 7.5 meters high, the general design has a cross-sectional shape of a trapezoid so that it will not easily collapse. That is, the thickness of a city wall can be very different depending on where the measurement is taken. The Skinner data contains three different thickness variables: (1) thickness at the top of the city wall, available for 274 cities; (2) thickness at the base of the city wall, available for 309 cities; and (3) thickness at an unspecified position of the city wall, available for 820 cities. Overall, there are 934 cities with at least one of the three thickness variables available.

Using these variables, we construct two thickness measures for regression analysis. The first one takes the thickness at base if it is available; if not, it takes the thickness at top if available; if both are unavailable, it takes the thickness at the unspecified position if it is available. We simply call this variable the "thickness of city wall" and construct two dummy variables to indicate if it measures thickness at top and if it measures thickness at an unspecified position (instead of at base). The second thickness variable also takes the thickness at base if it is available. If unavailable, we will estimate it when possible. For cities with both thickness at top and at base, we regress base thickness on top thickness. We then use this estimated equation to compute the thickness at base for cities for which only thickness at top is available. Similarly, we use the thickness at an unspecified position to estimate the thickness at base for cities for which only thickness at an unspecified position is available. We call this second variable the "estimated thickness of city wall."

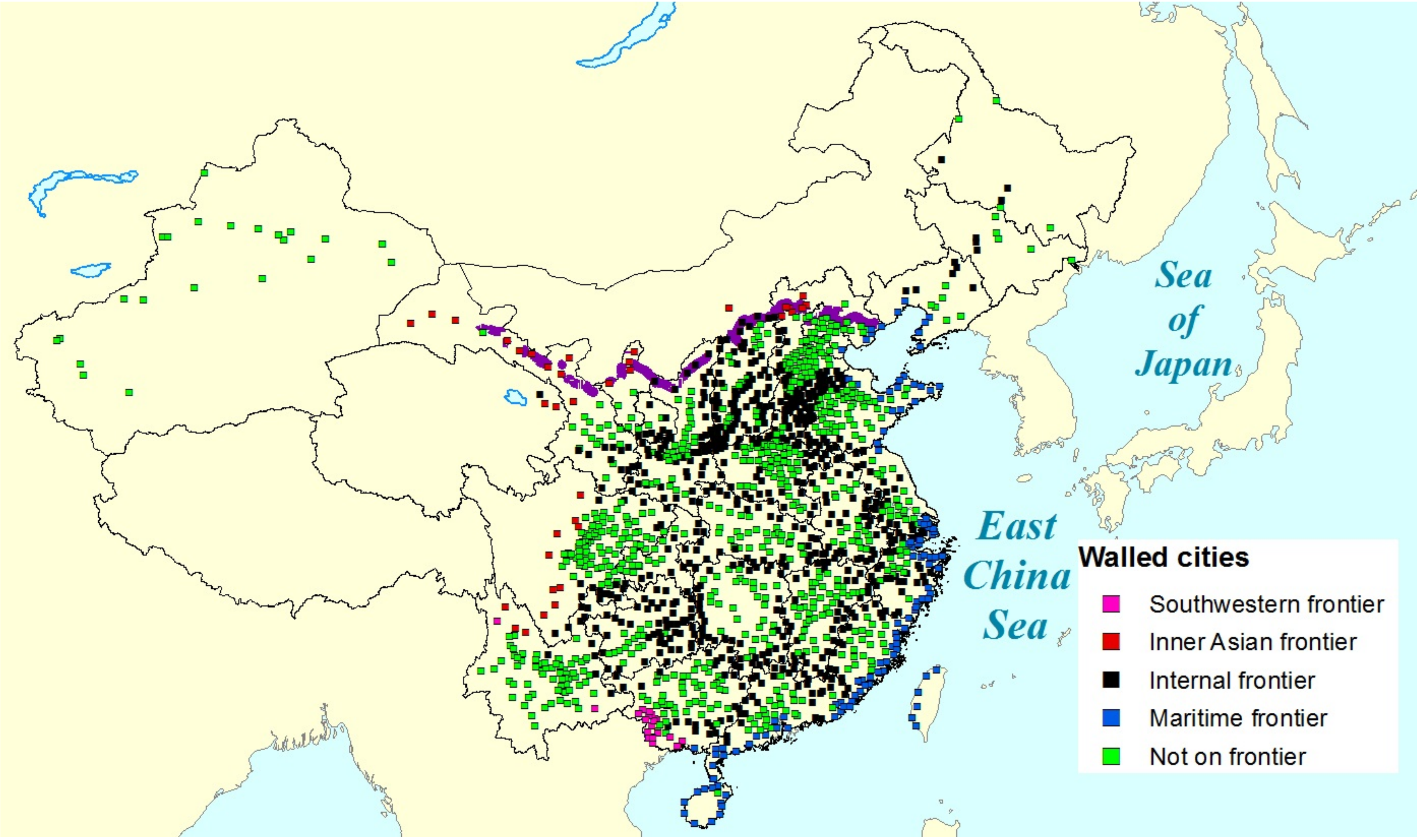
The ideal explanatory variable we need is the probability of being attacked for each city, which is not available. Instead, we will use each city's location relative to different frontiers to proxy the probability of being attacked. Skinner and associates carefully coded this information for each city in their database. Based on their categorization, we divided cities into five different groups:[19]

- On inner Asian frontiers: 38 cities;

---

[19] In the Skinner data, cities on internal frontiers are further divided into three subgroups: on macroregional frontiers only, on macroregional and provincial frontiers; and on provincial frontiers only. We combined all of them into one group of cities on internal frontiers. They also distinguished between cities on provincial and maritime frontiers and on maritime frontiers only. We combined them into a single group of cities on maritime frontiers.

Figure 3: Walled cities in Qing Dynasty and different frontiers



A total of 1,623 walled cities are shown on this map. The Ming Dynasty Great Wall is highlighted in purple. Black lines indicate the international and provincial borders of modern China.

- On southwestern frontiers: 26 cities;
- On maritime frontiers: 125 cities;
- On internal frontiers: 661 cities;
- Not on any frontiers: 733 cities.

Figure 3 shows all the 1,623 walled cities in the late Qing Dynasty, color coded according to their frontier types. Notice that cities close to international borders are not necessarily exposed to more attacks. For example, many of the cities in the southwestern Yunnan province were built on such rugged terrain that the mountains essentially protected them from enemies across the international border.

The first three groups are all close to the borders of China Proper. The inner Asian frontiers are along the borders of China Proper in the north, northwest, and west. These were traditionally the battlefields between Han Chinese and various belligerent ethnic groups, including the Tibetans, Xiongnu, Xianbei, Khitans, Tanguts, Jurchens, Mongols, and Manchus. Many of these minority peoples were nomadic or semi-nomadic; they frequently raided and pillaged the border regions. No wonder that the Great Wall was built on these frontiers and it was rebuilt over and over again throughout the history of imperial China. Cities on these frontiers faced the highest risks of confronting a strong and powerful enemy. In contrast, the southwestern frontiers were much less dangerous, partly because there was rugged terrain in these areas. There were only a few narrow passes in the mountains between China and regions on the Indochina peninsula, which were easy to defend. Throughout the history, relatively few battles were fought in these areas and countries on the Indochina peninsula were never a dangerous threat to China. The maritime frontiers used to be relatively safe too. However, between the 13th and 16th centuries, coastal areas of China were repeatedly invaded by pirates.[20] According to Gu's (2005) book, many cities in the coastal areas used to have no city walls. But after the pirates raided nearby villages and towns, the local governments started to build city walls for protection. In the Qing Dynasty, pirates were less a concern, but western countries started to invade China from the sea in the 19th century. Thus coastal cities in the Qing Dynasty still faced some risk of being attacked. For cities on internal frontiers, the risks come from domestic bandits, peasant rebellions, and regional military conflicts, which sometimes were as destructive as foreign invaders.

In Table 5, we regress city wall quality measures on dummy variables that indicate whether a city is on any of the frontiers. Cities not on any frontiers are used as the comparison group. For each regression, we control for administrative level dummies and the estimated area inside the city wall, i.e., the size of the city. For all four regressions,

---

[20] In history, these pirates were referred to as Wokou, meaning literally the "Japanese bandits." Recently, many scholars have come to the conclusion that the majority of Wokou were actually Han Chinese.

Table 5: Quality of city walls in frontier cities

	(1) DV: Height of city wall	(2) DV: Number of towers	(3) DV: Thickness of city wall	(4) DV: Est. thickness of city wall
On inner Asian frontiers	2.133***	1.625	1.849**	1.954***
	(0.592)	(1.204)	(0.677)	(0.664)
On southern frontiers	-3.185***	-2.626**	-2.959***	-3.250***
	(0.745)	(1.175)	(0.417)	(0.481)
On maritime frontiers	0.131	4.406***	-0.438	-0.432
	(0.489)	(1.332)	(0.455)	(0.445)
On internal frontiers	0.289	1.552**	-0.015	-0.030
	(0.378)	(0.593)	(0.320)	(0.320)
Area inside the wall	0.295***	1.271*	0.498***	0.516***
	(0.086)	(0.628)	(0.086)	(0.086)
Top of the wall			-2.719*	
			(1.520)	
Unspecified position of the wall			0.749	
			(1.294)	
Administrative level dummies	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes
Adjusted R^2	0.0844	0.1010	0.1712	0.1010
No. of observations	1,467	1,125	934	934

Cities on different frontiers are compared with cities “removed from any frontiers,” which is the excluded group. Area inside city wall is estimated as $\left(\frac{\textit{length of city wall}}{4}\right)^2$, assuming a square-shaped city. Standard errors in parentheses are clustered by province. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

area inside the city wall has a significant and positive coefficient. That is, larger cities have higher and thicker city walls with more towers, which is not surprising.

The first column examines the height of the city wall. It shows that city walls on inner Asian frontiers are on average 2.1 meters higher and that city walls on southwestern frontiers are on average 3.2 meters lower. City walls on maritime or internal frontiers are not significantly different in height. The second column shows that cities on southwestern frontiers have on average 2.6 fewer towers on the wall and that cities on maritime and internal frontiers have 4.4 and 1.6 more towers, respectively. Columns 3 and 4 investigate the thickness of the city walls, using two different measures. The results are similar: city walls on inner Asian frontiers are about 2 meters thicker and those on southwestern frontiers are about 3 meters thinner. City walls on maritime or internal frontiers are not significantly different in thickness.

Overall, we find that city walls on inner Asian frontiers are higher and thicker and that city walls on maritime and internal frontiers have more towers. We interpret these as evidence that cities facing higher risks built better city walls. City walls on southwestern frontiers are inferior in every respect: They are lower, thinner, and have fewer towers. We think this is because they were in mountainous areas and were unlikely to be attacked by enemies or bandits.[21]

4.4 Physical size distribution of walled cities

We now examine the physical size distribution of walled cities in both Qing and Ming Dynasties. Following the recent empirical literature, we focus on two questions in our analysis: (1) whether the physical size distribution of all walled cities is lognormal; and (2) whether the distribution of larger cities follows Zipf's law or Pareto law, more generally.

4.4.1 Accounting for measurement error

First of all, we explain why it is appropriate to work with the "land area inside city wall" estimated by assuming a perfect square. In principle, assuming the shape of the city to be a perfect square may over- or under-estimate the land area inside city wall. For example, if a city wall actually forms a circle, then treating it as a square will under-estimate the land area. In contrast, if the actual city shape is a rectangle, then assuming a square will over-estimate the land area. However, notice that in either case the sign of the estimating error does not depend on the actual size of the land area. Rather, the sign of the error is determined by the actual shape of the city (relative to a perfect square) and the magnitude of the error is proportional to the actual city size.

Let S be the true size (land area) of the city and $\hat{S}$ the estimated size, then one may

---

[21] The rugged terrain on southwestern frontiers also implies higher construction costs, which may also partly explain the lower quality of city walls there.

write

$$(1+\epsilon)\,S = \hat{S}, \tag{7}$$

where ϵ denotes the estimating error as a fraction of S. Let's assume ϵ is normally distributed, then $\ln(1+\epsilon) + \ln S = \ln \hat{S}$. Because ϵ is generally small, $\ln(1+\epsilon) \approx \epsilon$ and therefore $\epsilon + \ln S = \ln \hat{S}$. Suppose city size is lognormally distributed, then $\ln S$ is normal. Given the assumption of a normal ϵ, $\ln \hat{S} = \epsilon + \ln S$ should be normal. In other words, under the assumption of normally distributed ϵ, a lognormal $\hat{S}$ is a necessary and sufficient condition for a lognormal S. This is why it is informative to test whether $\hat{S}$ is lognormal.

What if S follows a Pareto distribution? In that case, its density function $f(S)$ and cumulative density function $F(S)$ can be written as:

$$\begin{aligned} f(S) &= \frac{b\underline{S}^b}{S^{b+1}} \qquad \forall S \geq \underline{S}; \\ F(S) &= 1 - \left(\frac{\underline{S}}{S}\right)^b \qquad \forall S \geq \underline{S}, \end{aligned}$$

where $\underline{S}$ is the smallest size and $b > 0$ a constant parameter. Zipf's law will be satisfied if data are drawn from a special case of the Pareto distribution with $b = 1$. Let $\mathcal{R}$ be the rank of a city with size S and I the number of cities in the sample (i.e., the rank of size $\underline{S}$), then in expectation

$$\mathcal{R} = I\,[1 - F(S)] = I\left(\frac{\underline{S}}{S}\right)^b.$$

Taking natural logs yields

$$\ln \mathcal{R} = a - b \ln S, \tag{8}$$

where $a \equiv \ln I + b \ln \underline{S}$ is a constant. Thus the common practice to test Zipf's law is to regress log rank on log size and check whether $b = 1$. A highly significant linear relationship with any $b > 0$ suggests a Pareto distribution of city size.

Suppose we do not observe S but instead $\hat{S} = (1+\epsilon)\,S$, where the proportional measurement error ϵ follows a normal distribution. Notice that in expectation, the rank of size S in the unobserved sample and the rank of size $\hat{S}$ in the observed sample should be the same.[22] Given that $\ln(1+\epsilon) \approx \epsilon$ for small ϵ, plug $S = (1+\epsilon)^{-1}\hat{S}$ into equation (8) to get

$$\ln \mathcal{R} = a' - b \ln \hat{S} + b(\epsilon - \bar{\epsilon}), \tag{9}$$

where $a' \equiv \ln I + b \ln \underline{S} + b\bar{\epsilon}$ is a constant and $b(\epsilon - \bar{\epsilon})$ is a normally distributed error with mean zero. That is, we may regress log rank on log size as in equation (9) using the mismeasured data on $\hat{S}$. As long as the measurement error is normal (as assumed here), the coefficient of $\ln \hat{S}$ in equation (9) will be the same as the the coefficient of $\ln S$ in equation (8), which can be used to test for Zipf's law.[23]

4.4.2 Size distribution of all walled cities

Cities in Qing Dynasty

Let us first plot the distribution of city sizes smoothed with a kernel, starting with the Qing Dynasty. Panel (a) of Figure 4 shows the density of city sizes in the Qing Dynasty using the full sample. As expected, there are few very large cities; most cities are rather small. Starting from the right end of the distribution, the density hardly increases as city size decreases. It takes a sharp turn and starts to rise quickly once moving below a certain size cutoff. But this does not continue all the way to zero; after another cutoff, the density loses its momentum and starts to fall. One important feature to notice is that, if we ignore the lower end of the distribution on the left side of the mode, the rest of the distribution indeed looks like the density of a Pareto distribution, which includes as a special case Zipf's law.

Panel (b) of Figure 4 shows the density of log city sizes using the full sample (the solid line). It looks like a normal distribution in that the density function is symmetric and bell-shaped. For comparison purpose, we add to Panel (b) the density of a normal distribution

---

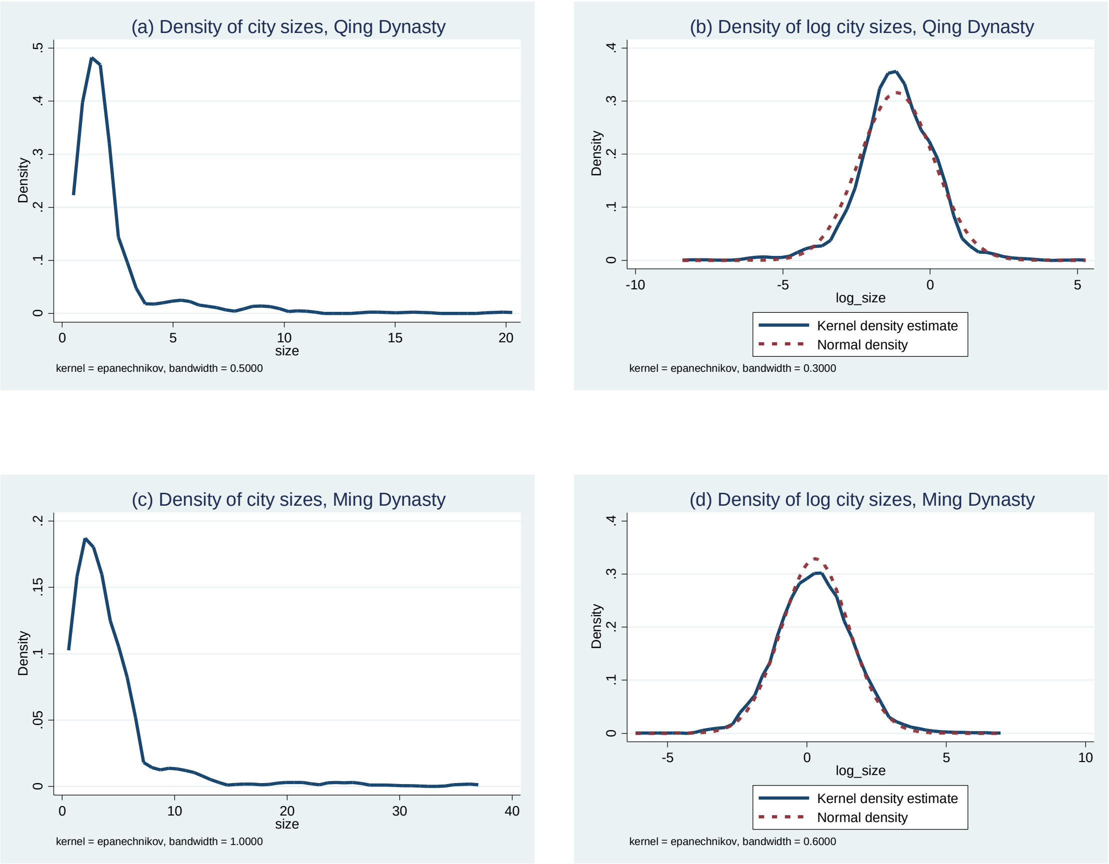
[22] As a heuristic example, consider a simple case with two cities only: i and j. Their sizes are measured with error: $\hat{S}_i = S_i(1+\epsilon_i)$ and $\hat{S}_j = S_j(1+\epsilon_j)$ are observed, where ϵ_i and ϵ_j are assumed to be independent draws from a normal distribution. The expected true rank of city i is $E[R_i] = 1 \times \text{Prob}[(S_i > S_j] + 2 \times \text{Prob}[(S_i \leq S_j]$. Let $\hat{R}_i$ be the rank of the observed size of city i. If $S_i > S_j$, then $E[\hat{S}_i] > E[\hat{S}_j]$ because of the independence assumptions and thus the rank order is preserved in expectation: $E[\hat{R}_i | S_i > S_j] = E[R_i | S_i > S_j] = 1$. Similarly, if $S_i \leq S_j$, we have $E[\hat{R}_j | S_i \leq S_j] = E[R_j | S_i \leq S_j] = 2$. Therefore,

$$\begin{aligned} E[\hat{R}_i] &= E[\hat{R}_i | S_i > S_j] \times \text{Prob}(S_i > S_j) + E[\hat{R}_i | S_i \leq S_j] \times \text{Prob}(S_i \leq S_j) \\ &= 1 \times \text{Prob}(S_i > S_j) + 2 \times \text{Prob}(S_i \leq S_j) = E[R_i]. \end{aligned}$$

When there are more than two cities, this logic applies to any arbitrary pair of two cities, thus in expectation the actual rank of city i and the rank of the observed city i are the same.

[23] There is an alternative way to justify our practice of running the Zipf regression using data with measurement errors. Notice that the proportional measurement error is equivalent to a growth rate. So the observed city sizes can be thought of as the actual city sizes that have experienced another round of random growth shocks. If indeed the actual city size follows the Zipf distribution, then we still have a Zipf distribution after moving forward for one more period because Zipf is a steady state under i.i.d. random growth. This is true even if our measurement errors and the random growth rates are drawn from different distributions (Gabaix 1999, p. 743). Thus the Zipf regression using the mis-measured city size data is informative about the distribution of actual city sizes. For this insightful observation, we thank the editor Gilles Duranton.

Figure 4: Density of city sizes in Qing and Ming Dynasties



For Qing Dynasty: sample size $I = 1,623$; unit of city size is square kilometer.
For Ming Dynasty: sample size $I = 1,182$; unit of city size is square *li*.

with the same mean and standard deviation (the dotted line). The two density functions resemble each other, although there are some discernible deviations especially around the mode of the density.

We then conduct two formal tests to check whether log city size follows a normal distribution. The first one is a one-sample Kolmogorov-Smirnov test; its test statistic is 0.071 with a p-value less than 0.001. The second one is a Skewness/Kurtosis test for normality, which yields a p-value less than 0.0001.[24] Both tests reject that the distribution of log city size is normal. Thus the evidence suggests that the physical size distribution of Chinese cities in the late Qing Dynasty does not follow a log normal distribution.

Despite the rejection of a lognormal distribution in these formal tests, it is rather remarkable how closely the density of log city size in our sample visually resembles the normal distribution. There is still a possibility that the actual city size S is indeed lognormal, and the estimated city size $\hat{S}$ is significantly different from lognormal only because the estimating error ϵ is far from normal. Without more data, there is no way to assess the likelihood of this possibility.

Cities in Ming Dynasty

We visualize the size distribution of Ming Dynasty cities in panels (c)-(d) of Figure 4. Size and log size densities are very similar to those plotted for the Qing Dynasty. Although the sample size of walled cities increased from 1,182 to 1,623 from the Ming to the Qing Dynasty, the overall city size distribution appears to be stable. The density of log city sizes looks remarkably close to a normal distribution. We again performed the one-sample Kolmogorov-Smirnov test and the Skewness/Kurtosis test, both again rejected the normality of log-size distribution.

4.4.3 Size distribution of larger walled cities

Cities in Qing Dynasty

We next check whether the physical sizes of larger cities obey Zipf's law (as often found to be the case in the literature). We first plot $\ln \mathcal{R}$ against $\ln \hat{S}$ using the full sample of walled cities in the Qing Dynasty; see panel (a) of Figure 5. The right portion of the plot indeed appears to be a decreasing linear relationship. The left tail is rather flat. There is clearly a sharp break in the slope. We therefore decide to locate the break first and then test whether the larger city sizes follow Zipf's law.

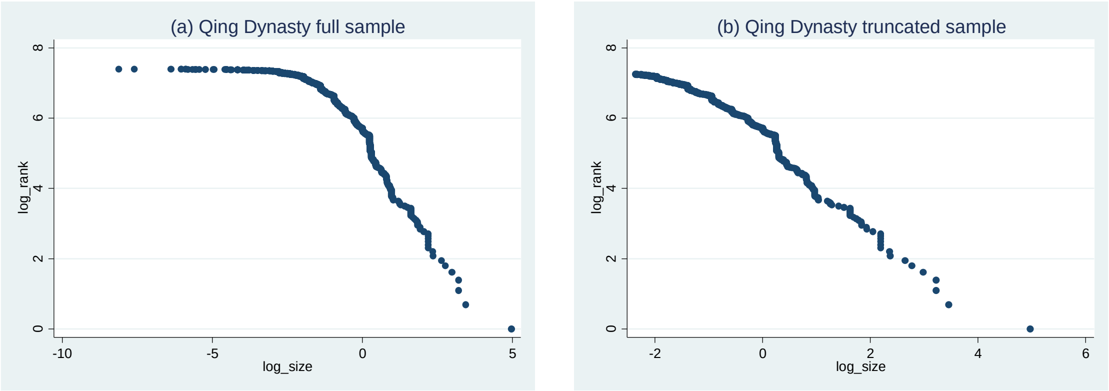
To identify the break in the slope, we run the following regression:

$$\ln \mathcal{R} = a - b_1 \ln \hat{S} - b_2 \left(1_{\mathcal{R} > \mathcal{R}^*} \ln \hat{S}\right) + e, \tag{10}$$

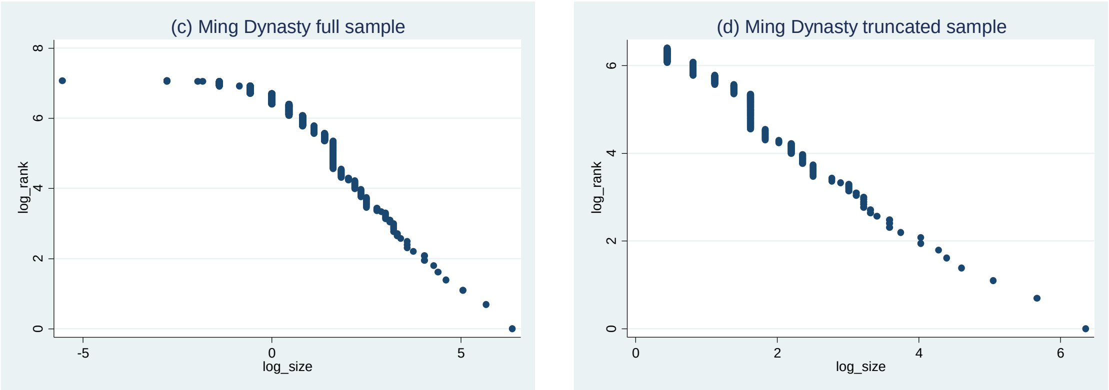
---

[24] Log city size in our data sample has a skewness of -0.384 and a kurtosis of 5.385, compared to a normal distribution's theoretical skewness of 0 and kurtosis of 3.

Figure 5: Log rank against log size for walled cities in Qing and Ming Dynasties


(a) Qing Dynasty full sample

(b) Qing Dynasty truncated sample


(c) Ming Dynasty full sample

(d) Ming Dynasty truncated sample

Qing Dynasty sample sizes: I =1,623 for panel (a) and 1,409 for panel (b).
Ming Dynasty sample sizes: I =1,182 for panel (c) and 602 for panel (d).

Table 6: Regressions of log rank on log size
Dependent variable: $\ln(\mathcal{R} - 1/2)$

A. Qing Dynasty regressions

	(1) Rank $\leq$1,409	(2) Rank $\leq$1,200	(3) Rank $\leq$1,000	(4) Rank $\leq$800	(5) Rank $\leq$600	(6) Rank $\leq$400
Constant	5.417***	5.408***	5.417***	5.447***	5.489***	5.559***
	(0.009)	(0.007)	(0.006)	(0.006)	(0.006)	(0.009)
Log size ($\ln \hat{S}$)	-1.015***	-1.116***	-1.200***	-1.295***	-1.367***	-1.456***
	(0.038)	(0.046)	(0.054)	(0.065)	(0.079)	(0.103)
R^2	0.939	0.957	0.966	0.975	0.975	0.975
No. of Obs.	1,049	1,200	1,000	800	600	400

B. Ming Dynasty regressions

	(1) Rank $\leq$1,000	(2) Rank $\leq$800	(3) Rank $\leq$602	(4) Rank $\leq$400	(5) Rank $\leq$200
Constant	6.534***	6.712***	6.889***	7.032***	6.913***
	(0.009)	(0.010)	(0.013)	(0.023)	(0.041)
Log size ($\ln \hat{S}$)	-0.985***	-1.109***	-1.209***	-1.279***	-1.249***
	(0.044)	(0.055)	(0.070)	(0.090)	(0.125)
R^2	0.939	0.960	0.967	0.964	0.958
No. of Obs.	1,000	800	602	400	200

Following Gabaix and Ibragimov (2011), we subtract 1/2 from the rank to account for the potential bias in small samples and calculate the standard error of the estimated slope $\hat{b}$ as $\hat{b}\sqrt{\frac{2}{I}}$, where I is the number of observations. ***: $p < 0.01$.

where a is a constant, b_1 and b_2 are coefficients, and e is the error term. $1_{\mathcal{R}>\mathcal{R}^*}$ is an indicator function that takes value 1 if rank $\mathcal{R}$ is greater than a particular rank $\mathcal{R}^*$ and value 0 otherwise. That is, this regression allows the coefficient of $\ln \hat{S}$ to be different above and below rank $\mathcal{R}^*$. We search for the break point between the 40th percentile and 90th percentile of the sample size 1,623, i.e., between ranks 650 and 1,460. For each rank $\mathcal{R}^* \in [650, 1460]$, we run the above regression. The $\mathcal{R}^*$ that gives the highest R^2 in the regression is considered the location of the break. This procedure identifies rank 1,409 as the break point. At this rank, city size is 0.09 square kilometers, with a city wall that is 1.2 kilometers long. Panel (b) of Figure 5 plots log rank against log size using the truncated sample of 1,409 observations. The negative linear relationship is obvious. We then regress log rank on log size using this sample of 1,409 larger cities; the results are in column (1) of panel A in Table 6. The coefficient of $\ln \hat{S}$ is -1.015, remarkably close to -1, suggesting that Zipf's law holds for this truncated sample of 1,409 cities.

Eeckhout (2004) proves that if the underlying distribution is lognormal, then the magnitude of the log-size coefficient in the log-rank-log-size regression should be increasing as one uses a smaller and smaller sample of the largest cities. Intuitively, the smaller the cutoff, the thinner the upper tail will be. This is empirically true in the 2000 U.S. census places data.

We conduct a similar analysis here with various cutoff points, regressing log rank on log size using samples of 1,200, 1,000, 800, 600, or 400 largest cities. The results are in columns (2)-(6). Comparing the coefficients of $\ln \hat{S}$ across different columns, we see that indeed the coefficient is increasing in absolute value as the sample size of largest cities decreases. This is consistent with Eeckhout's (2004) findings with 2000 U.S. census places data. Note that in all the regressions in panel A of Table 6, the R^2 is never lower than 0.939.[25] That is, the straight line always fits very well despite the varying slope.

Cities in Ming Dynasty

In panels (c)-(d) of Figure 5, we plot log rank against log city size for Ming Dynasty cities. First notice that there is a lot of round-number bunching in the data. As mentioned above, Gu Zuyu tended to use round numbers when recording the circumference of city walls. He frequently uses such language as "over eight *li*" or "close to five *li*." In these cases, one can do nothing but take the closest whole numbers as the approximate length, which is why there is so much bunching in the city size variable. The biggest vertical jump in the figure corresponds to the city-wall perimeter of nine *li*, which has a total of 116 observations.

Despite the data bunching problem, the overall plot exhibits similar properties as for the Qing Dynasty. The plot using the full sample, in panel (a), again shows a clear linear relationship for larger cities. There seem to be too few small cities, perhaps due to left censoring. That is, cities below a certain size cutoff are less likely to build a defensive wall, which makes perfect economic sense. We follow the same procedure to identify a break point in the log-rank-log-size plot and find it to be 602. Panel (b) shows the plot for 602 larger cities only, which gives a nearly perfect linear relationship.

We again regress log rank on log size, and the results are in panel B of Table 6. In addition to the 602 break point, we also tried samples of 1,000, 800, 400, and 200 largest cities for comparison purposes. There is still some concavity in the log rank–log size data since the absolute value of the log size coefficient tends to become bigger as we use fewer and fewer large cities. But overall, the coefficient varies within a smaller range around unity, between -0.985 and -1.279. That is, the rank-size distribution of walled cities in the Ming Dynasty is fairly close to Zipf's law.

4.4.4 Discussion

How to explain these results? Researchers have proposed different variations of Gibrat's law (growth rate is independent of size) to provide a foundation for observed city size distribu-

---

[25] To put this in perspective, we note that Eeckhout (2004, Table 3) reported regressions of log rank on log size for U.S. census places at different truncation points, with R^2's ranging from 0.860 to 0.997. When using subsamples from the upper tail (with 5000, 2000, or 135 largest places), he obtained R^2's higher than 0.98.

tions e.g., Gabaix (1999), Eeckhout (2004), Duranton (2006), Rossi-Hansberg and Wright (2007), Córdoba 2008), with the latest and most thorough contribution being Duranton and Puga (2014). Using any version of these "random growth models" to explain the physical size distribution of walled cities in late imperial China would face this problem: City walls last for hundreds of years and thus the proxy for city size used in this study rarely "grows" over a long period of time.[26] A random growth model may help explain the empirical findings here only if one believes that physical city sizes had long reached an equilibrium distribution by the Ming Dynasty.

Only a few static models have been proposed to explain city size distribution (Krugman 1996, Hsu 2012, Lee and Li 2013). Krugman (1996) suggests that the power law distribution of city sizes may simply reflect the "inhomogeneity" of the landscape on which cities emerged. Since the varying features of the landscape can generally be regarded as random, one could use this random variation to produce a power law. However, Krugman (1996) does not provide a full model to formalize this idea. As a concrete example of the "inhomogeneity" in nature, he shows that a plot of the log flow size of the 25 largest rivers in the United States against their log rank strongly suggests a power law distribution, with a coefficient of -0.949. In our case in this paper, the power law distribution of R (the parameter for urban productivity) may result from an accumulation of random productivity shocks over time, or simply reflects a distribution of natural advantages over space. Such a source of the power law distribution is consistent with both the random growth theory and the static theory in the spirit of Krugman (1996) or Lee and Li (2013).

We conclude this section by suggesting that future research could explore formal testing of more precise spatial aspects of the medieval urban system of China. Interestingly, G. William Skinner has argued that the spatial structure of cities in imperial China should be understood as an interaction between two "hierarchies of central places" — "one created and regulated by the imperial bureaucracy for purposes of field administration, the other given shape in the first instance by economic transactions" (Skinner 1977c, p.275). Skinner himself actually formulated a "central place theory" to model this spatial structure (Skinner 1977c). These ideas bring to mind Hsu (2012), which has yet to be directly tested empirically. We suspect that our data from the late imperial China may be useful in devising a test of "hierarchies of central places." This clearly deserves attention in future research.

4.5 Medieval walls and modern Chinese cities

We next examine whether medieval city walls have had persistent effects on the structure of Chinese cities today. In view of the discussion in section 2 above, we proceed to examine empirically whether the presence of medieval walls, even if they have been torn down (which

---

[26] This does not mean that city population was static in imperial China; it only implies that changes in population were mostly absorbed by varying population densities in walled cities.

Table 7: Density of economic activity and presence of walls in Qing Dynasty

	1984 Regressions		2013 Regressions	
	(1) DV: Log Population Density	(2) DV: Log Employment Density	(3) DV: Log Population Density	(4) DV: Log Employment Density
Presence of walls in Qing Dynasty	0.439***	0.342**	0.316***	0.381***
	(0.128)	(0.132)	(0.100)	(0.118)
Number of historical-cultural sites	0.00105	0.00518	-0.00730	0.00176
	(0.007)	(0.007)	(0.006)	(0.007)
Log GDP^a	-0.252*	-0.644***	0.235*	-0.042
	(0.135)	(0.139)	(0.124)	(0.145)
Log per capita GDP^a	0.999***	1.746***	-0.201*	0.561***
	(0.140)	(0.145)	(0.116)	(0.137)
Log number of industrial enterprises	0.625***	0.839***	0.164**	0.188**
	(0.135)	(0.140)	(0.077)	(0.091)
Employment share of service sector	4.463***	1.780*	-0.485	-1.934***
	(0.944)	(0.976)	(0.384)	(0.446)
Terrain ruggedness index^b	-0.0064*	-0.0097**	-0.0100***	-0.0100***
	(0.004)	(0.004)	(0.003)	(0.003)
Province capital dummy	-0.374	0.030	0.148	0.749***
	(0.240)	(0.248)	(0.212)	(0.247)
Seaport dummy	-0.0499	0.0445	0.126	0.299
	(0.209)	(0.216)	(0.156)	(0.187)
Constant	Yes	Yes	Yes	Yes
Adjusted R^2	0.487	0.634	0.329	0.507
No. of observations	288	288	284	275^c

Standard errors are in parentheses. *: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$. ^{a} For the 1984 regressions, "industrial output value" is used because at that time GDP statistics had not been adopted yet in China. ^{b} Average index within a 50 km circle around the city's centroid. ^{c} For a number of cities, the number of private sector employees is missing and thus total employment is unavailable.

Data sources: The number of historical-cultural sites is calculated using data from http://www.bjww.gov.cn/wbsj/zdwbdw.htm; other control variables are from the 1985 and 2014 editions of the *China Urban Statistical Yearbook* (CSSB 1985, 2014). The terrain ruggedness index is from Diego Puga's data archive (http://diegopuga.org/data/rugged/).

is indeed the case with most city walls in China) might affect the density of economic activity in Chinese cities at present. We thus examine whether there is a significant difference in density of employment or population between cities that had walls in the Qing Dynasty and those that did not. Our analysis sample here consists of prefectural level cities only, for which data are easily available from the annual publication of the *China Urban Statistical Yearbook*. Cities that did not have walls in the Qing Dynasty were insignificant places then; they have now become prefectural level cities because they emerged and expanded for various reasons over the last century.

Cities that did have walls have richer histories, where for a long time urban economic activities have found it inescapable to operate within the confines of those walls. Many areas inside the walls had established a tradition for certain types of uses. Urban planners tend to respect these traditions in order to maximize the value of urban land within walls and thus build densely in these city centers even if the walls are no longer standing up. This type of path dependence is demonstrated aptly by the example of Beijing. Within one mile of Tiananmen Square, Beijing's historical center, there are three highly popular, densely built shopping districts (Qianmen, Wangfujiang, and Xidan). From a modern city planner's perspective, this urban design may seem redundant. Indeed, the three shopping districts owe their existence today mainly to historical antecedents: by the time of the Qing Dynasty they were already shopping centers within the city wall. More generally, cities that had defensive walls tend to have an attractive traditional city center that holds back urban sprawl and generate positive feedbacks from increased density. That is, all else equal, an existing wall confines activities, which in turn makes it attractive for other activities to colocate, just as "crowded parties are the best!" In our discussion of long run effects of city walls in section 2, we hypothesize that such cities have higher population and employment densities today. This hypothesis is indeed confirmed using data for 1984 and 2013.

Table 7 reports results with two sets of regressions, for 1984 and 2013. In each period, we regress log population and employment densities on a city wall dummy and a set of city characteristics. In interpreting the results we ought to note that by the 1960s, almost all city walls had been torn down. The city wall dummy indicates whether a city had defensive walls in the Qing Dynasty. To separate the effect of a historical wall *per se* from that of a richer historical heritage, we include as a control variable the number of historical-cultural sites in the city today, but this variable is not statistically significant in any of the regressions. We also control for log GDP as a measure of city economic size. The respective coefficient is negative and significant for 1984, but positive and significant only for population density for 2013. We explain this as an indication of a changing role of urban economic activity during the industrialization of China. We include log per capita GDP in the respective year as an income control.[27] We find positive and statistically very significant effects in three

---

[27] In addition to this log per capita GDP as an income control, we have included log GDP in the regression as a city size control. An alternative city size control is log population; but because it is already in the dependent variable, we use log GDP instead. Consequently, if one wants to know the overall effect of log

of the four regressions, which confirms it as a determinant of population and employment density. We control for the structure of the local economy by including the log number of industrial enterprises and find significant and highly positive effects in all four regressions. The effect of the employment share of the service sector in the city has positive and very significant coefficients in 1984, but negative and in the case of log employment density in 2013 very significant coefficient. We interpret this result as indicating that as these cities have become quite large, they are also very diversified and their service sectors no longer playing significant roles. We also include the terrain ruggedness index in the regression, which as expected is negatively correlated with population and employment densities. And lastly, we control for whether a city is a province capital, using a capital dummy, and whether a city is a seaport, using a seaport dummy.

In all four regressions reported in Table 7, the city wall dummy has a positive and statistically very significant coefficient. That is, despite the fact that city walls were torn down many decades ago, cities that did have walls in history tend to have higher densities of economic activities today. This is a powerful demonstration of the lingering effect of a city wall as a man-made amenity that defines the historical boundary of a contemporary city's historical core and induces higher density.[28] The statistical fit is better for the employment density regressions, which accords with intuition because employment density is more likely to be sensitive to economic magnitudes than population density.[29]

5 Conclusion

Throughout the majority of human history and in different civilizations, cities are surrounded by defensive walls. However, city walls are not well understood from the economic perspective. The present paper offers a simple model to rationalize the existence of city walls. The model relates the sizes and qualities of city walls to a set of economic variables and geographic fundamentals, which provides a guide for empirical analysis of walled cities. Furthermore, specific features of the model, such as defense considerations yield a novel justification for the Pareto law of city sizes.

Our empirical work draws on two unique and previously unused (for economics research) data sources. The first one contains a wide range of characteristics of city walls in the Qing Dynasty, hand-collected by a group of researchers led by G. William Skinner. The second

---

GDP on density, one should recognize that log GDP is also in log per capita GDP and thus need to combine the two coefficients.

[28] An alternative interpretation of this effect is persistence of density. Existence of wall was more likely to confine economic activity and thus increase density back at the time it was built.

[29] One also wonders whether the physical size of medieval walled cities is correlated with their population size at present, if these cities still exist today. This turns out to be true: cities with longer walls in history tend to have larger populations today (this empirical result, although not presented in this paper, is available upon request). It implies that in some empirical exercises the length of city walls could serve as an instrumental variable for current city population in China. See Xing and Zhang (2015) for such an application.

one contains information on the circumference of city wall and jurisdiction population size in the Ming Dynasty, which we hand-collected from Gu Zuyu's book. Using these data, we have shown that the area inside the city wall is increasing in jurisdiction population, that the length of city wall is correlated with local economic fundamentals, and that measures of city wall quality are increasing in the risk of being attacked (proxied by the location of the city relative to different frontiers). All of these results are consistent with the predictions of our model.

We use these data to explore in greater depth the physical size distribution of walled cities. The existing literature on city size distributions focuses almost exclusively on population size. We draw attention to the fact that the land area inside the city wall is a natural measure of the physical size of the city. We show that the physical sizes of larger cities in both Qing and Ming Dynasties follow a Pareto distribution. Given that our analysis is concerned with a much earlier time period and based on a very different size measure, this empirical regularity of city size distribution appears to be even more robust than previously thought. Our findings suggest that the theorization of Pareto law in city size distribution needs to take into account its long history. Finally, we demonstrate that presence of city walls in medieval times has a lingering effect in population and employment density in Chinese cities at the present time.

Appendix

A. A technical note

We show that in equation (6), a power law distribution of R leads to a near-power-law distribution of N. Let $G_R(r)$ be the countercumulative distribution of R. If R is power-law distributed, then $G_R(r) = \underline{r} r^{-\zeta}$, where $\underline{r}, \zeta > 0$ are positive parameters. Let $\mathcal{N}(N)$ denote the function of N in the l.h.s. of (6), after it has been raised to the power of e:

$$\mathcal{N}(N) = N^{1-\alpha} \cdot \exp\left[\tau\sqrt{N}\left(\sqrt{1+\lambda}-1\right)/2\right].$$

Equation (6) may be rewritten as $\mathcal{N}(N) = \rho R$, where ρ is a function of all parameters in the r.h.s of (6) except R. Thus, we have:

$$\text{Prob}\left\{N \geq n\right\} = \text{Prob}\left\{\mathcal{N}^{-1}(\rho R) \geq n\right\} = \text{Prob}\left\{R \geq \rho^{-1}\mathcal{N}(n)\right\} = \underline{r}\left(\rho^{-1}\mathcal{N}(n)\right)^{-\zeta}.$$

The countercumulative distribution of N readily follows:

$$\text{Prob}\left\{N \geq n\right\} = \rho' n^{-\zeta(1-\alpha)} \cdot \exp\left[-\zeta\tau\sqrt{n}\left(\sqrt{1+\lambda}-1\right)/2\right],$$

where ρ' is a function of parameters. The deviation from the power law is clear. Numerical results with the last factor above show that it is important for small values of N. For large values of N, the factor is well approximated by a power function of N. The argument that properties of the landscape are transmitted to city size is presented by Krugman (1996) and adopted by Gabaix and Ioannides (2004).

B. Data on walled cities in Qing Dynasty

There is a long tradition in China that local governments publish gazetteers to document the history, geography, culture, and outstanding individuals in their local regions. The first gazetteer appeared in the Jin Dynasty (265-420). By the Ming Dynasty, gazetteers were so common that "for a county or monastery not to have a gazetteer was regarded as evidence that the place was inconsequential" (Brook 1997, p.237). A survey in 1976 revealed that more than 8,000 gazetteers survived in China; many places had multiple editions published at different points in history. One of the most commonly documented facts in a gazetteer is the physical structure of cities, which is why gazetteers are useful for collecting information on city walls.

Table A.1: Publication dates and types of gazetteers consulted when preparing city wall data

	Publication dates						
Types	1519-1599	1600-1699	1700-1799	1800-1899	1900-1974	Unknown	Total
Empire-wide				3			3
Provincial-level			5	14	7		26
Prefectural-level	1	1	10	15	2		29
County-level	10	28	136	347	339	3	863
Miscellaneous			1		7	2	10
Total	11	29	152	379	355	5	931

To construct the ChinaW dataset, Skinner and collaborators consulted a total of 931 gazetteers published during 1519-1974. Table A.1 shows the types and publication dates of these gazetteers. The bulk of these publications (93 percent) are county-level gazetteers, which were usually written by leading local intellectuals who had access to accurate information about the local region. Seventy-nine percent of these gazetteers were published after 1800, meaning that the information on most city walls was up to date in the late Qing Dynasty. Given that city walls are stable structures that often last several hundred years during peaceful times, it is perhaps true that even the information published a little earlier (e.g., in the 1700s) still accurately reflects the situation in the late 1800s.[30]

---

[30] Skinner (1977a) uses Suzhou as an example to illustrate the fact that walled cities had stable physical forms. He compares a map of Suzhou engraved on a stone in 1229 with an aerial photograph of the city taken in 1945 and finds that walls, moats, streets, and canals on the two maps are almost identical despite drastic population fluctuations in the city over that period.

Among the 1,869 relevant observations, there are 224 places for which the Skinner research team could not decide whether they had had city walls or not. Our own guess is that most of these places had no city walls at all, which is why no information about city walls could be found in historical records. Ninety-four percent (210 out of 224) of these places were county-level units. That is, even if they had city walls, they must have been rather small and would be at the lower tail of the city size distribution. For the rest of the 1,645 observations, it is known that they did have city walls.

Note that local gazetteers usually do not directly mention the land area inside a city wall. However, they almost always give the dimensions of the city wall. Most gazetteers specify the circumference of the city wall; others give the length of each section of the city wall from which the circumference can be calculated. Indeed, the circumference of the city wall is the most complete variable among all the city-wall attributes recorded in the Skinner data. Among the 1,645 cities that are known to have had city walls, the circumference variable is available for 1,623 cities; this variable is missing for only 22 cities. Twenty-one out of these 22 cases were county seats, and thus they were likely to be small cities. For this reason, our empirical analysis focuses on the sample of 1,623 cities with the city-wall circumference variable available and uses this information to estimate the land area inside city wall. It seems reasonable to believe that this sample contains almost all cities that had city walls in the late Qing Dynasty. The few observations with missing city-wall circumferences are most likely to be very small cities and thus only affect the distribution of city sizes at the lower end.

Before using the Skinner data to conduct empirical analysis, it is important to verify that the information on city-wall circumferences is reliable. As a precautionary check, we arbitrarily chose four prefectural-level gazetteers that Skinner's team of researchers used as data sources, including those for Dingzhou, Guangping Fu, Hangzhou Fu, and Tianjin Fu which were published in 1849, 1894, 1922, and 1899 respectively. We read these four gazetteers and were able to find information on 27 walled cities, for all of which the city-wall circumference variable was available. In every single case, the information we found in the gazetteers agrees with the value recorded in the Skinner data (in a few cases, up to a rounding error). Thus, the city-wall circumference information in the Skinner data is very accurate.

How to estimate the land area inside city wall is a tricky issue, especially that for most cities the shape of the walled area is unknown. Early historical records indicate that many ancient cities were square-shaped. An ancient Chinese book on science and technology, *The Records of Examination of Craftsman* (*Kao Gong Ji*), described the monarchy's central city as a perfect square. This book was later (in the Han Dynasty, 202 BC – 220 AD) included in a Confucius classic and became a must-read among Chinese intellectuals for two thousand years. It had an important impact on the design of cities in Chinese history, because the book made people believe that an ideal city should be square-shaped. According to Zhang

(2003, p. 293), more than 70 percent of Chinese cities had square-shaped city walls. In northern China, where flat land was abundant, city walls were almost always designed to form a square or a rectangle close to a square. Departures from rectangularity might take the form of one or two curving sides (usually along a river) or a truncated corner. In the south, where city walls were often built on rugged terrains, many cities had to deviate from the ideal and ended up with irregular shapes (Chang 1977). The Skinner data include an "estimated intramural area" variable, which equals the square of one quarter of the city-wall circumference. That is, the estimate simply assumes that every city was a perfect square. Without reliable information on exact city shapes, there is no obviously better way to estimate the land area inside each city wall. Thus we use this estimate as the city size.[31] In Section 4.4, we explicitly specify the conditions under which we may use this inaccurate estimate to draw inferences about the actual city-size distribution.

C. Data on walled cities in Ming Dynasty

Data on Ming Dynasty cities are collected from Gu Zuyu's book *Important Notes on Reading the Geography Treatises in the Histories.* Gu grew up in a well educated family during the slow collapse of the Ming Dynasty. He witnessed the conquest of China by the Manchus, a minority group and, like many other intellectuals in that period, felt ashamed by it. As a result, Gu decided to write a book on the geography and history of local jurisdictions as delineated in the late Ming Dynasty. He sought to document the geographic features of military importance for all places in China and thus provide a guide to patriots to better protect China in the future. Gu had access to one of the best private libraries at his time. So he read extensively formal histories, historical documents, and local gazetteers.[32] He also collected first-hand information by traveling to different places. Gu spent more than thirty years working on his book. The final product was essentially an encyclopedia of the geography and history of late-Ming-Dynasty local jurisdictions. A 2005 republication of Gu's book was divided into 12 volumes and together had 6,294 pages. It remains one of the most important references for the study of local jurisdictions in the Ming Dynasty.[33]

Gu organized his book according to the government structure of Ming Dynasty. Below the central government were a number of provinces. In each province, there were prefectures (*fu*) followed by subprefectures (*zhou*). The lowest unit was the county (*xian*). There were two large areas that belonged to no province, but were metropolitan areas (*jing*) attached

---

[31] Local gazetteers also tend to mention city population sizes. However, such numbers are almost always jurisdiction population instead of population inside the city wall. Skinner's team of researchers did not record such population numbers. Instead, they tried to estimate population residing inside the city wall for all the cities. They decided to use a discrete variable that has 11 population size categories. Such a crude measure of population is not very useful for studying city size distribution. Skinner (1977b) discussed the rank-size distribution of Qing Dynasty cities based on estimated population sizes.

[32] The wide range of references Gu consulted is evident from his extensive citations. However, he did not provide a complete documentation of all the references, some of which did not survive.

[33] See, e.g., Liang (2008, pp. 282-336) who uses the information in Gu's book to calculate population in the Ming Dynasty.

to Nanjing and Beijing. In early years, the capital of the empire was Nanjing (*Yingtian Fu*, 1368–1421); the third emperor moved the capital to Beijing (*Shuntian Fu*, 1421–1644). For each local jurisdiction (down to the county level), Gu recorded its population size and important historical and geographical information on the city (or cities) in the jurisdiction.

A "city" here refers to the capital of a prefecture, subprefecture, or county. City populations in Gu's book includes those who lived inside the city wall and those outside the city wall within the jurisdiction. Since city wall was of major military importance, Gu always commented on it. The circumference of city wall was almost always recorded in Gu's book. Sometimes the number of gates and building materials were also recorded. In many cases, a brief history of the city wall is sketched, indicating when it was first built and at what time it was destroyed, rebuilt, repaired, fortified, etc.

Based on the circumference of city wall, the ten largest cities in the Ming Dynasty were Nanjing, Beijing, Fengyang, Xi'an, Hangzhou, Suzhou, Taiyuan, Quanzhou, Zhenjiang, and Chengdu. Seven of them were still among the top ten in the Qing Dynasty. The interesting case is Fengyang, which was the third largest city in the Ming Dynasty but dropped out of the top ten in the Qing Dynasty. Fengyang was the hometown of Zhu Yuanzhang, the first emperor of the Ming Dynasty. In 1369, one year after Zhu became the emperor, he started to build Fengyang aggressively with the intention to eventually move his capital there. The plan was later abandoned; the oversized Fengyang could not be sustained by economic forces and declined over time.

References

[1] Black, Duncan, and Vernon Henderson. 2003. "Urban Evolution in the USA." *Journal of Economic Geography* 3, 343–372.

[2] Bleakley, Hoyt and Jeffrey Lin. 2012. "Portage and Path Dependence." *Quarterly Journal of Economics* 127 (2), 587–644.

[3] Brook, Timothy. 1997. "Native Identity under Alien Rule: Local Gazetteers of the Yuan Dynasty." In *Pragmatic Literacy, East and West, 1200-1330*, ed. R. Britnell, 235–246. Woodbridge: Boydell & Brewer.

[4] Brooks, Leah and Byron Lutz. 2014. "Vestiges of Transit: Urban Persistence at a Micro Scale." Unpublished manuscript, Trachtenberg School of Public Policy and Public Affairs, George Washington University.

[5] Buck, Pearl S. 1931. *The Good Earth.* New York: John Day Company.

[6] Chang, Sen-Dou. 1977. "The Morphology of Walled Capitals." In *The City in Late Imperial China*, ed. G. W. Skinner, 75–100. Stanford: Stanford University Press.

[7] China State Statistical Bureau (CSSB). 1985. *China Urban Statistical Yearbook 1985.* Beijing: China Statistical Press. (in Chinese)

[8] China State Statistical Bureau (CSSB). 2014. *China Urban Statistical Yearbook 2014.* Beijing: China Statistical Press. (in Chinese)

[9] Córdoba, Juan-Carlos. 2008. "On the Distribution of City Sizes." *Journal of Urban Economics* 63, 177–197.

[10] Desmet, Klaus, and Jordan Rappaport. 2013. "The Settlement of the United States, 1800 to 2000: The Long Transition Towards Gibrat's Law." Research Working Paper 13-02, Federal Reserve Bank of Kansas City.

[11] Dittmar, Jeremiah. 2011. "Cities, Markets, and Growth: The Emergence of Zipf's Law." Unpublished manuscript, Department of Economics, American University.

[12] Dobkins, Linda H., and Yannis M. Ioannides. 2000. "Dynamic Evolution of the U.S. City Size Distribution". In *The Economics of Cities, Theoretical Perspectives*, ed. J.-M. Huriot and J.-F. Thisse, 217–260. Cambridge: Cambridge University Press.

[13] Duranton, Gilles. 2006. "Some Foundations for Zipf's Law: Product Proliferation and Local Spillovers." *Regional Science and Urban Economics* 36, 542–563.

[14] Duranton, Gilles, and Diego Puga. 2014. "The Growth of Cities." In *Handbook of Economic Growth Volume 2*, ed. P. Aghion and S. Durlauf, 781–853. Amsterdam: Elsevier.

[15] Eeckhout, Jan. 2004. "Gibrat's Law for (All) Cities." *American Economic Review* 94, 1429–1451.

[16] Gabaix, Xavier. 1999. "Zipf's Law for Cities: An Explanation." *Quarterly Journal of Economics* 114, 739–767.

[17] Gabaix, Xavier, and Rustam Ibragimov. 2011. "Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents." *Journal of Business Economics and Statistics* 29, 24–39.

[18] Gabaix, Xavier, and Yannis M. Ioannides. 2004. "The Evolution of City Size Distributions." In *Handbook of Regional and Urban Economics Volume 4: Cities and Geography*, ed. J. V. Henderson and J.-F. Thisse, 2341–1278. Amsterdam: Elsevier.

[19] Ge, Jianxiong, Songdi Wu, and Shuji Cao. 1997. *The History of Migration in China.* Fuzhou: People's Press of Fujian. (in Chinese)

[20] Gu, Zuyu. 2005. *Du Shi Fang Yu Ji Yao.* Beijing: Zhonghua Shuju Press. (in Chinese)

[21] Hammel, Eugene A. 2009. George William Skinner 1925-2008. Online Collection of Biographical Memoirs, National Academy of Sciences. Available online at: http://www.nasonline.org/publications/biographical-memoirs/memoir-pdfs/skinner-g-william.pdf.

[22] Holmes, Thomas J., and Sanghoon Lee. 2009. "Cities as Six-by-Six-Mile Squares: Zipf's Law?" In *The Economics of Agglomerations*, ed. E.L. Glaeser, 105–132. Chicago: University of Chicago Press.

[23] Hornbeck, Richard and Daniel Keniston. 2014. "Creative Destruction: Barriers to Urban Growth and the Great Boston Fire of 1872." Unpublished manuscript, Department of Economics, Harvard University.

[24] Hsu, Wen-Tai. 2012. "Central Place Theory and City Size Distribution." *Economic Journal* 122, 903–932.

[25] Ioannides, Yannis M. 2013. *From Neighborhoods to Nations: The Economics of Social Interactions.* Princeton, NJ: Princeton University Press.

[26] Ioannides, Yannis M., and Henry G. Overman. 2003. "Zipf's Law for Cities: An Empirical Examination." *Regional Science and Urban Economics* 33, 127–137.

[27] Ioannides, Yannis, and Spyros Skouras. 2013. "US City Size Distribution: Robustly Pareto, But Only in the Tail." *Journal of Urban Economics* 73, 18–29.

[28] Krugman, Paul. 1996. "Confronting the Mystery of Urban Hierarchy." *Journal of the Japanese and International Economies* 10, 399–418.

[29] Lee, Sanghoon, and Qiang Li. 2013. "Uneven Landscapes and City Size Distributions." *Journal of Urban Economics* 78, 19–29.

[30] Liang, Fangzhong. 2008. *Zhong Guo Li Dai Hu Kou Tian Di Tian Fu Tong Ji.* Beijing: Zhonghua Shuju Press. (in Chinese)

[31] Nunn, Nathan, and Diego Puga. 2012. "Ruggedness: The Blessing of Bad Geography in Africa." *Review of Economics and Statistics.* 94: 20–36.

[32] Rossi-Hansberg, Esteban, and Mark L. J. Wright. 2007. "Urban Structure and Growth." *Review of Economic Studies* 74, 597–624.

[33] Rozenfeld, Hernán D., Diego Rybski, Xavier Gabaix, Hernán A. Makse. 2011. "The Area and Population of Cities: New Insights from a Different Perspective on Cities." *American Economic Review* 101, 2205–2225.

[34] Siodla, James. 2015. "Razing San Francisco: The 1906 Disaster as a Natural Experiment in Urban Redevelopment." *Journal of Urban Economics* 89, 48–61.

[35] Sirén, Osvald. 1924. *The Walls and Gates of Peking: Researches and Impressions.* New York: Orientalia.

[36] Skinner, G. William. 1977a. "Introduction: Urban Development in Imperial China." In *The City in Late Imperial China*, ed. G. W. Skinner, 3–31. Stanford: Stanford University Press.

[37] Skinner, G. William. 1977b. "Regional Urbanization in Nineteenth-Century China." In *The City in Late Imperial China*, ed. G. W. Skinner, 211–249. Stanford: Stanford University Press.

[38] Skinner, G. William. 1977c. "Cities and the Hierarchy of Local Systems." In *The City in Late Imperial China*, ed. G. W. Skinner, 275–351. Stanford: Stanford University Press.

[39] Soo, Kwok Tong. 2005. "Zipf's Law for Cities: A Cross-Country Investigation." *Regional Science and Urban Economics* 35, 239–263.

[40] Waln, Nora. 1933. *The House of Exile.* Boston: Little, Brown and Company.

[41] Xing, Chunbing, and Junfu Zhang. 2015. "The Preference for Larger Cities in China: Evidence from Rural-Urban Migrants." Unpublished manuscript, Department of Economics, Clark University.

[42] Yue, Zumou, G. William Skinner, and Mark Henderson. 2007. "ChinaW Dataset: Nineteenth-Century Cities, Yamens, and County-level Units." Unpublished data manual.

[43] Zhang, Yuhuan. 2003. *The History of Chinese Walled Cities.* Tianjin, China: Baihua Literature and Art Publishing House. (in Chinese)

[44] Zipf, George K. 1949. *Human Behavior and the Principle of Least Effort.* Cambridge, MA: Addison-Wesley.