Walled Cities in Late Imperial China

Yannis Ioannides and Junfu Zhang[*]

June 28, 2014

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. To rationalize the existence of city walls, we propose a simple model that relates the dimensions of city walls to a set of economic variables. Guided by this model, we conduct an empirical analysis using hand-collected and previously unused data on city walls in the Ming (1368–1644) and Qing (1644-1911) Dynasties. Consistent with the model, we find that the circumference of a city wall is positively correlated with population size in the jurisdiction and that frontier cities subject to a higher probability of attack tended to have stronger city walls. Since a city wall imposes a physical boundary around a city, the land area inside the city wall provides a natural proxy of city size. We examine the physical size distribution of walled cities in late imperial China. We find that city sizes above a certain cutoff follow Zipf's law, although the Zipf coefficient is sensitive to the choice of the cutoff point. This result complements findings in the existing literature that focuses almost exclusively on the population size distribution of cities.

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

JEL Classification: R12.

[Preliminary draft; comments are welcome.]

---

[*]Ioannides is the Max and Herta Neubauer Chair and Professor in Economics at Tufts University (Address: Department of Economics, Tufts University, Medford, MA 02155; Tel: 617-627-3294; Email: yannis.ioannides@tufts.edu). Zhang is an Associate Professor of Economics at Clark University (Address: 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.

1

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)

[The Achaeans] built a high wall to shelter themselves and their ships; they gave it strong gates that there might be a way through them for their chariots, and close outside it they dug a trench deep and wide, and they planted it within with stakes.   Homer's Iliad (Book VII)

1   Introduction

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. City walls represented a most salient feature of Chinese cities throughout history until the mid-twentieth century, when the government led a movement 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 for example 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 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.[1]

City walls were also common in other civilizations. According to Homer's Iliad, an epic based on the Trojan War which scholars date to about 3,200 years ago, the city of Troy had strong walls with high towers and great gates. At the archaeological site of Troy, 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.

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

---

[1]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.

Figure 1: City wall of Beijing in late 19th century
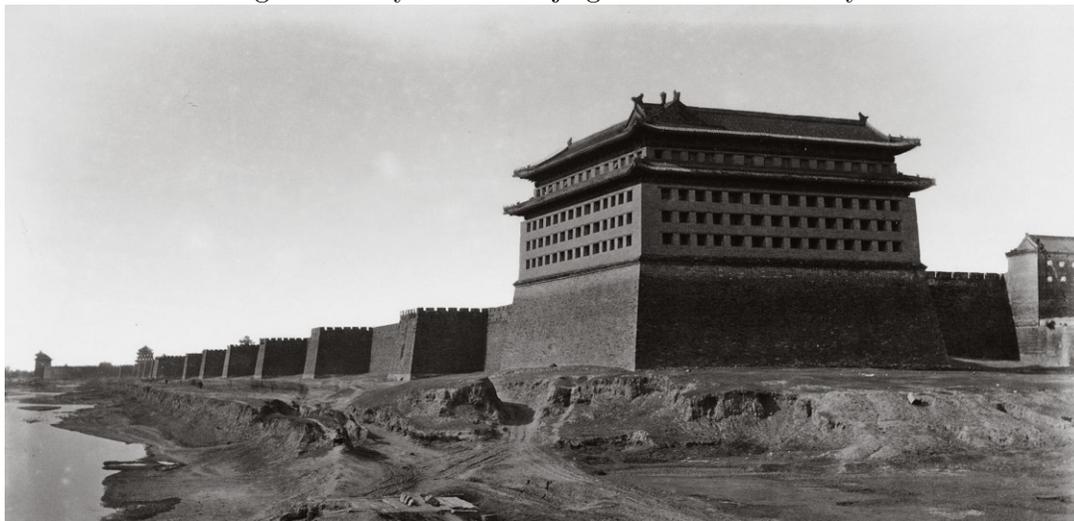
Source: http://www.photographium.com/south-gate-beijing-china-1874, with authors' modifications.

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 this study, we rationalize the existence of city walls using a simple monocentric city model. In our model, city walls are 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.

To test the predictions of the model, we use two unique and hitherto unutilized data sources. The first dataset was constructed by hand-collecting 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 the early Qing (1644-1911) Dynasty scholar Gu Zuyu (1631-1692). In his book, Gu sought to cover the history and geography of all places in China in the late Ming Dynasty. We were able to code data 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 suggested by the model.

The second dataset was 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, another prediction of the model.

Since city walls were built to protect urban residents and properties from outside attacks, they were the physical boundaries of cities. Therefore, the land area inside a city wall is a natural measure of 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 the model. For

3

both periods, we find evidence that above a certain size cutoff the physical size of walled cities indeed follows a power law. This part of the analysis contributes to a large body of the literature on city size distribution.

There has been well-documented empirical evidence that city size tends to be inversely proportional to city rank. That is, in an economy such as the U.S., the second largest city is roughly half of the size of the largest one, the third largest city is roughly a third of the size of the largest one, etc. If we let n be the rank of the nth largest city and regress log rank on log city size, we tend to find a linear relationship with a coefficient of negative one. This empirical regularity is called the rank-size rule or Zipf's law of city size distribution (Zipf 1949).

Despite the statistical significance of Zipf's law, traditionally it received little attention from economists and the standard urban economic theory of city sizes is hard-pressed to justify it (e.g., Henderson 1988). Krugman (1995, p. 44) famously remarked that the rank-size rule was "a major embarrassment for economic theory." Since then, urban economists have devoted a considerable amount of research effort to this topic and there have been some serious attempts on the theory side to provide a microeconomic foundation for the size distribution of cities.[2]

In the meantime, the empirical literature on city-size distribution has continued to grow.[3] Economists have assembled more and higher-quality data and applied more advanced econometric techniques to characterize the distribution of city sizes. One important development is the examination of city size distribution over much smaller cities. Using U.S. data on "census places" in 2000, Eeckhout (2004) shows that the size distribution of all cities is log normal. He argues that earlier evidence on city size distribution was always based on relative small samples of the largest cities and that Zipf's law (or, more weakly, the Pareto distribution) is supported only because the upper tail of the lognormal distribution is difficult to distinguish from the Pareto distribution. While Eeckhout's claim of a lognormal distribution is debated (Levy 2009; Eeckhout 2009), more evidence has appeared in support of the Pareto law for the upper tail (Rozenfeld *et al.* 2011; Ioannides and Skouras 2013).

The empirical research on city size distribution has focused almost exclusively on population size, perhaps because population data are rather accurately measured and easily available in all modern societies. However, population data are generally aggregated based on political or administrative boundaries, and these boundaries do not necessarily coincide with the boundaries of a city as an economic entity. One way to deal with this problem is to conduct empirical analysis using alternative definitions of cities and place more confidence in the more robust results (e.g., Ioannides and Skouras 2013). Another approach is to build

---

[2]See, e.g., Krugman (1996), Gabaix (1999), Eeckhout (2004), Duranton (2006), Córdoba (2008), Hsu (2012), and Lee and Li (2013).

[3]See, e.g., Dobkins and Ioannides (2000), Black and Henderson (2003), Ioannides and Overman (2003), Eeckhout (2004), Soo (2005), Holmes and Lee (2009), Rozenfeld *et al.* (2011), and Ioannides and Skouras (2013). Gabaix and Ioannides (2004) provide an excellent survey of the earlier empirical and theoretical literature.

Figure 2: A square-shaped city and surrounding rural area
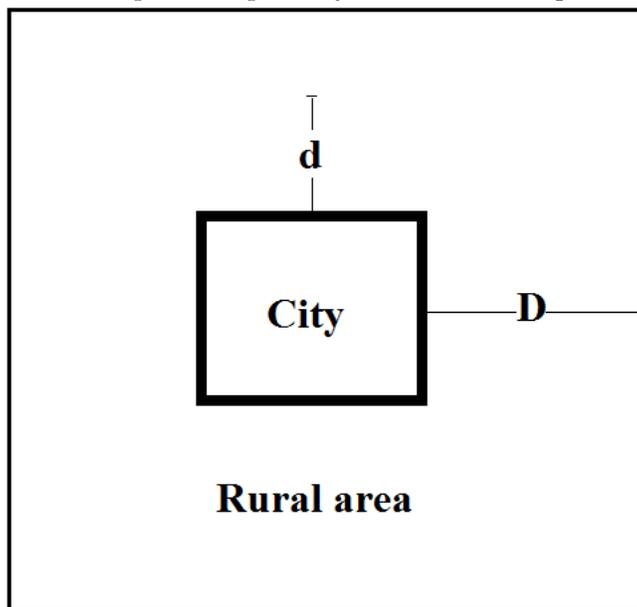
cities "from the bottom up," i.e., to construct cities using high-resolution micro spatial data without adopting arbitrary political boundaries (Rozenfeld *et al.* 2011).

Our analysis of the size distribution of walled cities in late imperial China adds a new perspective to this literature. Our key innovation is to measure city size using the land area it occupies.[4] The long Chinese tradition of building protective walls around cities has made this approach feasible. We show rather strong evidence that the physical size distribution of larger walled cities follows a power law. This finding, combined with more recent contributions by Dittmar (2011) and Desmet and Rappaport (2013), provides a much deeper historical dimension to the empirical regularities of city size distribution. It not only helps us better understand walled cities, but also has important implications for the proper way to model cities.

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:

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

---

[4]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 law.

where Y_c is the total output of clothing and R is a production amenity. For the moment, we may think of R as river size. It is assumed that R follows a power law, which is supported by evidence from both the U.S. and China.[5] Land is not used in the manufacturing production, but each worker demands inelastically one unit of residential land in the city. One possible interpretation is that production occurs at the household level. We will use clothing as the numeraire good: $P_c = 1$.

Workers are paid according to the value of their marginal product, so urban wages 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 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, a farmer can produce x units of food. Through a share-cropping contract, the landlord (the local government in the city) will pay the farmer θx and keep $(1-\theta)x$, where $0 < \theta < 1$. As we will see shortly, for spatial equilibrium θ is 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

Every individual, whether a worker or a farmer, has the same utility function $U = A F^\beta C^{1-\beta}$, 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.[6] 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

---

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

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

also suggesting a power law distribution.

[6]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 rural production.

Some of the food produced in rural area will be 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; there is no shipping cost if goods are moved parallel to the city edge.[7] In particular, if a good is sold for price P at the location of production, to offset 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 inside the city. Then in the city a worker's 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 farmers trade with workers in the city. 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 $x P_f e^{-\tau D}$. Clothing price at distance D is $P_c e^{\tau D} = e^{\tau D}$. Then the utility at distance D is

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

Spatial equilibrium requires that $V_D = V_u$: $x P_f^{1-\beta} e^{-2(1-\beta)\tau D} = \alpha R N^{\alpha-1} P_f^{-\beta}$. That 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 $\theta(d) x P_f e^{-\tau d} = x P_f e^{(2\beta-2)\tau D + (1-2\beta)\tau D}$. At price

---

[7]This is an assumption to make the calculation easier under a square-shaped city. No need for this if we deal with a circular or a linear structure.

$P_f e^{-\tau d}$, this farmer's demand for food is

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

That is, each farmer consumes a fraction β of the food output as compensation he receives (according to the share-cropping contract), and sells the remainder $(1-\beta)\theta(d)x$ on the market so that he can buy 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 this amount of revenue a farm receives from selling food depends on distance d. Since the total number of farmers at distance d is $\frac{4(\sqrt{N}+2d)}{\lambda}$, the total food revenue for all farmers is

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

This is simply the total expenditure on food by all the workers in the city. Remember that each worker's income is $\alpha RN^{\alpha-1}$. Utility maximization require a β-fraction of the income to 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

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

It defines an equilibrium maximum distance as a function of city population and other parameters as the unique root of the above equation: $D(N,\lambda,\tau,\beta)$. Obviously, $D_N(\cdot) > 0$ 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 $\underline{V}$.[8] These people will move to the city as long as a worker's utility is higher than $\underline{V}$. Thus in equilibrium, a worker's utility will be exactly $\underline{V}$. From equations (1) and (2), this implies

---

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

that

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

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

$$(1-\alpha) \ln N + 2\beta\tau D(N, \lambda, \tau, \beta) = \ln \alpha + \ln R + \frac{\beta}{1-\beta} \ln x - \frac{1}{1-\beta} \ln \underline{V}. \qquad (4)$$

This closes the model and allows to solve implicitly for equilibrium city population $N^* = N(R, \alpha, x, \lambda, \tau, \underline{V})$. Since $D_N(\cdot) > 0$ and $D_\lambda(\cdot) > 0$, we know 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 make intuitive sense.

From equation (4), 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 solve for $D(N, \lambda, \tau, \beta)$ explicitly. 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 (4) becomes

$$(1-\alpha) \ln N + \tau\sqrt{N} \left(\sqrt{1+\lambda} - 1 \right) / 2 = \ln \alpha + \ln R + \ln x - \frac{1}{2} \ln \underline{V}. \qquad (5)$$

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 $\ln N$.[9] We have

---

[9]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 r.h.s. of (5), 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 (5) may be rewritten as $\mathcal{N}(N) = \rho R$, where ρ is a function of all parameters in the r.h.s of (5) 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:

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

viewed R as river size, a natural production amenity. But actually, R can be anything that affects urban productivity. In particular, R can be interpreted as the accumulation of all productivity shocks in the past. 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 "profit" 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. Note that neither the surplus in the urban sector nor the surplus in the rural sector 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 total food surplus is

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

Total surplus in the two sectors is

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

Since D is an increasing function of N, it is clear that $S'(N) > 0$.

City wall

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 a high density area and all the surplus is stored in the city. A city wall will reduce the loss when an attack happens.

The circumference of the wall (or equivalently, the area inside the wall) is determined by N^*, thus a power law distribution of N^* implies a power law distribution of (the circumference of or the area inside) city wall.

---

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 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 a city will only retain a fraction π of its total surplus if an attack happens. We assume $0 < \pi(N^*, h) < 1$, $\pi_h > 0$, and $\pi_{hh} < 0$; 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: $\pi_{N^*} > 0$ and $\pi_{N^*N^*} < 0$. Both the size and the quality of the city wall is 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.

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:

$$\max_h : \quad \gamma\pi(N^*, h)S(N^*) + (1 - \gamma)S(N^*) - c(N^*, h).$$

The first order condition $\gamma\pi_h(N^*, h^*)S(N^*) - c_h(N^*, h^*) = 0$ 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 $h_\gamma = \frac{\pi_h S}{c_{hh} - \gamma\pi_{hh}S} > 0$. 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 growth in the future, 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 in what would have a lognormal distribution of city sizes. That is, if the cost of maintaining a city wall contains a *fixed* component, then very small cities would be infeasible. If the random factors invoked by Lee and Li (2013) were to point downwards, this consideration acts as a reflective barrier, preventing cities from becoming too small. This has a dramatic effect in that it transforms a lognormal to a Pareto distri-

bution, with a mode at the lower tail, and an upper tail that would be fatter than that of a lognormal. That is, as Duranton and Puga (2014, p. 836) argue in the context of random growth, 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 be suitably modified to accommodate a lower bound. 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.

Summary

We have presented a simple model to explain why city walls existed. It has four 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.* City wall is built to protect urban population, so it is positively correlated with urban population. Given $D_N(N, \lambda, \tau, \beta) > 0$ and $D_\lambda(N, \lambda, \tau, \beta) > 0$, where λ is the fixed density in rural areas, urban population is positively correlated with rural population. Thus the sum of urban and rural population is positively correlated 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.*

3   Data

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

The first dataset used in this study result from a long-term research project led by the late G. William Skinner. Professor Skinner, before he passed away in 2008, was widely considered "the most eminent anthropological sinologist in the United States."[10] He was best known for his spatial approach to Chinese history. In later years of his long academic

---

[10]For a biographical memoir of Professor Skinner, see Hammel (2009).

12

Table 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

career, 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.[11] 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 documented in Yue, Skinner, and Henderson (2007), 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.

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.

To construct the ChinaW dataset, Skinner and collaborators consulted a total of 931 gazetteers published during 1519-1974. Table 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.[12]

Among the 1,869 relevant observations, there are 224 places for which the Skinner research team could not decide whether they 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 be 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 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. To proceed, we will focus on the sample of 1,623 cities with the city-wall circumference variable available and use 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). That is, 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

---

[12]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.

Table 2: 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

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.

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 a square. 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 will proceed by using this estimate as the city size. In the next section, we will explicitly specify the conditions under which we may use this inaccurate estimate to draw inferences about the actual city-size distribution.[13]

In Table 2, we present some descriptive statistics on city walls. The key variable, area inside city wall, 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.

---

[13]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.

Notice that the average city wall is 7.47 meters high and 7.26 meters thick at base. It has 4 gates and 8 towers. With a 96 percent chance there were also moats surrounding the city wall. 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.

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

The second dataset contains information on the circumferences of city walls and jurisdiction population sizes. We hand-collected these data from a 130-chapter publication entitled *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). Gu grew up in a well educated family during the slow collapse of the Ming Dynasty. He witnessed the conquest of China by the minority group Manchus 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 intended 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 in formal histories, historical documents, and local gazetteers.[14] 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.[15]

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 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.

---

[14]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.

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

Table 3: 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 lived both inside and outside the city wall.

A "city" here refers to the capital of a prefecture, subprefecture, or county. City population in Gu's book includes those 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.

For our empirical analysis, we only collect data on jurisdiction population and the circumference of city wall.[16] 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.[17] 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 the whole number but indicated in our data file whether the number is rounded up or down.

Table 3 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 more 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

---

[16]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).

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

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.

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, as listed above. 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.

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 4. 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 the first column, we regress the length of city wall on jurisdiction population. We then estimate the area inside a city wall using the formula $area = \left(\frac{length\ of\ city\ wall}{4}\right)^2$ (i.e., assuming a square-shaped city) and regress this estimated area 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 separate 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 provinces dummies. When using the full sample in columns 1-3, we also control for administrative level dummies. Standard errors are computed by clustering observations

18

Table 4: Circumference of 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. ***: $p < 0.01$.

at the province level.

Across different specifications, the size of city wall is always positively correlated with jurisdiction population. That is, when a larger population pay taxes to the local government, the local government tend 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.

4.2 The qualities of city wall

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:[18]

- On inner Asian frontiers: 38 cities;

- 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 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

---

[18]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
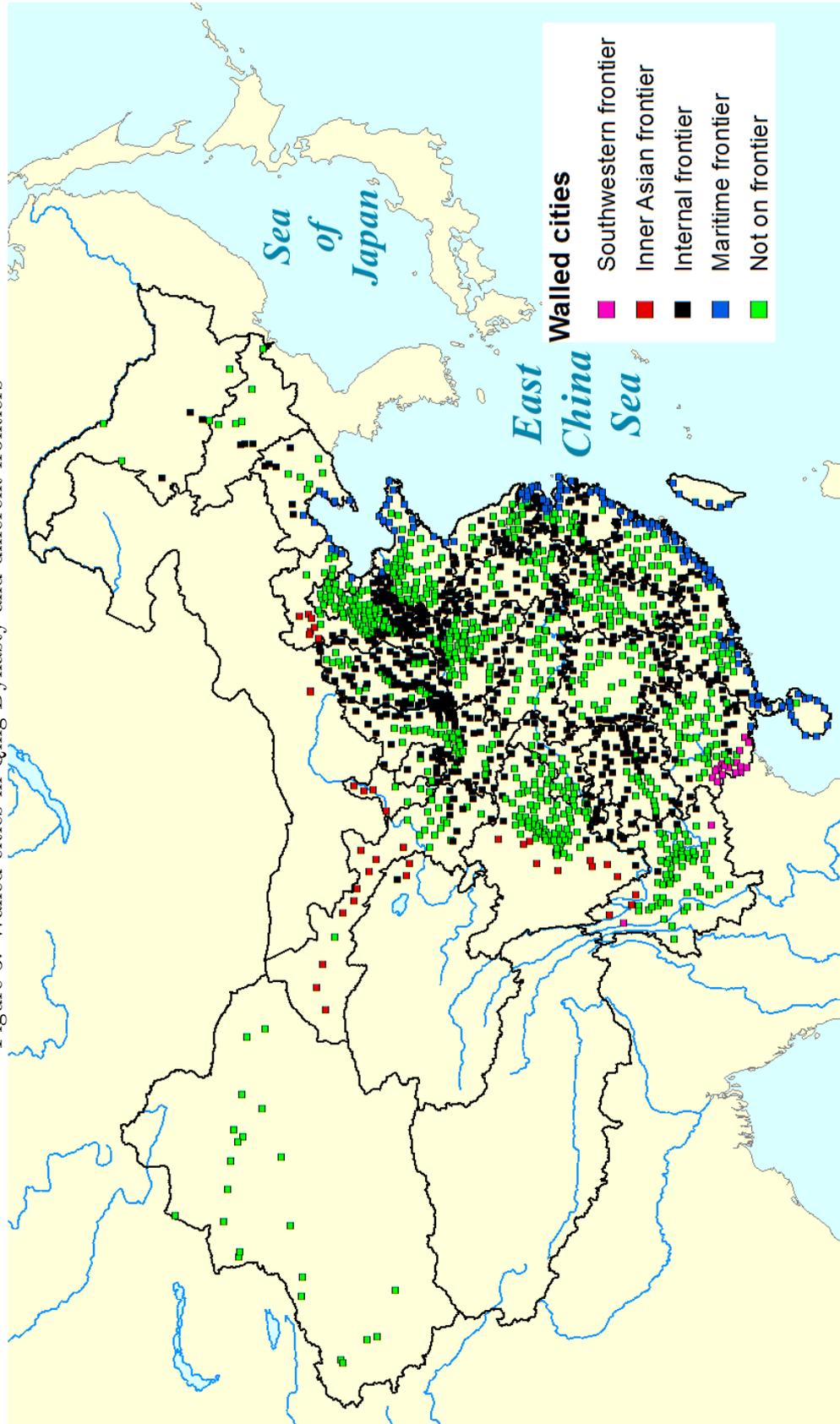
A total of 1,623 walled cities are shown on this map. Black lines indicate the international and provincial borders of modern China.

Walled cities

Southwestern frontier
Inner Asian frontier
Internal frontier
Maritime frontier
Not on frontier

Sea of Japan

East China Sea

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.[19] 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, 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

---

[19]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 $\left(\frac{length\ of\ city\ wall}{4}\right)^2$, assuming a square-shaped city. Standard errors in parentheses are clustered by province. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

enemies or bandits.[20]

4.3 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.

4.3.1 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 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 write

$$(1 + \epsilon) S = \hat{S}, \tag{6}$$

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:

$$f(S) = \frac{b\underline{S}^b}{S^{b+1}} \qquad \forall S \geq \underline{S};$$

$$F(S) = 1 - \left(\frac{\underline{S}}{S}\right)^b \qquad \forall S \geq \underline{S},$$

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 R be the

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

rank of a city with size S and N the number of cities in the sample (i.e., the rank of size $\underline{S}$), then in expectation

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

Taking natural logs yields

$$\ln R = a - b\ln S, \tag{7}$$

where $a \equiv \ln N + b\ln\underline{S}$ 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 have an inaccurately measured $\hat{S} = (1 + \epsilon)S$, 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. Given that $\ln(1+\epsilon) \approx \epsilon$ for small ϵ, plug $S = (1 + \epsilon)^{-1}\hat{S}$ into equation (7) to get

$$\ln R = a' - b\ln\hat{S} - b(\epsilon - \bar{\epsilon}), \tag{8}$$

where $a' \equiv \ln N + b\ln\underline{S} - b\bar{\epsilon}$ 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 (8) 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 (8) will be the same as the the coefficient of $\ln S$ in equation (7), which can be used to test for Zipf's law.
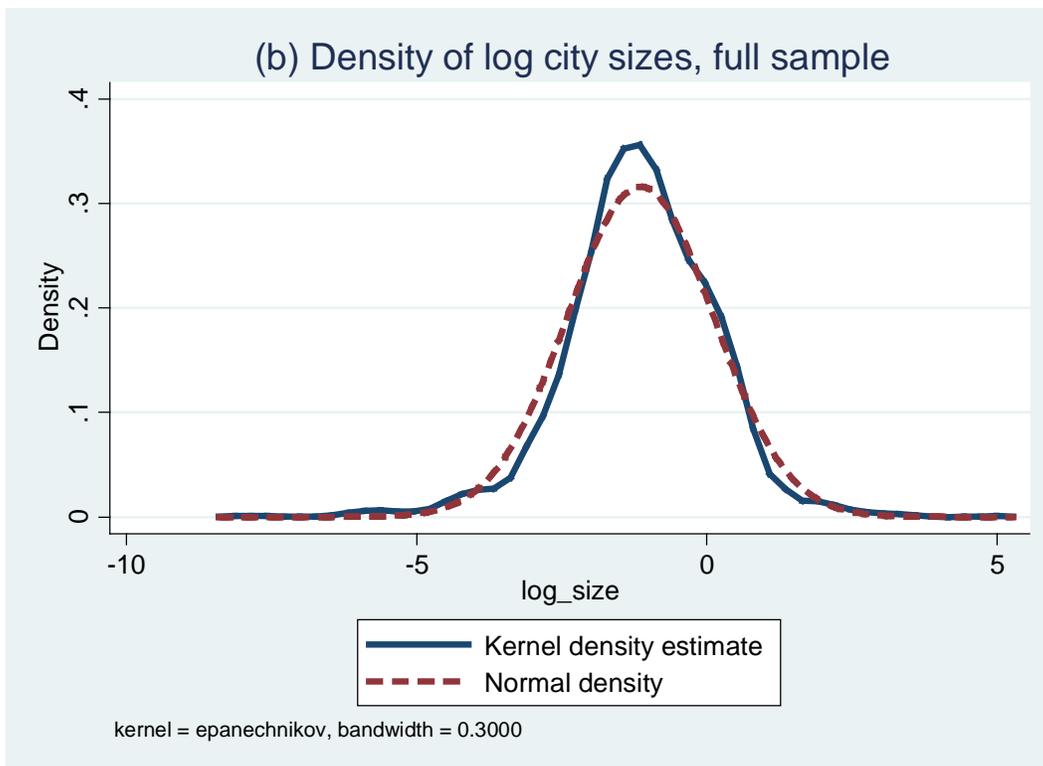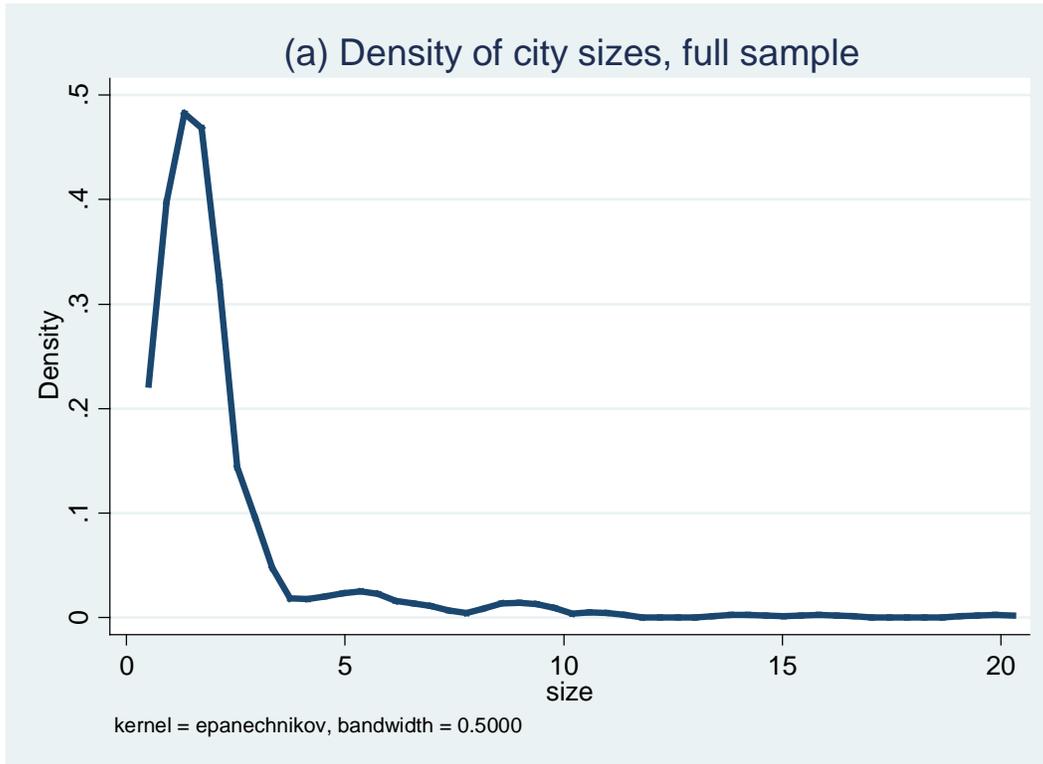
4.3.2 Size distribution of all walled cities

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 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

Figure 4: Density of city sizes in Qing Dynasty

(a) Density of city sizes, full sample

kernel = epanechnikov, bandwidth = 0.5000

(b) Density of log city sizes, full sample

Kernel density estimate

Normal density

kernel = epanechnikov, bandwidth = 0.3000

$N = 1,623$. Unit: square kilometer.

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 equal to 0.000. The second one is a Skewness/Kurtosis test for normality, which yields a p-value equal to 0.0000.[21] 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.[22]

Cities in Ming Dynasty

We visualize the size distribution of Ming Dynasty cities in Figure 5. 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.

4.3.3 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 $\ln R$ against $\ln \hat{S}$ using the full sample of walled cities in the Qing Dynasty, which is in panel (a) of Figure 6. The right portion of the plot indeed appears to be a negative 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.

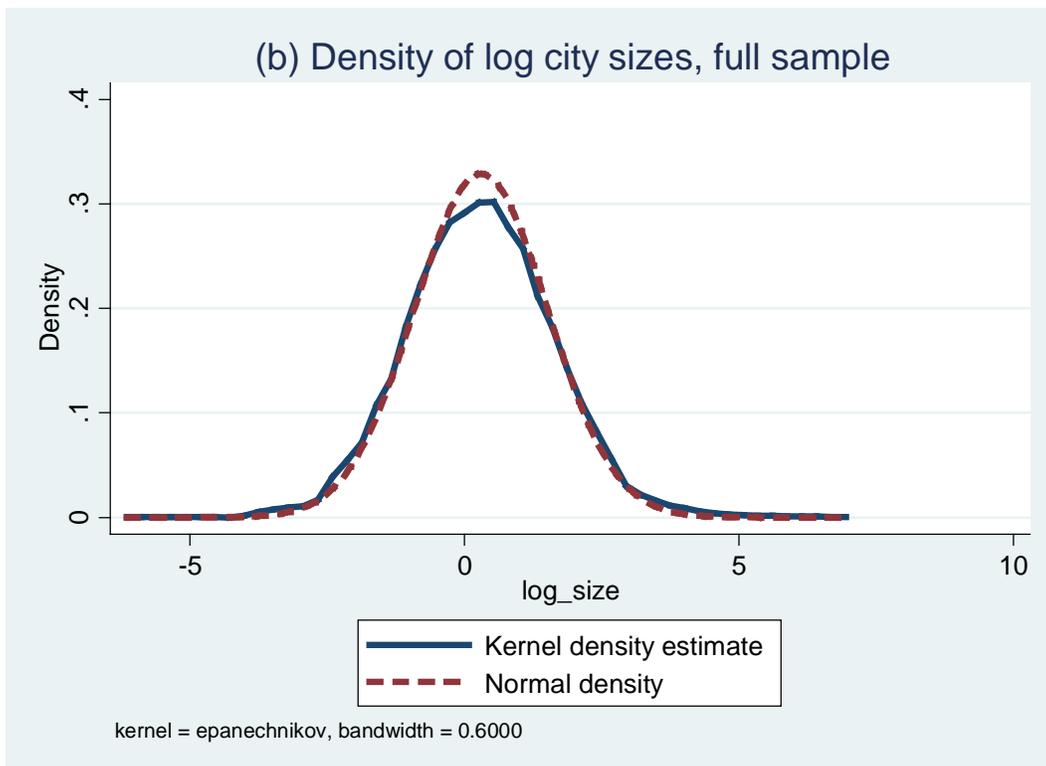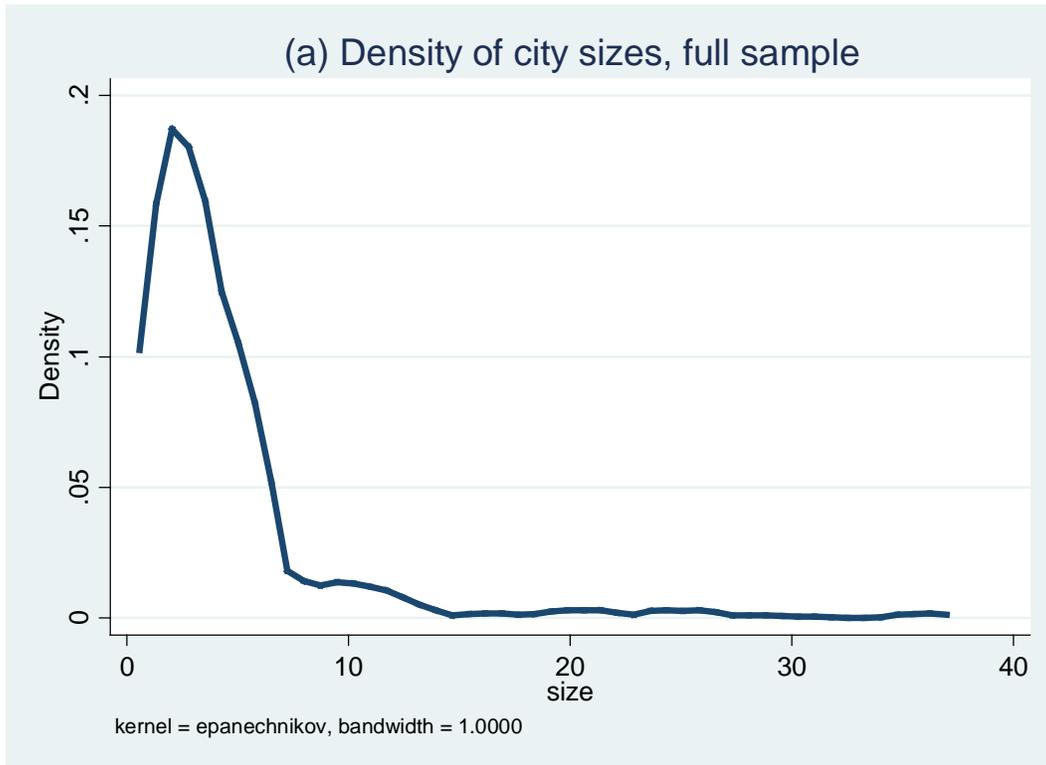To identify the break in the slope, we run the following regression:

$$\ln R = a - b_1 \ln \hat{S} - b_2 \left(1_{R>R^*} \ln \hat{S}\right) + e, \tag{9}$$

where a is a constant, b_1 and b_2 are coefficients, and e is the error term. $1_{R>R^*}$ is an indicator function that takes value 1 if rank R is greater than a particular rank R^* and value 0 otherwise. That is, this regression allows the coefficient of $\ln \hat{S}$ to be different above and below rank R^*. We search the break point between the 40th percentile and

---

[21]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.

[22]Despite the rejection of a lognormal distribution in the formal tests, it is rather remarkable how closely the density of log city size in our sample 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.

Figure 5: Density of city sizes in Ming Dynasty

(a) Density of city sizes, full sample

kernel = epanechnikov, bandwidth = 1.0000

(b) Density of log city sizes, full sample

Kernel density estimate
Normal density

kernel = epanechnikov, bandwidth = 0.6000

$N = 1,182$. Unit: square li.

Figure 6: Log rank again log size for walled cities in Qing Dynasty

(a) Log rank against log size, full sample

(b) Log rank against log size, truncated sample
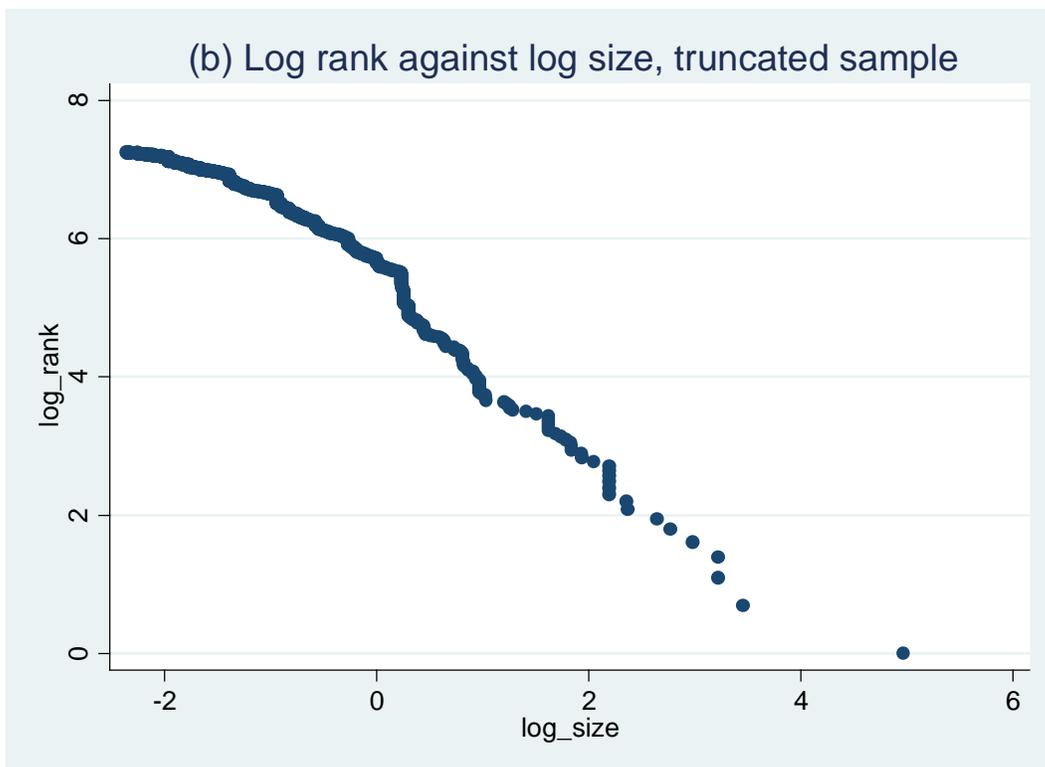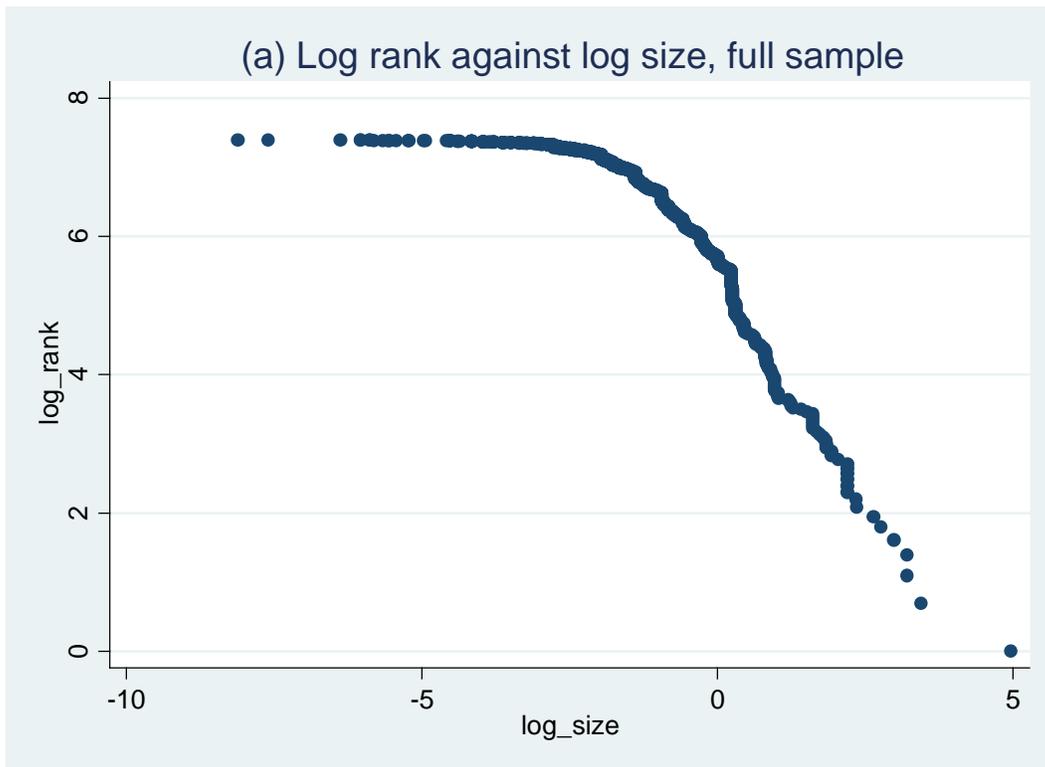
$N = 1{,}623$ for panel (a) and 1,409 for panel (b).

Table 6: Regressions of log rank on log size, Qing Dynasty

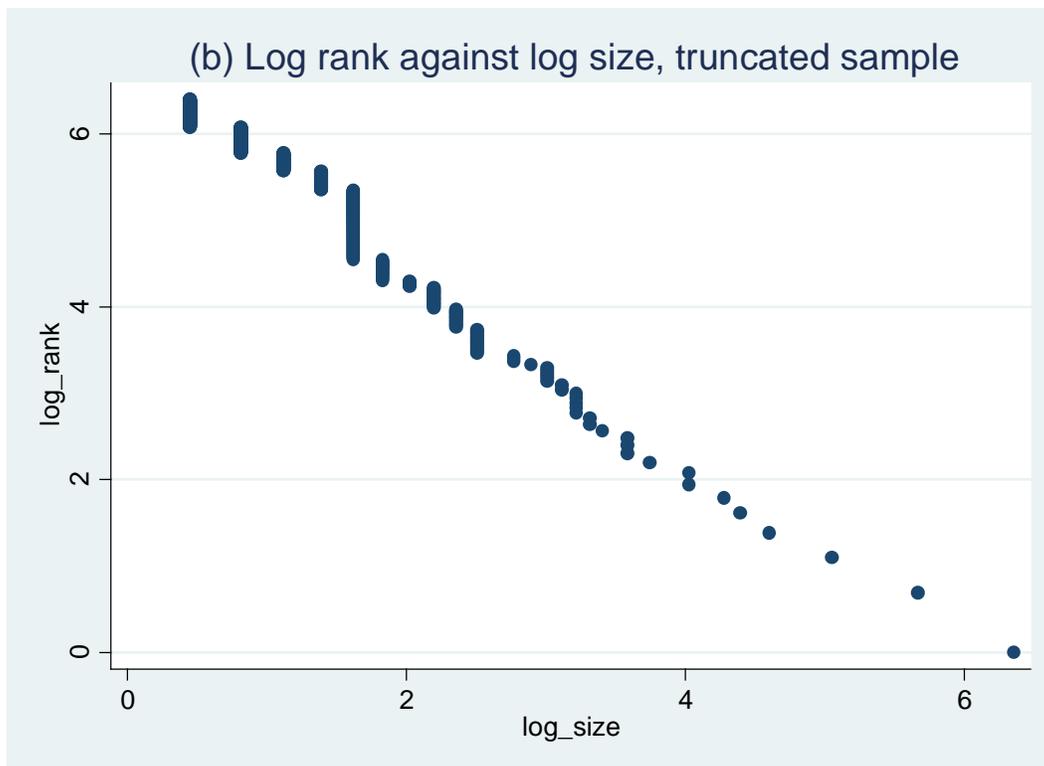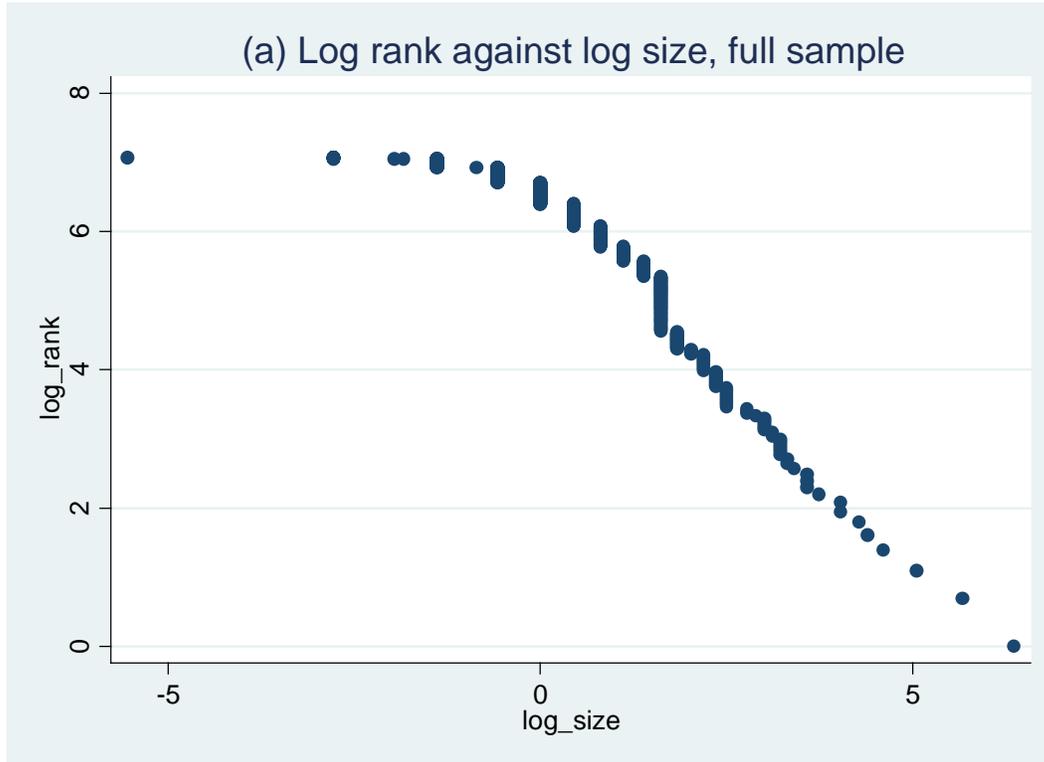|	Dependent variable: Log rank ($\ln R$)					
	(1)	(2)	(3)	(4)	(5)	(6)
	Rank	Rank	Rank	Rank	Rank	Rank
	$\leq 1,409$	$\leq 1,200$	$\leq 1,000$	≤ 800	≤ 600	≤ 400
Constant	5.427***	5.418***	5.427***	5.456***	5.493***	5.553***
	(0.008)	(0.007)	(0.006)	(0.006)	(0.007)	(0.010)
Log size ($\ln \hat{S}$)	-1.006***	-1.105***	-1.186***	-1.276***	-1.341***	-1.418***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.009)	(0.013)
	[0.038]	[0.045]	[0.053]	[0.064]	[0.077]	[0.100]
R^2	0.941	0.958	0.967	0.974	0.973	0.969
No. of Obs.	1,049	1,200	1,000	800	600	400

Standard errors are in parentheses. Gabaix-Ioannides (2004) corrected standard errors, $\hat{b}\sqrt{\frac{2}{N}}$, are in brackets. ***: $p < 0.01$.

90th percentile of the sample size 1,623, i.e., between ranks 650 and 1,460. For each rank $R^* \in [650, 1460]$, we run the above regression. The 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 6 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 Table 6. The coefficient of $\ln \hat{S}$ is -1.006, 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 rank-size regression should be increasing as one uses a smaller and smaller sample of the largest cities. He empirically demonstrates that this is true in the 2000 U.S. census data and interprets it as evidence that population sizes of "census places" follow a lognormal distribution. More recent analysis has confirmed that results from the rank-size regression are sensitive to the choice of the cutoff point (Fazio and Modica 2012).

We conduct 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 on U.S. cities using 2000 census data. Note that in all the regressions in Table 6, the R^2 is never lower than 0.94. That is, the straight line always fits very well despite the varying slope.

Figure 7: Log rank again log size for walled cities in Qing Dynasty

(a) Log rank against log size, full sample

(b) Log rank against log size, truncated sample

$N = 1,182$ for panel (a) and 602 for panel (b).

Table 7: Regressions of log rank on log size, Ming Dynasty

	Dependent variable: Log rank ($\ln R$)				
	(1)	(2)	(3)	(4)	(5)
	Rank	Rank	Rank	Rank	Rank
	≤1,000	≤800	≤602	≤400	≤200
Constant	6.533***	6.710***	6.868***	7.009***	6.866***
	(0.009)	(0.010)	(0.014)	(0.023)	(0.048)
Log size ($\ln \hat{S}$)	-0.975***	-1.098***	-1.187***	-1.255***	-1.210***
	(0.008)	(0.008)	(0.009)	(0.013)	(0.021)
	[0.044]	[0.055]	[0.068]	[0.089]	[0.121]
R^2	0.941	0.961	0.965	0.960	0.942
No. of Obs.	1,000	800	602	400	200

Standard errors are in parentheses. Gabaix-Ioannides (2004) corrected standard errors, $\hat{b}\sqrt{\frac{2}{N}}$, are in brackets. ***: $p < 0.01$.

Cities in Ming Dynasty

In figure 7, 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 Table 7. 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 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 -0.975 and -1.210. That is, the rank-size distribution of walled cities in the Ming Dynasty is fairly close to the Zipf's law.

4.3.4 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 distributions (e.g., Gabaix 1999; Eeckhout 2004; Duranton 2006; Rossi-Hansberg and Wright 2007; and Córdoba 2008). Using any version of these "random growth models" to explain the physical size distribution of walled cities in the late imperial China would face this problem: City walls last for hundreds of years and thus city sizes as measured in this study rarely "grow" over a long period of time.[23] 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.

As discussed above, 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. Above, we adapted the argument by Lee and Li (2013) to claim that as long as these factors are random across cities and only weakly correlated, equilibrium city-size distribution would be asymptotically lognormal, were it not for the fact that it is inherent in the need for cities to defend themselves. In the latter case, simple intuition suggests that as long as defense costs include a fixed component, very small cities would be infeasible and instead of the lognormal, we end up with Pareto distribution for city sizes. This, we claim, is a novel justification of the Pareto distribution.

Our simple model generates the power law distribution from the power law distribution of the productivity parameter R. As noted above, the power law distribution of R may result from an accumulation of random productivity shocks over time, or simply reflects a distribution of natural advantages over space. That is, the way we model the 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).

Hsu (2012) formalizes the central place theory using an equilibrium entry model in

---

[23]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.

which cities produce different goods to satisfy the demands of farmers that are uniformly distributed over space. The heterogeneity in scale economies among different goods generates a hierarchical system of cities of different sizes in equilibrium. Under fairly reasonable assumptions about the distribution of scale economies, Hsu proves that in equilibrium there exists only one next-layer city between neighboring larger cities. This "central place property" of the equilibrium leads to a power law distribution of city sizes. Hsu's model allows for a slight concavity in the log-rank-log-size plot without implying a lognormal distribution of all city sizes. This seems to provide an alternative and parsimonious explanation of our findings here.[24] Note that Hsu (2012) remains to be tested empirically. We suspect that our data from the late imperial China may be useful in devising a test of Hsu's model, because the Chinese administrative system seems to give a best example of a "central place hierarchy" featured in Hsu's model.

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, 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. The second 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 city wall is increasing in jurisdiction population and that city wall quality are increasing in the risk of being attacked (proxied by the location of the city relative to different frontiers). Both results are consistent with the predictions of our model.

More importantly, we use these data to explore in greater depth the physical size distribution of walled cities. Existing literature on city size distribution 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

---

[24]Interestingly, G. William Skinner had always 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). He himself actually tried to formulate a "central place theory" to model this spatial structure (Skinner 1977c).

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.

References

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

[2]   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.

[3]   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.

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

[5]   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.

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

[7]   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.

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

[9]   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.

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

[11] Eeckhout, Jan. 2009. "Gibrat's Law for (All) Cities: Reply." *American Economic Review* 99, 1676–1683.

[12] Fazio, Giorgio, and Marco Modica. 2012. "Pareto or Log-Normal? A Recursive-Truncation Approach to the Distribution of (All) Cities." Working Papers no. 2012_10, Business School - Economics, University of Glasgow.

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

[14] 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.

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

[16] 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.

[17] Henderson, J. Vernon. 1988. *Urban Development: Theory, Fact and Illusion.* Oxford: Oxford University Press.

[18] 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.

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

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

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

[22] Krugman, Paul. 1995. *Development, Geography, and Economic Theory.* Cambridge, MA: MIT Press.

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

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

[25] Levy, Moshe. 2009. "Gibrat's Law for (All) Cities: Comment." *American Economic Review* 99, 1672–1675.

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

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

[28] 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.

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

[30] 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.

[31] 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.

[32] 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.

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

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

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

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