Crowning the Metropolis: Skylines, Land Values, and Urban Population

David Albouy[*]
University of Illinois and NBER

Mauricio Arango
University of Illinois

Minchul Shin
Federal Reserve Bank of Philadelphia

August 10, 2020
PRELIMINARY - Do not cite

[*]We thank Gabriel Ahlfeldt, Xavier Gabaix Laurent Gobillon, Steven Malpezzi, and Dan McMillen for inspiration and assistance. Please e-mail any questions or comments to albouy@illinois.edu.

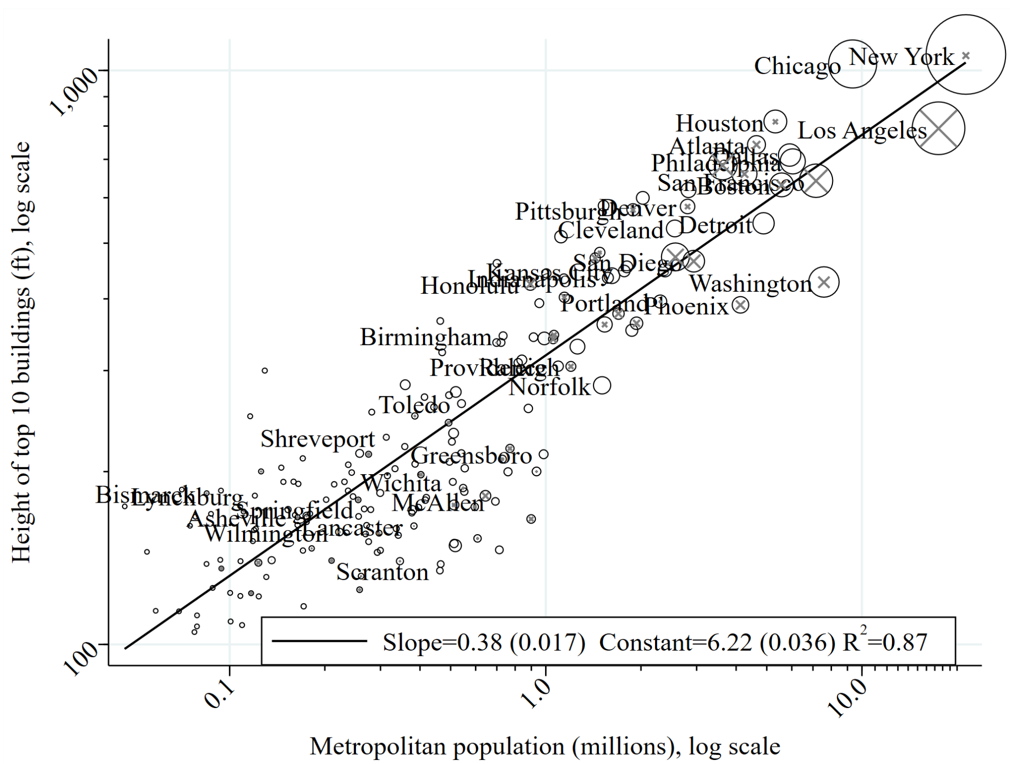
Abstract

In the U.S., the height of cities' tallest buildings is strongly correlated with their greater metropolitan area's population. This is explained through land prices, which rise proportionally with population and income in a monocentric city model, while decreasing proportionally with the arc at which a city can expand. These prices in turn raise building heights less than proportionally through a production function for skyscrapers, mitigated by construction costs and land-use regulations. Using a system of recursive simultaneous equations, we endogenize income with agglomeration economies and test these economic relationships, providing a novel and intrinsically interesting instrumental variables framework for skyscraper heights.

1 Introduction

A curious (and novel) fact about U.S. cities is that the height of their tallest buildings predicts the population of their entire metropolitan area quite accurately. The correlation coefficient between population and the average height of the tallest ten buildings – both expressed in logarithms – is 0.93. The regression line, seen in Figure 1, finds that a 10 percent higher population predicts almost 4 percent taller buildings across most of the sample. This reflects that skyscrapers vary in range by a factor of 10 — from one hundred to one thousand feet — while metro populations vary by a factor 400 — from 50 thousand to 20 million. Conversely, by observing the height of a city's skyline, one can predict fairly accurately how many people would potentially commute there.

Figure 1: Skyscraper Heights and Metropolitan Population

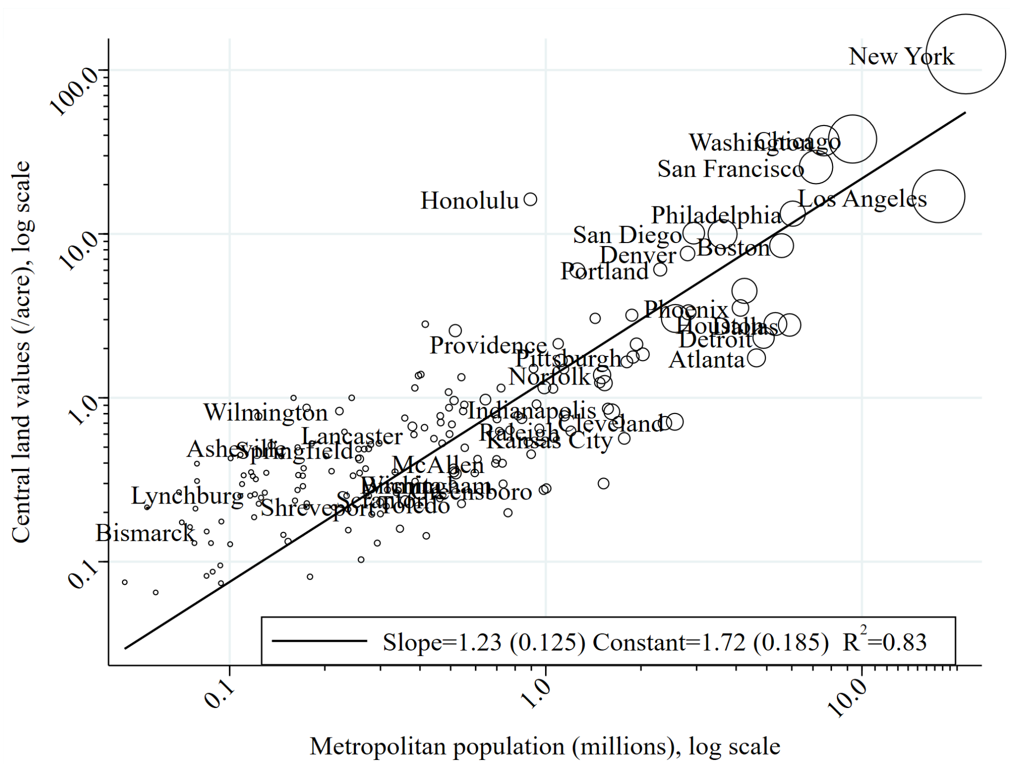


Metropolitan areas refers to 2010 consolidated metropolitan statistical areas (CMSAs). Heights from skyscraper-page.com for 2010. Standard errors in parentheses. Regression weighted by inverse population rank. The size of the marker is proportional to the metro population times the arc of expansion a city can develop away from its center, as justified in the text. Markers with an "X" have a Wharton regulatory index for their central city above the national average: the greater the proportion filled in, the higher the index.

Of course no policy or law – neither legal or natural – ordains a close connection between metro population and building heights. Instead, we propose the two quantities are connected by prices, namely land values. More populous cities have greater central land values, Those greater land values cause builders to build higher, i.e. use more capital, to provide floor space to customers.

The central land value estimates provided by Albouy et al. (2018) are the first to allow researchers to tie this urban economic relationship together.[1]. Plotted against metropolitan population in figure 2, we see central land values vary by a factor of roughly a thousand — from as low as one hundred thousand an acre, to a hundred million. While these value numbers are somewhat noisy, the correlation coefficient is still quite high at 0.91, with a regression slightly greater than one.

Figure 2: Central Land Values and Metropolitan Population



Central land values reported in Albouy et al. (2018) refers to an area of 1 mile radius around city hall. In large CMSAs with multiple centers, the maximum is shown. Standard errors in parentheses. All regressions are weighted by the inverse of the population rank. Metropolitan population and size of marker explained in figure 1

The fact that metro population and central land values track each so closely is consistent with the canonical monocentric city model of **?**, Mills (1967) and Muth (1969). Central land values are higher in larger cites as they provide residents a greater savings from commuting all the way to

[1]Previous estimates have focused on residential land, which is typically much more peripheral.

the city's edge, where land is almost uniformly cheap. In fact, we argue below that a coefficient just above one is consistent with a version of the model outlined in an unpublished early version of Combes et al. (2016), also taking into account enodogenous urban agglomeration.[2]

This monocentric model implies that land values should fall proportionally with a city's arc of expansion from its center: fixing population, a city centered on a straight coast (like Chicago), should have double the land values of a city on an open plain. This echoes some of the more recent work on city shape such as Harari (2020). Using a less known measure of this arc by Malpezzi (????), the rather tight prediction finds support in the data.[3]

The relationship between metro population and land values provides an economic motivation for an instrumental variable estimation strategy, illustrated in Figure 3. This shows how skyscraper height is related to land values. If metro population alone does not influence building heights, except through greater land values, than it may serve as an instrumental variable (IV) for them. Since we expect land values to measured with greater error than population, it is sensible that the IV coefficient is greater than the standard ordinary least squares (OLS). Moreover, as argued in Ahlfeldt and McMillen (2018), this slope may provide a clue on the elasticity of substitution builders face between land and capital in construction for providing floor space. The arc of expansion, which lowers land values, may also provide an additional IV for estimating this parameter if proper identification restrictions hold, which may be tested for using an overidentification test. Namely a metro area's arc of expansion should lower building heights by almost the same degree that population raises them.

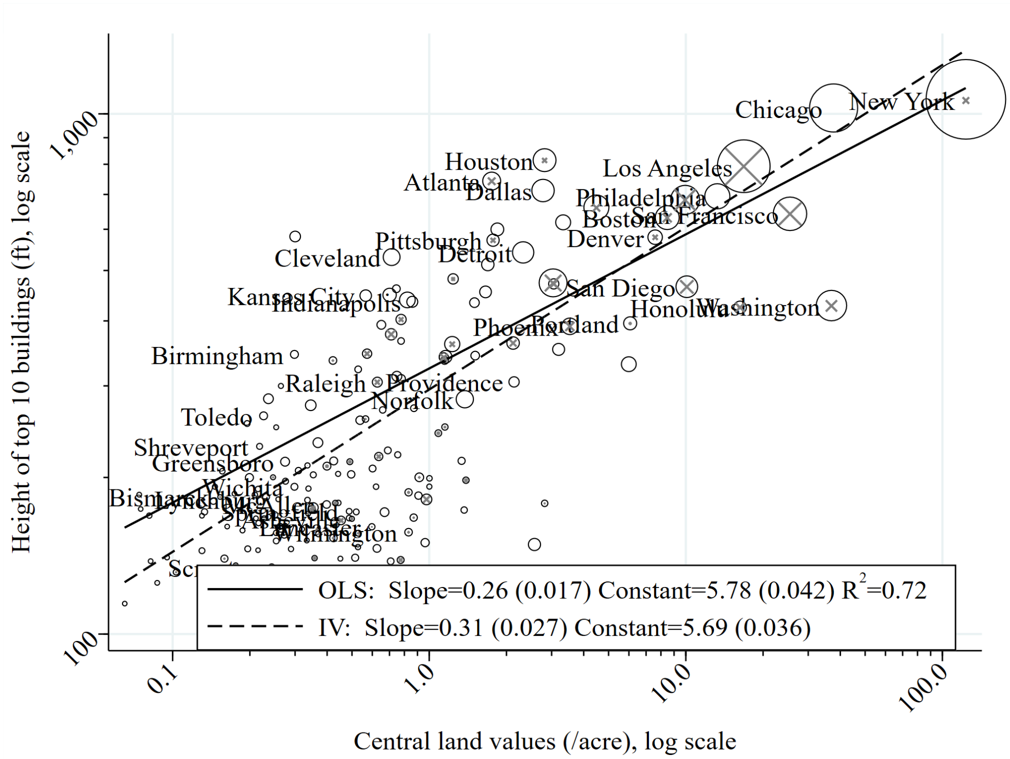
The body of the paper fleshes out the structural and econometric model described here, adding a few flourishes. First, through urban agglomeration economies in production, income depends on metro population endogenously. This makes land values rise more than proportionately with population. Second, construction costs should rise with construction wages, which tied to local income, and thus population. With data across cities, this is potentially important, as construction costs vary widely. Third, land-use regulation imposed by local governments may lower building heights relative to the height predicted by optimising behaviour. Indeed, this seems to explains some of the variation in outliers in figures 1 and 3.

In all we build a recursive system of four econometric equations with three key estimated parameters, and up to five testable restrictions based on the structural model. To extend the analysis, we reconsider the relationship using Zipf's Law for cities to instrument for metro population.

[2]Technically, this is Combes et al. (2012).

[3]Using The measure is slightly different than that measured by Albert (2010), which is the fraction of land over water or with a slope greater than 15 percent within 50 kilometers. Ours is more focused on the city center.

Figure 3: Central Land Values and Metropolitan Population



IV estimate from instrumenting central land values with population. Building heights are from roof to base, leaving aside spires and antennas.

2 Data

To get overlapping data across all domains, we restrict the analysis to cover the cross section of metropolitan areas in 2010. The data are organized into 181 metropolitan areas for which we have data from all sources.

Metropolitan areas refer to 2010 consolidated metropolitan statistical areas (CMSAs). Population numbers and mean family income are taken from the Census and American Community Survey.

Building heights for large cities are taken from skyscraperpage.com, for smaller cities we use data form emporis.com. For skyscraper height, we take the average height of the top ten tallest buildings in each metropolitan area, although we consider alternative groupings. Since we do not have land values across cities for more than a few years, as a control, we also collect the age of the building.

Central land values reported in Albouy et al. (2018) refers to an area of 1 mile radius around city hall. In large CMSAs with multiple centers, the maximum is used.

Wharton regulatory index is taken for the central city only, as opposed to the more common index. While the response rate to this survey was well below 100 percent, the data on central cities in our dataset is much more complete.

Construction costs are provided by R.S. Means and cover both materials and installation costs.

3 Structural Model

The structural model is a recursive system of four equations determining each metro area $j's$ income, m^j; construction costs, v^j, central land values r^j, and building heights H^j. It takes each metropolitan population N^j and arc of expansion Θ^j as exogenous. To follow the trail of causality intuitively, we present this system in reverse order.

3.1 Monocentric City Structure

Cities expand from a central business district (CBD), where the skyscrapers are built. Due to data limitations, we do not try to use land value variation within the CBD. Away from the CBD, land is available for residential purposes along an effectively endless disc, which can expand only along a fixed (exogenous) arc of expansion, Θ_j, which measured in radians may be as high as 2π; cities with their CBD on a linear coast have $\Theta_j = \pi$.

3.2 Determining Skyscraper Height

Skyscrapers are built in the central business district of a monocentric city.[4] Builders there take the price of central land, $r_j(0)$, and construction costs, v_j. The production function for skyscrapers is given by

$$d \ln H_j = \sigma^* \left(d \ln r_j(0) - d \ln v_j + d \ln B_j \right) \tag{1}$$

The * is to denote that this is not the true elasticity of substitution, but rather a transformation. This number is altered by changes in land footprint, as well as changes in the marginal of building square footage. Using data from Chicago, Ahlfeldt and McMillen (2018) find that accounting for

[4]We assume all buildings have the same shape, so there is a one-to-one relation between floor area and building height.

these increases the estimated elasticity of substitution by roughly 30 percent. [5] B_j captures any differences in factor bias — favouring capital over land — that might differ across cities.

3.3 Determining Central Land Values

The CBD abuts land, and may even include, land used for residential purposes. Residents consume housing, non-housing goods, according to Cobb-Douglas preferences, implying a fixed housing expenditure share. They also pay commuting costs which fall with distance, z from the CBD according to a (similar) power law. Land values fall with distance z from the CBD to compensate households for commuting costs. Housing producers generate housing services from land and construction costs according to a Cobb-Douglas technology – at least as a "close enough" approximation. This implies a cost share of land, and in turn a fixed derived share of household expenditures spent on land. [6]

The assumptions imply that central land values increase proportionally with population and income, and decrease proportionally, with the arc of expansion:

$$d \ln r(0) = d \ln N + d \ln m - d \ln \Theta - d \ln F \tag{2}$$

The additional F or reductions in commuting costs, captured in F, which would lower them.[7]

Intuitively, this equation relies on the idea that land values are proportional to aggregate income in a Cobb-Douglas economy. The multiplicative commuting disutility costs keeps the share proportional across the land value gradient. Commuting acts as a quality of life disamenity. Higher incomes should raise land values.

3.4 Endogenous Construction Costs

The model also requires addressing local income and construction costs, which may be endogenous.[8] If construction inputs consists of local labor costs in installation, while material costs are uniform,

[5] As shown in the appendix $\sigma^* = \sigma(1 + \zeta - \xi)^{-1}$, where ζ is the elasticity of capital to floorspace, with respect to building height, and ξ is the elasticity of the land relative to the footprint of the building.

[6] These are clearly approximations. Evidence from and imply that the elasticities are probably close enough to one for this to be a workable approximation.

[7] The appendix shows that increases in the expenditure share on housing s, would also raise central land values proportionally. An elasticity of substitution of less than one, would make this higher in more expensive cities, which would then bias the coefficient on value raising characteristics, such as population, upwards.

[8] The appendix also covers the case where they are exogenous.

then we expect the two to be related through a cost function

$$d \ln v = a d \ln m - d \ln A_v \tag{3}$$

The parameter a is related closely to the cost share of land, while $d \ln A_v$ captures any potential productivity shifters in costs. [9]

3.5 Endogenous Income

Urban economists have long studied urban agglomeration economies that cause incomes to rise. The standard relationship between income and population is generally given by the power function

$$d \ln m = \gamma d \ln N + d \ln A_m \tag{4}$$

where γ is the agglomeration parameter and $d \ln A_m$ account for productivity differences unaccounted for by population.

Providing new ways of estimating γ is not the goal here. Rather, that if population and arc of expansion may be taken as exogenous, a recursive system of equation follows. Population raises income in equation (4), and through it construction costs in (3) by the elasticity $a\gamma$ and land values in (2) by the elasticity $1 + \gamma$. Finally, the building heights equation (1) takes into account both of these channels, producing an elasticity of $\sigma^*(1 + \gamma - a\gamma$. The chain for the arc of expansion is simpler as it only affects land values, and through them building heights.

3.6 Incorporating Zipf's Law to Instrument for Population

Finally, it is worth addressing the endogeneity of population.

$$d \ln N^j = d \ln Z^j + e_Z^j \tag{5}$$

where Z^j is the inverse of the metro area's rank in the population distribution. e_Z^j is an error term. If we assume this is a relationship,

[9]Note Cobb-Douglas technology provides an exact first-order approximation in logarithms for any cost function.

4 Econometric Model

The econometric model may be presented and tested through an unrestricted reduced form of four equations. The two key exogenous variables provide eight parameters, which in the strictest version may be explained by only three parameters. Following the standard Cowles notation $\mathbf{BY} = \mathbf{\Pi Z} + \varepsilon$, denote the rectangular system as

$$\ln H^j = \pi_{HN} \ln N^j + \pi_{H\Theta} \ln \Theta^j + X^j \beta_H + \varepsilon_H^j \tag{6a}$$

$$\ln r(0)^j = \pi_{rN} \ln N^j + \pi_{r\Theta} \ln \Theta^j + X^j \beta_r + \varepsilon_r^j \tag{6b}$$

$$\ln v^j = \pi_{vN} \ln N^j + \pi_{v\Theta} \ln \Theta^j + X^j \beta_v + \varepsilon_v^j \tag{6c}$$

$$\ln m^j = \pi_{mN} \ln N^j + \pi_{m\Theta} \ln \Theta^j + X^j \beta_m + \varepsilon_m^j \tag{6d}$$

where X^j are control variables that include a constant. The model is overidentified as there are eight free parameters and only three structural parameters. The reduced form parameters are

$$\pi_{HN} = \sigma^*(1 + \gamma - a\gamma) \qquad \pi_{H\Theta} = -\sigma^* \tag{7a}$$

$$\pi_{rN} = 1 + \gamma \qquad \pi_{r\Theta} = -1 \tag{7b}$$

$$\pi_{vN} = a\gamma \qquad \pi_{v\Theta} = 0 \tag{7c}$$

$$\pi_{mN} = \gamma \qquad \pi_{m\Theta} = 0 \tag{7d}$$

which reading upwards relays the recursive structure of the model.

The testable restrictions may be grouped according to their interest and importance in identifying key relationship. Arguably two most interesting and demanding restrictions regard land values, which should fall proportionally with the arc of expansion, $\pi_{r\Theta} = -1$ and rise by slightly more due to the agglomeration parameter $\pi_{r\Theta} = 1 + \gamma$, which is typically small. Indeed these restrict parameters to have an exact value, implying a strong degree of external validity, and not just that the resemble each other.

The other key restriction is through the building heights relationship. It is non-linear and covers three equations $\pi_{H\Theta} = (1 + \pi_{mN} - \pi_{vN})\pi_{HN}$. This is equivalent to stating that the elasticity of substitution σ^* is the same using the population instrument as the using the arc of expansion instrument. This is resembles a standard Sargan over-identification test, except that it accounts for endogenous income and construction cost shifts.

There are reasons why this overidentification test may fail to hold. One reason is that a limited arc of expansion may change the demand for building height. Cities with a low arc of expansion may have views that are more, or less desirable.[10] If the goal is to have the building seen from far away, a greater arc of expansion may provide more nearby viewers.

Less critical is the restrictions that the arc of expansion in construction costs and income $\pi_{v\Theta} = 0$. It is possible that constrained cities suffer additional cost factors in construction. Thus, we consider models that relax these last two restrictions. An addition, since $a \leq 1$, we should have that $\pi_{vN} < \pi_{mN}$

The arc of expansion may also be correlated with unaccounted for productivity boosters. This restriction is not essential either way.

Finally, control variables may affect our outcomes, and so we focus on those that might affect skyscraper height. Land-use restrictions would be expected to reduce building height if they create a factor bias limiting capital's productivity. Older buildings may also be shorter, reflecting more expensive building technologies, not to mention a potentially dynamic mismatch due to their durability. As buildings age and land appreciates, they may not reflect fully the value of the land underneath them.

5 Results

The reduced form results are shown in Panel A of Table 1, without controls in columns (1) through (4), corresponding in number to the order for the structural equations determining the outcomes. Controls for the Wharton index are added in the next four columns. The five tests for the restrictions are shown in Panel B, starting with the single coefficient restrictions, and moving to the multiple equation ones.

Starting with income determinants in columns (4) and (8), the estimated coefficient on population is positive, significant and small. The estimated coefficient on the arc of expansion is not significantly different from zero, as predicted.

For construction costs in (3) and (7), the coefficient on the population is less than the one for income, as predicted. However the coefficient on the arc of expansion is negative, meaning that construction costs are higher in more constrained cities, conditional on population.

The key land value equations in columns (2) show the relationship seen in Figure 2 augmented with additional variables. Remarkably, the coefficient on the arc of expansion is very close to minus one, as predicted. The coefficient on population is above one by an amount fairly close to the population elasticity for income. The closeness of the estimates to these tight predictions is remarkable.

[10]While views of nature may be in higher demand, it is contestable whether this is true of business operations.

Table 1: Determinants of Skyscraper Heights and Land Values
Reduced Form Regression Results and Structural Test

Outcome	Skyscr. height $\ln H^j$ (1)	Land value $\ln r^j$ (2)	Const cost $\ln v^j$ (3)	Hhold income $\ln m^j$ (4)	Metro pop. $\ln N^j$ (5)	Skyscr. height $\ln H^j$ (6)	Land value $\ln r^j$ (7)	Const cost $\ln v^j$ (8)	Hhold income $\ln m^j$ (9)	Metro pop. $\ln N^j$ (10)
Panel A: Reduced-Form Regression Results										
Pop. rank, inverse $\ln Z^j$	0.362	1.004	0.050	0.087	0.958	0.341	0.990	0.053	0.083	0.920
	(0.038)	(0.073)	(0.009)	(0.008)	(0.044)	(0.039)	(0.061)	(0.009)	(0.009)	(0.039)
Arc of expan, radians $\ln \Theta^j$	-0.177	-0.756	-0.074	0.016	0.085	-0.081	-0.813	-0.078	0.017	0.138
	(0.227)	(0.456)	(0.044)	(0.045)	(0.118)	(0.148)	(0.475)	(0.044)	(0.036)	(0.085)
Wharton Regulatory index, central city						-0.095	0.236	-0.009	0.027	0.092
						(0.035)	(0.096)	(0.011)	(0.020)	(0.031)
Age of bldgs, years						-0.480	0.396	0.011	0.022	-0.153
						(0.167)	(0.293)	(0.035)	(0.043)	(0.056)
R^2	0.739	0.841	0.541	0.700	0.968	0.789	0.849	0.541	0.710	0.973
# of metros	262	262	262	262	226	226	226	226		
Panel B: Tests of the Structural Model										
I: $\pi_{NZ} = 1$	0.3309					0.0355				
II: $\pi_{rZ} = \pi_{NZ} + \pi_{mZ}$	0.5313	Joint Test of I-VII (Tight Model)				0.8393	Joint Test of I-VII (Tight Model)			
III: $\pi_{r\Theta} = -1$	0.5906	0.0198				0.6905	0.0008			
IV: $\pi_{N\Theta} = 0$	0.4697					0.1019				
V: $\pi_{m\Theta} = 0$	0.7179	Joint Test of I-V (Loose Model)				0.6334	Joint Test of I-V (Loose Model)			
VI: $\pi_{v\Theta} = 0$	0.0942	0.2494				0.0741	0.0867			
VII $\pi_{H\Theta}(\pi_{NZ} + \pi_{mZ} - \pi_{vZ}) = \pi_{HZ}\pi_{NZ}$	0.5113					0.1548				

Table 2: Determinants of Skyscraper Heights and Land Values Using Log Population
Reduced Form Regression Results and Structural Test

Outcome	Skyscr. height $\ln H^j$ (1)	Land value $\ln r^j$ (2)	Const cost $\ln v^j$ (3)	Hhold income $\ln m^j$ (4)	Skyscr. height $\ln H^j$ (5)	Land value $\ln r^j$ (6)	Const cost $\ln v^j$ (7)	Hhold income $\ln m^j$ (8)
Panel A: Reduced-Form Regression Results								
Population $\ln N^j$	0.386	1.001	0.046	0.088	0.383	1.021	0.051	0.086
	(0.036)	(0.071)	(0.010)	(0.009)	(0.036)	(0.075)	(0.011)	(0.010)
Arc of expan, radians $\ln \Theta^j$	-0.187	-0.966	-0.094	0.001	-0.109	-1.066	-0.098	-0.003
	(0.207)	(0.491)	(0.049)	(0.045)	(0.131)	(0.506)	(0.048)	(0.038)
Wharton Regulatory index, central city					-0.134	0.155	-0.012	0.020
					(0.030)	(0.102)	(0.012)	(0.020)
Age of bldgs, years					-0.413	0.513	0.014	0.032
					(0.171)	(0.316)	(0.035)	(0.046)
R^2	0.784	0.816	0.486	0.691	0.835	0.820	0.490	0.689
# of metros	262	262	262	262	226	226	226	226
Panel B: Tests of the Structural Model								
I: $\pi_{rN} = 1 + \pi_{mN}$	0.2055	Joint Test of I-V (Tight Model)			0.3408	Joint Test of I-V (Tight Model)		
II: $\pi_{r\Theta} = -1$	0.9448	0.0062			0.8944	0.0120		
III: $\pi_{m\Theta} = 0$	0.9750	Joint Test of I-III (Loose Model)			0.9427	Joint Test of I-III (Loose Model)		
IV: $\pi_{v\Theta} = 0$	0.0517	0.2249			0.0380	0.7333		
V: $\pi_{H\Theta}(1 + \pi_{mN} - \pi_{vN}) = \pi_{HN}$	0.4383				0.1019			

The equation for building heights shows results that are precise for population, but less clear for the arc of expansion. The cross equation restriction passes at a p-value of 5 or 10 percent without controls, but not with them. This suggests that while the arc of expansion may be a good predictor of land values, it may not be a good instrument for land values in the building heights equation, if population is. Alternatively, the Sargan test suggests the two potential IVs have limited value.

The control variables do not appear to matter in the outcome predictions, except where they were most expected, namely in affecting building height. Greater land-use restrictions and building age both lower height, as expected.

Table 3 presents the structural estimates of the parameters using Generalized Method of Moments (GMM). It presents estimates using two sets of restrictions. The columns with "Tight" restrictions (1) and (3) uses all five restrictions, while the columns with "Loose" restrictions allow the arc of expansion to impact construction costs and building heights.

Table 3: Structural Estimation Results Using Population rank

	Controls		No controls	
Restrictions	Tight (1)	Loose (2)	Tight (3)	Loose (4)
Population elasticity of income	0.076	0.075	0.093	0.092
$\gamma = \pi_{mn}$	(0.004)	(0.004)	(0.005)	(0.005)
Const. cost share of labor	0.804	0.851	0.605	0.512
$a = \pi_{vn}/\pi_{mn}$	(0.077)	(0.116)	(0.045)	(0.076)
Elasticity of substitution on bldg hgt	0.309	0.321	0.342	0.374
$\sigma = \pi_{hn}/(1 + \pi_{mn} - \pi_{vn})$	(0.018)	(0.020)	(0.014)	(0.025)
Arc of expansion on height		0.317		0.247
		(0.113)		(0.136)
Arc of expansion on const. cost		-0.083		-0.077
		(0.041)		(0.033)
Overidentification test	0.0013	0.1818	0.0045	0.2036

In the tight model, all constraints hold ($\pi_{r\Theta} = -1$, $\pi_{v\Theta} = 0$, $\pi_{m\Theta} = 1$, $\pi_{rN} = 1 + \pi_{mN}$, and $\pi_{hN} = -\pi_{h\Theta} * (1 + \pi_{mN} - \pi_{vN}))$.

The implied agglomeration parameter, γ is around 9 percent. This is slightly on the high side relative to most recent estimates, but not implausible for this parameter. The labor share in construction parameter a is in most specifications close to one half, which is not far from most accounting procedures.

The elasticity of substitution in building height is near 0.35 in almost all of the specifications. This is close to comparable estimates based on land-value variation in Chicago estimated by Ahlfeldt and McMillen (2018). Taking their estimates to convert this into a true elasticity of substitution of capital and land produces values slightly above 0.5.

6 Alternative Empirical Approaches

To be filled in.

6.1 Taking Income as Exogenous

6.2 Commuting times

6.3 Instrumenting with Pure Amenities

6.4 Non-linearities

7 Conclusion

We demonstrate that tight the relationship between skyscraper height and land values is explained through land values by marrying two canonical models. The first is a monocentric city model that predicts that central land values increase proportionally with aggregate income, reflecting potential savings in commuting costs. We test a particular variant that restricts the relationship to be one to one. Furthermore, we test the prediction that central land values fall proportionally with a city's arc of expansion, namely the degree to which geography allows metro areas to expand radially from the center in all directions. Remarkably, the data appear consistent with the model's exact one-to-one predictions despite its radical simplicity.

The second model relates land values to building heights through a production relationship. The magnitude of this relationship is determined by the elasticity of substitution between land and capital in production. This parameter is unconstrained, although the data produce an estimate consistent with the literature. More interestingly, the two models combined justify a testable instrumental variable model. If the exclusion restrictions hold, a metro area's arc of expansion should lower building heights to the same degree that aggregate income raises them. In other words, skyscrapers are taller in places where geography makes it hard for cities to sprawl, and this operates through higher land values. While this prediction proves to be less than perfect, overall the model does well in light of a battery of tests, and performs well with controls.

Furthermore, the model provides a tractable and plausible system of equation's that tracks several important economic urban economic phenomena together. It highlights visible features of cities, taking together several complex phenomena, including urban agglomeration economies in a transparent framework. Thus, we hope it provides a good starting point for teaching and learning as well as future research.

Appendix

7.1 Economic Model

7.1.1 Land Value Determination

Based on mobility within city, the equilibrium price of housing at distance z from the center falls in proportion to the increase in commuting costs, $f(z)$

$$p(z) = p(0)\left[\frac{f(0)}{f(z)}\right]^{\frac{1}{s}}$$

Housing prices relate to land price via the following unit cost function, where ϕ is the cost share of land in housing:

$$p(z) = \frac{[r(z)]^{\phi}\, v^{1-\phi}}{A_Y \phi^{\phi}(1-\phi)^{1-\phi}}$$

Housing demand is

$$N(z)y(z) = N(z)s\frac{m}{p(z)} = N(z)\frac{sm}{p(0)}\left[\frac{f(z)}{f(0)}\right]^{\frac{1}{s}} = N(z)\frac{smA_Y\phi^{\phi}(1-\phi)^{1-\phi}}{[r(0)]^{\phi}\, v^{1-\phi}}\left[\frac{f(z)}{f(0)}\right]^{\frac{1}{s}}$$

Note that the price gradient does not depend heavily on the assumption of constant income and cost shares: they are valid first-order approximations. The demand and supply quantities depend on the elasticity of substitution even in a first-order approximation, making them much more dependent on the Cobb-Douglas assumption.

Land values are determined by the derived demand for housing

$$r(z) = [p(z)A_Y]^{\frac{1}{\phi}}\,\phi\left(\frac{1-\phi}{v}\right)^{\frac{1-\phi}{\phi}} = \underbrace{[p(0)A_Y]^{\frac{1}{\phi}}\,\phi\left(\frac{1-\phi}{v}\right)^{\frac{1-\phi}{\phi}}}_{r(0)}\left[\frac{f(z)}{f(0)}\right]^{-\frac{1}{s\phi}}$$

Housing supply at distance z

$$Y(z) = z\Theta A_Y\left[\frac{r(z)}{v}\right]^{1-\phi}\left(\frac{1-\phi}{\phi}\right)^{1-\phi} = z\Theta A_Y\left[\frac{r(0)}{v}\right]^{1-\phi}\left(\frac{1-\phi}{\phi}\right)^{1-\phi}\left[\frac{f(z)}{f(0)}\right]^{-\frac{1-\phi}{s\phi}}$$

Taking the ratio of supply to demand to solve for the population at distance z

$$N(z) = \frac{Y(z)}{y(z)} = r(0)\frac{\Theta z}{sm}\left[\frac{f(z)}{f(0)}\right]^{-\frac{1}{s\phi}}$$

I think the equation above is missing ϕ, see equation below (check ϕ in $Y(z)$ and $y(z)$, (8) is not affected by this Mauricio)

$$N(z) = \frac{Y(z)}{y(z)} = r(0)\frac{\Theta z}{sm\phi}\left[\frac{f(z)}{f(0)}\right]^{-\frac{1}{s\phi}}$$

The population at any distance is proportional to the central land value

$$N = \int_0^\infty N(z)dz = r(0)\frac{\Theta}{sm}\underbrace{\int_0^\infty\left[\frac{f(z)}{f(0)}\right]^{-\frac{1}{s\phi}} zdz}_{\equiv F}$$

Remarkably the commuting cost function may be taken out if commuting cost structures are similar across cities.[11] Solving for $r(0)$ and taking the differential provides the following cost-based equation

$$d\ln r(0) = d\ln N - d\ln\Theta + d\ln m + d\ln s - d\ln F \tag{8}$$

7.1.2 Building Heights

Neither building height function nor floorspace is equivalent to capital. Rather, assume that capital investment rises proportionally with floor space F, but increases with building height as governed

[11] If there is an outside option

by the elasticity ζ

$$d\ln K = d\ln F + \zeta d\ln H \tag{9}$$

A building has a certain footprint T, however the land required may rise with height, possibly from zoning codes

$$d\ln L = d\ln T + \xi d\ln H = d\ln F + (\xi - 1)d\ln H \tag{10}$$

where the second equation uses the identity that height is merely floorspace divided by the footprint, $H = F/T \rightarrow T = F/H$ Thus we have that floorspace drops out capital to land ratio is

$$d\ln K - d\ln L = (1 + \zeta - \xi)d\ln H$$

We assume that skyscrapers compete for land with central residences. Unlike the latter, allow building of skyscrapers to follow a more flexible technology, with the elasticity of substitution between capital and land given by σ

$$d\ln K - d\ln L = \sigma(d\ln r(0) - d\ln v + d\ln B) \tag{11}$$

and $d\ln B$ captures any differences due to factor bias. $\ln v$ is a construction cost index.

We then have

$$d\ln H = \underbrace{\frac{\sigma}{1 + \zeta - \xi}}_{\sigma^*}(d\ln r(0) - d\ln v + d\ln B) \tag{12}$$

We lack the data to estimate beyond σ^*.

Substituting in the land value equation, provides a benchmark for how population raises and the arc of expansion lowers building heights.

$$d\ln H = \sigma^*(d\ln N - d\ln\Theta + d\ln m - d\ln v + d\ln B + d\ln s - d\ln F) \tag{13}$$

One difficulty is that incomes and construction costs are likely collinear due to similarities in pay differentials. To a lesser extent, population and income may also be collinear due to agglomeration economies.

Endogenous income

$$\ln m = \gamma\ln N + \ln A_m \tag{14}$$

Endogenous construction cost

$$\begin{aligned}\ln v &= a \ln m + \ln A_v \\ &= \gamma a \ln N + a \ln A_m + \ln A_v\end{aligned}$$

7.2 Econometric Model

7.2.1 Structural Model - Endogenous Income and Construction Costs

The structural equations $\mathbf{BY} + \mathbf{\Gamma Z} = \mathbf{U}$ may be given as (check)

$$\begin{bmatrix} 1 & -\sigma^* & \sigma^* & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -a \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \ln H \\ \ln r(0) \\ \ln v \\ \ln m \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -1 & 1 \\ 0 & 0 \\ -\gamma & 0 \end{bmatrix} \begin{bmatrix} \ln N \\ \ln \Theta \end{bmatrix} = \begin{bmatrix} d \ln B \\ \ln s + d \ln F \\ d \ln A_v \\ d \ln A_m \end{bmatrix}$$

Some of the controls may be used to absorb the residual.

7.2.2 Reduced Form

$\mathbf{Y} = \mathbf{\Pi Z} + \mathbf{V}$, where $\mathbf{\Pi} = -\mathbf{B^{-1}\Gamma}$, which means that the reduced form coefficient matrix is

$$\begin{bmatrix} \pi_{HN} & \pi_{H\Theta} \\ \pi_{rN} & \pi_{r\Theta} \\ \pi_{vN} & \pi_{v\Theta} \\ \pi_{mN} & \pi_{m\Theta} \end{bmatrix} = \begin{bmatrix} -1 & -\sigma^* & \sigma^* & -(1-a)\sigma^* \\ 0 & -1 & 0 & -1 \\ 0 & 0 & -1 & -a \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ -1 & 1 \\ 0 & 0 \\ -\gamma & 0 \end{bmatrix} = \begin{bmatrix} (1+\gamma-a\gamma)\sigma^* & -\sigma^* \\ 1+\gamma & -1 \\ a\gamma & 0 \\ \gamma & 0 \end{bmatrix}$$

So we have 5 restrictions and 3 free parameters. The five restrictions may be grouped. First the zero restrictions.

$$\pi_{m\Theta} = 0 \tag{15a}$$

$$\pi_{v\Theta} = 0 \tag{15b}$$

Then for land we need the two, the latter slightly augmented from before (but the right is close to zero:

$$\pi_{r\Theta} = -1 \tag{15c}$$

$$\pi_{rN} + \pi_{r\Theta} = \pi_{mN} \tag{15d}$$

In the building heights there is one non-linear restriction, which lets the

$$\pi_{HN} = -(1 + \pi_{mn} - \pi_{vN})\pi_{H\Theta} \tag{15e}$$

which is close to $\pi_{HN} = -\pi_{H\Theta}$ Some restrictions may be imposed.

The structural parameter estimates are

$$\gamma = \pi_{mN} \tag{16a}$$

$$a = \frac{\pi_{vN}}{\pi_{mN}} \tag{16b}$$

$$\sigma^* = -\pi_{H\Theta} = \frac{\pi_{HN}}{1 + \pi_{mN} - \pi_{vN}} \tag{16c}$$

(16d)

Note imposing $\gamma = 0$, which is not too far from reality, creates the simpler model we did before.

7.2.3 Structural Model - Endogenous Income and Construction Costs

The structural equations $\mathbf{BY} + \mathbf{\Gamma Z} = \mathbf{U}$ may be given as (check)

$$\begin{bmatrix} 1 & -\sigma^* \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \ln H \\ \ln r(0) \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & \sigma^* \\ -1 & 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} \ln N \\ \ln \Theta \\ \ln m \\ \ln v \end{bmatrix} = \begin{bmatrix} \ln s & d\ln F & 0 \\ 0 & 0 & d\ln B \end{bmatrix}$$

Some of the controls may be used to absorb the residual.

7.2.4 Reduced Form - Exogenous Income and Construction Costs

$\mathbf{Y} = \mathbf{\Pi Z} + \mathbf{V}$, where $\mathbf{\Pi} = -\mathbf{B^{-1}\Gamma}$, which means that the reduced form coefficient matrix is

$$\begin{bmatrix} \pi_{HN} & \pi_{H\Theta} & \pi_{Hm} & \pi_{Hv} \\ \pi_{rN} & \pi_{r\Theta} & \pi_{rm} & \pi_{rv} \end{bmatrix} = \begin{bmatrix} -1 & -\sigma^* \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & \sigma^* \\ -1 & 1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} \sigma^* & -\sigma^* & \sigma^* & -\sigma^* \\ 1 & -1 & 1 & 0 \end{bmatrix}$$

The restrictions may be grouped. In the land equation, we can test two equality of coefficients, and then the restriction of unity:

$$\pi_{rN} = -\pi_{r\Theta} \tag{17a}$$

$$\pi_{rN} = \pi_{rm} \tag{17b}$$

$$\pi_{rN} = 1 \tag{17c}$$

For construction costs we have

$$\pi_{rv} = 0 \tag{17d}$$

although measurement problems may make us want to impose this. One possibility is that higher construction costs are associated with worse infrastructure/greater commuting costs.

In the building heights

$$\pi_{HN} = -\pi_{H\Theta} \tag{18a}$$

$$\pi_{HN} = \pi_{Hm} \tag{18b}$$

$$\pi_{HN} = -\pi_{Hv} \tag{18c}$$

So we have 7 restrictions and 1 free parameterm σ^*. Some restrictions may be imposed.

7.2.5 Collinearity of constructions costs

If $\ln m = \ln v$

$$\begin{bmatrix} 1 & -\sigma^* \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \ln H \\ \ln r(0) \end{bmatrix} + \begin{bmatrix} 0 & 0 & \sigma^* \\ -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} \ln N \\ \ln \Theta \\ \ln m \end{bmatrix} = \begin{bmatrix} \ln s & d\ln F & 0 \\ 0 & 0 & d\ln B \end{bmatrix}$$

The reduced form coefficient matrix is

$$\begin{bmatrix} \pi_{HN} & \pi_{H\Theta} & \pi_{Hm} \\ \pi_{rN} & \pi_{r\Theta} & \pi_{rm} \end{bmatrix} = \begin{bmatrix} -1 & -\sigma^* \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 0 & \sigma^* \\ -1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} \sigma^* & -\sigma^* & 0 \\ 1 & -1 & 1 \end{bmatrix}$$

The restrictions are fewer now as they exclude thew zero in the land equation

$$\pi_{rN} = -\pi_{r\Theta} \tag{19a}$$

$$\pi_{rN} = \pi_{rm} \tag{19b}$$

$$\pi_{rN} = 1 \tag{19c}$$

In the building heights, there are only two

$$\pi_{HN} = -\pi_{H\Theta} \tag{20a}$$

$$\pi_{HN} = \pi_{Hm} \tag{20b}$$

In this model there are 5 restrictions and 1 free parameter, σ^*. Some restrictions may still be imposed.

7.2.6 Population and Income Imposed the Same

If $\ln N + \ln m$ is our variable for whatever justification

$$\begin{bmatrix} 1 & -\sigma^* \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \ln H \\ \ln r(0) \end{bmatrix} + \begin{bmatrix} 0 & 0 & \sigma^* \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \ln N + \ln m \\ \ln \Theta \\ \ln v \end{bmatrix} = \begin{bmatrix} \ln s & d\ln F & 0 \\ 0 & 0 & d\ln B \end{bmatrix}$$

The reduced form coefficient matrix is

$$\begin{bmatrix} \pi_{HN} & \pi_{H\Theta} & \pi_{Hv} \\ \pi_{rN} & \pi_{r\Theta} & \pi_{rv} \end{bmatrix} = \begin{bmatrix} -1 & -\sigma^* \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 0 & \sigma^* \\ -1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \sigma^* & -\sigma^* & -\sigma^* \\ 1 & -1 & 0 \end{bmatrix}$$

Three restrictions for land

$$\pi_{rN} = -\pi_{r\Theta} \tag{21a}$$

$$\pi_{rN} = 1 \tag{21b}$$

$$\pi_{rv} = 0 \tag{21c}$$

In the building heights, there are two

$$\pi_{HN} = -\pi_{H\Theta} \tag{22a}$$
$$\pi_{HN} = -\pi_{Hv} \tag{22b}$$

In this model there are 5 restrictions and 1 free parameter, σ^*. Some restrictions may still be imposed.

7.2.7 Population and Income Imposed the Same and No Construction Costs

If there is no meaningful difference in construction costs, just eliminate the last two.

7.2.8 Endogenizing population

Rosen-Roback model, no taxes. Amenities and income provide willingness-to-pay at center.

$$\ln Q + \ln m = s \ln p(0) \tag{23}$$
$$= s(\phi \ln r(0) + (1-\phi)\ln v - \ln A_Y) \tag{24}$$

Or in terms of income, another determinig equation.

$$\ln m = s\phi \ln r(0) + s(1-\phi)\ln v - \ln Q - s\ln A_Y \tag{25}$$

This provides the essential structural equation. We could probably go from here. Note this would bring in two more potentially free parameters, s and ϕ. In a sense, and $\beta_{mv} = s(1-\phi)\beta_{mr} = s\phi$, so that

$$s = \beta_{mr} + \beta_{mv} \tag{26}$$
$$\phi = \frac{\beta_{mr}}{\beta_{mr} + \beta_{mv}} \tag{27}$$

Substituting in all our expressions

$$\ln Q + (\gamma \ln N + \gamma A_m) = s(\phi(\ln N - \ln\Theta + \gamma\ln N + \gamma A_m + \ln s - \ln F) \tag{28}$$
$$+ (1-\phi)(\gamma a \ln N + a \ln A_m - \ln A_v) - \ln A_Y) \tag{29}$$

Rearranging gives the essence of the structure in reduced form equation

$$[s\phi(1+\gamma) - \gamma + s(1-\phi)\gamma a] \ln N = \ln Q + s\phi \ln \Theta \tag{30}$$

$$+ [1 - s\phi - s(1-\phi)a]\gamma \ln A_m \tag{31}$$

$$+ A_Y + (1-\phi)s \ln A_v - s\phi(\ln s + \ln F) \tag{32}$$

Instruments should be amenities and the arc of expansion. We need excludable amenities that are unrelated to productivity in housing or tradeable labor.

Zipf's law (Table 2). Zipf's law implies that

$$\ln N_j = a - \ln(rank(N_j))$$

where $rank()$ returns a rank of input variable. If the population follows Zipf's law perfectly, then what it implies in our context is that the strength (or, quality) of the log population as the instrument for the central land value is exactly the same as those based on the log rank population.

In our data set, Zipf's law holds for the population. Therefore, our empirical results should go through with the log rank population if it is used as an instrument. Results are quite similar to those based on the log population as it should be.

On imposing Zipf's law. Our structural model with Zipf's law.

The econometric model may be presented and tested through an unrestricted reduced form of four equations. The two key exogenous variables provide eight parameters, which in the strictest version may be explained by only three parameters. Following the standard Cowles notation $\mathbf{BY} = \mathbf{\Pi Z} + \varepsilon$, denote the rectangular system as

$$\ln H^j = \pi_{HN} \ln N^j + \pi_{H\Theta} \ln \Theta^j + X^j \beta_H + \varepsilon_N^j \tag{33a}$$

$$\ln r(0)^j = \pi_{rN} \ln N^j + \pi_{r\Theta} \ln \Theta^j + X^j \beta_r + \varepsilon_r^j \tag{33b}$$

$$\ln v^j = \pi_{vN} \ln N^j + \pi_{v\Theta} \ln \Theta^j + X^j \beta_v + \varepsilon_v^j \tag{33c}$$

$$\ln m^j = \pi_{mN} \ln N^j + \pi_{m\Theta} \ln \Theta^j + X^j \beta_m + \varepsilon_m^j \tag{33d}$$

$$\ln N_j = \pi_{NZ} \ln rank(N_j) + \varepsilon_N^j \tag{33e}$$

where X^j are control variables that include a constant. The model is overidentified as there are eight free parameters and only three structural parameters. The reduced form parameters are

$$\pi_{HN} = \sigma^*(1 + \gamma - a\gamma) \qquad \pi_{H\Theta} = -\sigma^* \tag{34a}$$

$$\pi_{rN} = 1 + \gamma \qquad \pi_{r\Theta} = -1 \tag{34b}$$

$$\pi_{vN} = a\gamma \qquad \pi_{v\Theta} = 0 \tag{34c}$$

$$\pi_{mN} = \gamma \qquad \pi_{m\Theta} = 0 \tag{34d}$$

which reading upwards relays the recursive structure of the model.

8 Data

8.1 Building heights

We use data on building heights from skyscraperpage.com. An important part of our approach is having a measure of height consistent with the floor area, that is why we measure heights from the ground to the roof, leaving aside ornamental features as spires and antennas. For part of the sample our source only provided the floor count but not height in feet. For those building we uses the predicted height bases on the floor count. Figure (4) shows the relation between height and the number of floors, our estimate is that the average height of a floor is 11.4 feet.

Table 4: Summary Statistics

	Mean	Std.Dev.	Obs
Height top 10 buildings	217.17	168.77	262
Land values	1.78	8.56	262
Population (millions)	0.83	2.08	262
Arc of expansion (percentage)	0.91	0.19	262
Age top 10 high-rises by 2020	52.41	21.02	261
Construction cost	89.70	9.75	262
Household income	59504.70	9091.29	262
Wharton city index	-0.25	0.90	223
Observations	262		

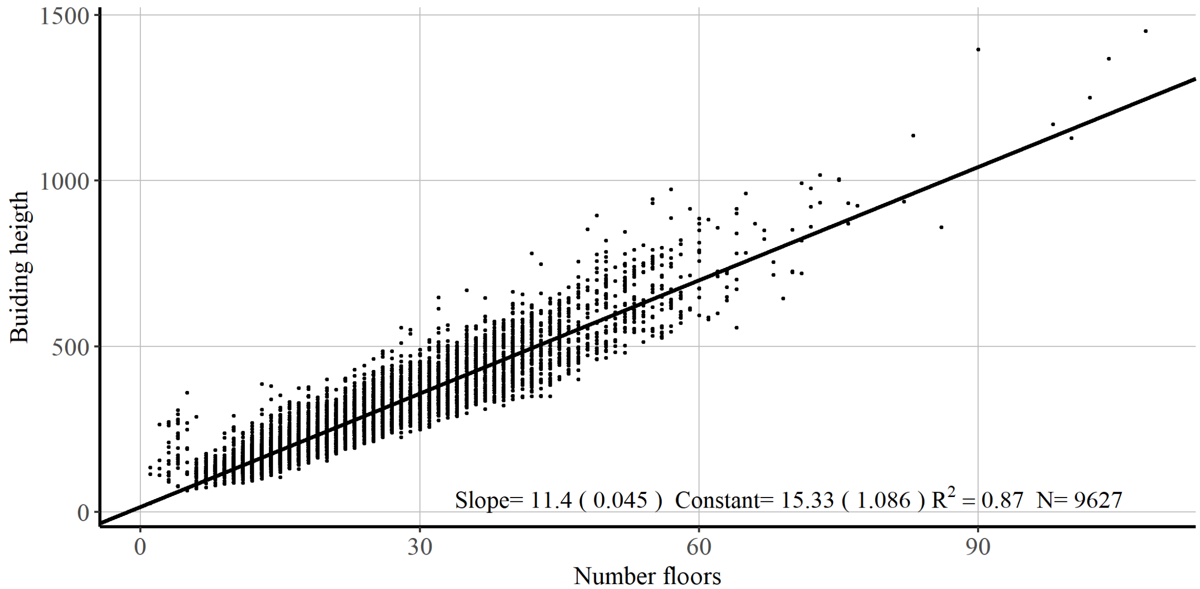
Table 5: Determinants of Skyscraper Heights and Land Values Using Saiz's Land Availability
Reduced Form Regression Results and Structural Test

Outcome	Skyscr. height $\ln H^j$ (1)	Land value $\ln r^j$ (2)	Const cost $\ln v^j$ (3)	Hhold income $\ln m^j$ (4)	Metro pop. $\ln N^j$ (5)	Skyscr. height $\ln H^j$ (6)	Land value $\ln r^j$ (7)	Const cost $\ln v^j$ (8)	Hhold income $\ln m^j$ (9)	Metro pop. $\ln N^j$ (10)
Panel A: Reduced-Form Regression Results										
Pop. rank, inverse $\ln Z^j$	0.386	1.051	0.054	0.084	0.949	0.353	1.054	0.058	0.082	0.910
	(0.030)	(0.042)	(0.009)	(0.006)	(0.048)	(0.037)	(0.045)	(0.009)	(0.007)	(0.042)
Saiz's availability	0.102	-0.530	-0.077	-0.039	-0.010	0.145	-0.474	-0.099	-0.032	0.078
	(0.210)	(0.408)	(0.036)	(0.056)	(0.086)	(0.114)	(0.366)	(0.036)	(0.045)	(0.074)
Wharton Regulatory index, central city						-0.084	0.182	-0.020	0.022	0.099
						(0.040)	(0.120)	(0.011)	(0.018)	(0.031)
Age of bldgs, years						-0.517	0.400	0.025	0.036	-0.150
						(0.182)	(0.260)	(0.031)	(0.047)	(0.061)
R^2	0.739	0.844	0.569	0.714	0.968	0.792	0.847	0.588	0.721	0.973
# of metros	260	260	260	260	224	224	224	224		
Panel B: Tests of the Structural Model										
I: $\pi_{NZ} = 1$	0.2938					0.0313				
II: $\pi_{rZ} = \pi_{NZ} + \pi_{mZ}$	0.6599		Joint Test of I-VII (Tight Model)			0.2392		Joint Test of I-VII (Tight Model)		
III: $\pi_{r\Theta} = -1$	0.2460		≤ 0.0001			0.1454		≤ 0.0001		
IV: $\pi_{N\Theta} = 0$	0.9085					0.2901				
V: $\pi_{m\Theta} = 0$	0.4821		Joint Test of I-V (Loose Model)			0.4688		Joint Test of I-V (Loose Model)		
VI: $\pi_{v\Theta} = 0$	0.0312		0.2912			0.0052		0.0076		
VII $\pi_{H\Theta}(\pi_{NZ} + \pi_{mZ} - \pi_{vZ}) = \pi_{HZ}\pi_{NZ}$	0.0421					0.0006				

Table 6: Determinants of Skyscraper Heights and Land Values Using Log Population and Saiz's Land Availability Reduced Form Regression Results and Structural Test

Outcome	Skyscr. height $\ln H^j$ (1)	Land value $\ln r^j$ (2)	Const cost $\ln v^j$ (3)	Hhold income $\ln m^j$ (4)	Skyscr. height $\ln H^j$ (5)	Land value $\ln r^j$ (6)	Const cost $\ln v^j$ (7)	Hhold income $\ln m^j$ (8)
Panel A: Reduced-Form Regression Results								
Population $\ln N^j$	0.412	1.064	0.051	0.086	0.398	1.106	0.058	0.087
	(0.029)	(0.051)	(0.012)	(0.007)	(0.033)	(0.071)	(0.012)	(0.009)
Saiz's availability	0.113	-0.575	-0.083	-0.041	0.119	-0.592	-0.107	-0.041
	(0.202)	(0.429)	(0.038)	(0.056)	(0.104)	(0.397)	(0.039)	(0.045)
Wharton Regulatory index, central city					-0.126	0.087	-0.024	0.014
					(0.035)	(0.133)	(0.012)	(0.018)
Age of bldgs, years					-0.447	0.503	0.027	0.046
					(0.185)	(0.288)	(0.031)	(0.049)
R^2	0.783	0.812	0.504	0.706	0.837	0.812	0.531	0.703
# of metros	260	260	260	260	224	224	224	224
Panel B: Tests of the Structural Model								
I: $\pi_{rN} = 1 + \pi_{mN}$	0.6525	Joint Test of I-V (Tight Model)			0.7778	Joint Test of I-V (Tight Model)		
II: $\pi_{r\Theta} = -1$	0.3192	≤ 0.0001			0.2986	≤ 0.0001		
III: $\pi_{m\Theta} = 0$	0.4618	Joint Test of I-III (Loose Model)			0.3628	Joint Test of I-III (Loose Model)		
IV: $\pi_{v\Theta} = 0$	0.0272	0.2820			0.0051	0.3237		
V: $\pi_{H\Theta}(1 + \pi_{mN} - \pi_{vN}) = \pi_{HN}$	0.0203				0.0001			

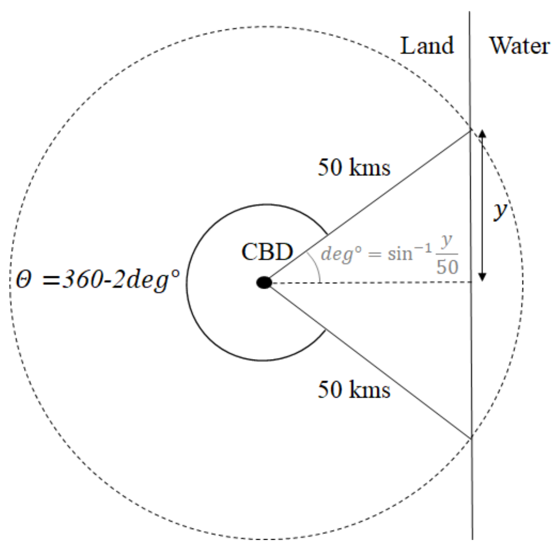
Figure 4: Building Height and the Number of Floors



8.2 Arc of Expansion

Data on the arc of expansion from Malpezzi (????) was complemented with our own calculations using Google maps. The method is simple, for cities that have shore with a large bodies of water (oceans or Great Lakes), we draw a circle with a radius no greater than 50 miles. Drawing a triangle as described in Figure (5), we calculate the degrees of expansion lost to the body of water.

Figure 5: Measure Arc of Expansion



References

Ahlfeldt, G. M. and McMillen, D. P. (2018). Tall buildings and land values: Height and construction cost elasticities in Chicago, 1870-2010. *Review of Economics and Statistics*, 100(5):861–875.

Albert, S. (2010). The geographic determinants of housing supply. *Quarterly Journal of Economics*, 125(3):1253–1296.

Albouy, D., Ehrlich, G., and Shin, M. (2018). Metropolitan land values. *Review of Economics and Statistics*, 100(3):454–466.

Combes, P., Duranton, G., Gobillon, L., Puga, D., and Roux, S. (2012). The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection. *Econometrica*, 80(6):2543–2594.

Combes, P.-P., Duranton, G., and Gobillon, D. (2016). The costs of agglomeration: House and land prices in french cities. *mimeo*.

Harari, M. (2020). Cities in bad shape: Urban geometry in india. *American Economic Review*, 110(8):2377–2421.

Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *American Economic Review*, 57:197–210.

Muth, R. F. (1969). *Cities and housing; the spatial pattern of urban residential land use*, volume 91. University of Chicago Press, Chicago,.