Abstract

Micro-geographic data capturing the spatial distribution of economic activity within cities in history are difficult to access. This paper discusses how tall and durable buildings can be exploited as a source of “big data” to trace the history of the spatial structure of cities. To this end, we provide stylized evidence on how building heights correlate with land values over space and time within cities, review the related nascent literature, and suggest future research areas.

Key words: Density, economics, history, skyscrapers, urban

JEL: R3, N9
1 Introduction

The formal analysis of the cities’ internal structure has a tradition that dates back to von Thuenen (1826). During this period, cities have become the natural habitat of humankind as the level of urbanization increased from less than 10% to more than 50% (Ritchie, 2018). This rapid urban growth was accompanied by radical changes in the way economic agents produce, consume, and interact in space. During the 20th century, the decentralization of manufacturing activity owing to reductions in transport costs, along with the rise of highly agglomerated knowledge-based tradable services, has changed the shape of cities and given birth to sophisticated polycentric spatial configurations in which edge cities and sub-centres thrive alongside central business districts (Ahlfeldt et al., 2020). Tracing the spatial history of cities, however, is often difficult since micro-geographic information on wages, rents, population, and employment densities are not always readily accessible.

In this paper, we argue that data on durable and tall structures represent an underutilized resource that can be exploited as historical “big data” to study the spatio-temporal evolution of cities. The theoretical case is easily made since, at any given point in time, all factors that affect location-specific demand such as production and consumption amenities capitalize in the price of floor space. As long as the cost of producing height exceeds the returns to height at the margin, there is a direct mapping from rents (and hence demand) to profit-maximizing building heights. Since the evidence supports the canonical height model (Ahlfeldt and Barr, 2020), inferring the historical demand for locations from the heights of old buildings is an avenue worth exploring.

Compared to alternative sources of “big data” such as lights at night (Florida et al., 2008; Mellander et al., 2015; Henderson et al., 2012) or geocoded social media data (Glaeser et al., 2018; Indaco, 2020; Chauvet et al., 2016), historic building heights have the advantage that they are available at a fine-grained spatial level and for the more distant past. As an example, detailed historic fire insurance atlases often provide information on all buildings in existence at the time of publication. Commercial databases such as that produced by Emporis also contain buildings that have been demolished. But even a contemporary snapshot of tall buildings along with their completion dates gives limited cause for sample-selection concerns since tall buildings are extremely durable.

In Section 2, we begin by briefly reviewing the history of technologies that gave rise to increasingly tall buildings. We then summarize the theoretical literature on the determinants of building heights in Section 3. In Section 4, we illustrate how
tall buildings preserve the spatial history of cities, similar to the way annual rings of old trees provide information about historical environmental conditions. Exploiting some readily accessible data sources, we find that tall buildings are suitable for the inference of within-city economic cores even in settings where height regulations and construction technology rule out skyscrapers.

We then review the growing but small literature that has used building heights to infer on other outcomes in Section 5. Next, in Section 6, we provide an overview of historical data sources that, upon digitization, may draw a more comprehensive picture of building heights in history, especially when merged with other historical data. In Section 7, we conclude with the suggestion that these data sources be used to systematically evaluate the internal structure of cities over the course of the 20th century.

2 The origins of tall buildings

In this section, we briefly review the technological history of tall buildings.

Two seminal innovations during the second half of the 19th century allowed buildings to surpass their typical height of five floors. First, the elevator turned the higher floors in a building from a disamenity into an amenity (Barr, 2016; Ahlfeldt and Barr, 2020). Secondly, steel beams and columns provided a more efficient way to carry a building’s weight. Before that, taller buildings required thicker masonry walls to bear the structure’s load, which then eroded rentable space on the lower floors (Barr, 2016).

Reviewing the suite of the technologies needed to build tall office buildings, historian, Carl Condit, concludes, “If we are tracking down the origins the skyscraper we have certainly reached the seminal stage in New York and Chicago around the year 1870” (Condit, 1988, p. 22). By the early 1890s, the key innovations—the steel-framed skeletal structure and the electric elevator—were in place to remove the barriers to height. So that from that time forward, skyscraper height decisions were based on balancing the costs with the revenues and were not so much determined by engineering issues per se.

Over the course of the 20th century, advances in material science and mainframe computing brought down the weight and the cost of tall buildings. As a result, the heights of the tallest buildings have increased at a rate of about 1% per year, with the number of tall buildings exceeding 150 meters growing at a rate of 5% (Ahlfeldt and Barr, 2020). Yet, tall structures have remained extremely persistent. Out of more than 4,000 tall buildings with non-missing information, Emporis only lists a
handful of buildings as no longer existing. In Hong Kong, for example, only 0.16% of 100 meters or taller skyscrapers have been demolished. In the class of buildings exceeding 250 meters, the World Trade Centre Twin Towers are the only buildings in the world that fall in this category.

3 Theoretical background

In this section, we briefly review the economics theories that link the spatial distribution of economic activity to the urban height profile.

Traditionally, urban economists have not modelled height explicitly. That said, the density of development has been added into the standard horizontal bid-rent framework (Alonso, 1964) by Mills (1967) and Muth (1969) already. Since then, a range of stylized (Brueckner, 1987) and quantitative (Ahlfeldt et al., 2015) models feature some notion of structural density, broadly defined as building services per land unit (Eppl et al., 2010). Intuitively, where the price of residential or commercial building services is higher, developers have an incentive to use more capital per unit of land in the production of building services as they can afford higher marginal costs. While structural density is technically closer to the floor-area-ratio (FAR) than height, the two measures are mechanically correlated since there are natural bounds for the site occupancy index.

Ahlfeldt and McMillen (2018) provide a simple framework for the evaluation of optimal building heights that engages explicitly with the economics of tall buildings. Developers face average construction costs that are convex in height since taller buildings require more sophisticated structural engineering (to withstand lateral wind loads), facilities such as elevators, and expensive building materials. Average revenues also increase since height has an amenity and signaling value (Liu et al., 2018, 2020). If costs of height exceed returns at the margin, there is a finite solution for the economic height and a one-to-one mapping from rent to height and land price, which—in a competitive market—is simply the capitalized present value of the building profit per land unit.

There are several reasons for why observed building heights may deviate from the economic height such as holdout problems (Strange, 1995), the durability of building stock (Brueckner, 2000), option values (Williams, 1991), height competition (Helsley and Strange, 2008), national vigor (Bradsher, 2004) and floor area regulation (Bertaud and Brueckner, 2005) and building height caps (Barr, 2013; Weiss, 1992). Sometimes, it may even be tall buildings that spur local economic development rather than the other way round (Al-Kodmany, 2017). Yet, the evidence, overall,
suggests that economic fundamentals that capitalize in floor space prices are powerful predictors of building heights (Barr, 2010; Barr and Luo, 2020). For a more detailed discussion of the theoretical framework and the supportive evidence, please see Ahlfeldt and Barr (2020).

4 Stylized evidence

In this section, we provide stylized evidence of how tall buildings serve as living memories of the spatial history of cities. Since the theoretical predictions we exploit are not use-specific, we do not distinguish between commercial and residential buildings in the remainder of this section.

In Figure 1, we present stylized versions of the urban height profiles of the most vertical U.S. cities, New York and Chicago. We use a data set from Emporis, previously utilized by Ahlfeldt and McMillen (2018). The company is arguably the leading commercial provider of data on tall buildings, but the coverage is less comprehensive for shorter structures.\(^1\) We view the data set as reasonably representative of the height of tall buildings, but not necessarily the average building.

For this reason, we aggregate the data set to 500m×500m grid cells, keeping the tallest building within each cell at any given point in time to minimize measurement error. Panel (a) nicely illustrates how Downtown and Midtown were hosts of economic activity in New York City 100 years ago. Consistent with improvements in construction activity, building heights in New York City have generally increased, but, in relative terms, the two main clusters continue to stand out.

Panel (b) documents that although Chicago features some of the earliest tall buildings completed in the wake of the Great Fire of 1871, its tallest skyscrapers did not rival those in New York City until the second half of the 20\(^{th}\) century. Even in 1965, the tallest structures where still roughly at the level of the tallest structures in New York City 50 years ago, in part due to building height restrictions imposed from 1893 to 1956 (Barr, 2013). Throughout its history, Chicago has remained a largely monocentric city, although several sub-centers have emerged (McMillen, 2001).

In Figure 2, we similarly compare the histories of the urban height profiles of London, arguably the most vertical city in Europe, to Hong Kong, currently the most vertical city of the world. In 1915, Hong Kong had a population of about half a million. In contrast, inner London alone had about ten times as many inhabitants. Hence, it is not surprising that the Emporis data base contains virtually no tall construction for Hong Kong up until 1915.

\(^1\)Jedwab et al. (2020) use the entire data set from https://www.skyscrapercenter.com/ and find that for structure of 80m+ there is a strong correlation between the two data sets.
By 1965, Hong Kong had grown to a population of about three million, which roughly corresponded to inner London, and which had experienced significant population decline. Still, Greater London was still more than twice as large. At around 100 meters, heights in both cities remained significantly below the tallest buildings in vertical U.S. cities. The absence of peak heights in both cities is consistent with a rather restrictive British planning system which extended to Hong Kong during the
British lease period. By 2015, Hong Kong had grown to roughly the size of Greater London (close to eight million). Not only do the tallest structures rival those in New York or Chicago, but there also are very tall structures throughout the city, and areas where buildings exceed 100 meters are the rule rather than the exception. While London also developed into the third dimension, development falls short of Hong Kong. Heights barely exceed 200 meters. It also maintains an approximately monocentric structure, with the secondary financial centre Canary Wharf being the visible exception.

In Figure 3, we summarize how the spatial structure of Chicago and New York change over 140 years from 1870 to 2010, measured in terms of heights of completed buildings within a decade and land values at the beginning of a decade. As before, we focus on the maximum heights (and for consistency also maximum land values) within 500m × 500m grid cells, as we presumably measure the height of the tallest building with less error than the height of the average building. Concretely, we present the point estimates and confidence bands of regressions of either the log of building height or the log of land value against the log of the distance from the CBD.

We observe that the land price gradient remains stable in New York and flattens out up until 1940 in Chicago, consistent with reductions in transportation costs owing to the increasing availability of mass transit and the automobiles. During the second half of the 20th century the trend has turned, consistent with the rise of knowledge-based tradable services that agglomerate in dense clusters. The 2010 estimate for Chicago is based on a relatively small sample of vacant land sales from the beginning of the decade, which were affected by the Great Recession. As with the 1940 estimate, which reflects the impacts of the Great Depression, it is presumably an outlier. That said, the land price gradient in New York started flattening in the 2000s suggesting a spatial shift in demand that is unrelated to the financial crisis.

As with the land price gradient, the height gradient in New York became steeper over time. These trends can be rationalized with an increase in central-city floor space rents within the canonical model (Ahlfeldt and Barr, 2020). In Chicago, the pattern is more nuanced. Up until the 1930s, the height gradient remains quite stable. This is not too surprising considering that Chicago initially responded to the new technology by regulating building heights. The more relaxed zoning regime, which is floor-area based and does not limit heights explicitly, has been in place since the 1950s. Indeed, throughout the second part of the 20th-century height and land price gradients follow a similar trend, consistent with the canonical model.

Figure 4 shows that the correlation between land values and building heights was striking even a century ago. Consistent with the canonical model, demand for space
in Downtown and Midtown materializes in tall buildings and high land prices (panel a). Historic Chicago, in panel (b), is an interesting case because skyscrapers were not yet a feature of the skyline. There are some selected tall structures in remote areas such as the 14-story Nichols Tower, built in 1906 at Homan Square, west of the CBD. The tower was part of the Sears, Roebuck & Co. complex before the company moved downtown in 1973, giving name to the tallest skyscraper in the world, at the
As such, the company history makes for a nice anecdote on how some private and corporate developers appear to have an intrinsic motivation to build tall, possibly for reasons unrelated to the canonical neoclassical economics framework. Overall, however, the supply of tall buildings follows the value of location closely. Where we observe tall buildings and low land values, this is usually in “emerging” locations with high fundamental value, either near downtown along representative boulevards, or near the shore of Lake Michigan.

In Figure 5, we analyse the relationship between heights and land values more systematically by correlating the log of the former with the field rank of the latter. Since building height data are sparse for shorter buildings, in particular for Chicago, we cannot use heights to predict land values in low-density areas. However, in areas where we observe tall construction, the field rank of the tallest completed building within a 500m\(\times\)500m grid cell is a powerful predictor of the land value throughout the 20th century. Hence, historic building heights are particularly useful to locate the centres of urban economic activity within cities. Admittedly, due to the durability of capital, tall historic buildings will be more helpful in tracing the history of urban growth than decline.
5 Tall buildings as “big data” in the literature

In this section, we review the nascent literature that has used building heights as sources of information to infer other outcomes.

While the availability of spatial data sets suitable for urban economics research has improved, there are still significant limitations. Population and employment data are often available for relatively small spatial units such as municipalities, but not necessarily at a truly micro-geographic scale, e.g., city blocks. Property transactions data is often geocoded at the building level, but often sparse and sometimes hard to access. Compiling micro-geographic data sets that cover multiple countries remains a challenge.

In contrast, building heights are becoming increasingly available at a global scale due to remote sensing (LIDAR) and commercial (Emporis) and non-profit initiatives (CTBUH Skyscraper Center) that operate beyond national borders. To date, a handful of studies have used these data sets on building heights to proxy for local economic activity at fine-grained levels.

Arribas-Bel et al. (2020) explore a detailed data set containing building heights of 12 million Spanish buildings from the Spanish Cadaster. To delineate urban areas, they use a machine learning algorithm that groups buildings within portions of space of relatively high density. With this approach, they obtain spatial units that resemble commuting patterns more closely than administrative boundaries, suggesting that the approach is suited for the identification of functional urban areas.

In many developing cities, there is a formal sector characterized by relatively tall buildings, whereas buildings in the informal sector are typically flat (Henderson et al., 2016). Harari and Wong (2019) exploit this feature by using building heights as a proxy for formality. Their study is also innovative in that they show how to collect a data set on building heights from photos (from Google Street View and fieldwork) if suitable data is not readily available.

There are also a couple of cases where researchers infer the spatial distribution of economic activity in history. To identify the CBD in Chicago, Ahlfeldt and McMillen (2018) estimate a non-linear least squares model that treats the geographic coordinates of the CBD, as well as the rate of decay in the distance gradient as parameters. They find that throughout the 20th century the nucleus of the land price gradient and the nucleus of the height gradients are very close. Looking at 125 global cities, Ahlfeldt et al. (2020) show that the height gradient peaks near the city hall in 1900, which in turn co-locates with knowledge-based tradable services. Hence, it appears that building height data may be suitable for the identification of...
business centres in the absence of better data.

Buringh et al. (2020) compile a multi-national data set of about 1,700 churches built between 700 and 1500 CE in Western Europe. During this period, churches were typically the tallest structures and, in the absence of modern construction technology, expensive to build. They use this data set on tall structures to generate an index of economic development for this economically formative, but poorly documented, era.

Lastly, research by Jedwab et al. (2020) uses the entire building database from The Skyscraper Center to study the impacts of building height regulations for countries around the world. From this data set, they are able to compare total building height (for all structures 80 meters or taller) for each country to what they “should" have built in the face of lighter regulations. They then investigate how these “gaps” correlate with sprawl, housing prices, and air pollution.

6 Historic data on tall buildings

In this section, we review sources of data on tall buildings that are suitable for the analysis of urban spatial structure.

Current databases on skyscrapers include Emporis (Ahlfeldt and McMillen, 2018; Barr, 2010, 2012, 2013), The Skyscraper Center (Jedwab et al., 2020), and Skyscraperpage.com (Barr, 2010). Information on construction costs, building use, or net floor area are frequently missing, but there is decent coverage of building heights, construction dates, and geographic location, at least for tall structures.

Increasingly, comprehensive data sets covering built structures are available from city administrations. For example, since 2002, New York City Department of City Planning uploads data sets, which include the building and lot information for every tax lot in New York City (Barr and Cohen, 2014). Brooks and Lutz (2019) use a similar data set that contains information on structural density for 2.3 million properties in Los Angeles County from 1999 to 2011. Similarly, other cities like Chicago, Boston, and San Francisco generate building-level data sets generated from LIDAR or other forms of remote sensing. A typical limitation of such data is that they are compiled for individual cities and relatively recent cross-sections, exclusively.

While older buildings in contemporary data sets conserve information about economic conditions at the time when height decisions were made, there is an obvious concern about the non-random selection of buildings that survive history. For very tall structures such as skyscrapers, the durability of the structural capital mitigates this concern. In general, however, historical data on building heights are the pre-
ferred alternative, if available. Indeed there are suitable sources whose potential could be further exploited in urban economics research. Randall and Randall (1999) and Leslie (2013), for example, list important tall buildings in Chicago. The most comprehensive source on nearly all ten story or taller buildings in the U.S. comes from Friedman (2014), who lists all ten-story or taller buildings ever constructed in the U.S. from 1900 and earlier.

Recently, some archives such as the New York Public Library (NYPL) and the Library of Congress have begun digitizing historical maps and atlases, which will likely result in a vast new trove of useful historical data. As an example, Sanborn Map Company regularly published detailed atlases since the mid-19th century for several cities. Since they were used by fire insurance companies to help set rates, they include information about floor heights and material types. The G.W. Bromley & Co. also issued detailed atlases for cities like New York (see https://digitalcollections.nypl.org/collections/atlases-of-new-york-city).

In addition, the NYPL has a program that allows volunteers to upload georeferenced scans of historic map excerpts, with the aim of ultimately generating digitized historical maps of entire cities (see http://maps.nypl.org/warper/). These maps frequently contain data on uses, height, and building materials (such as wood or masonry). Examples of research that uses these data for historical urban research can be found in Baics and Meisterlin (2016) and Baics and Meisterlin (2019). At the same time, the extraction of building-level data from historical data sources such as atlases is becoming a more feasible for smaller groups of researchers, thanks to advances in geographic data science. For a discussion of techniques to digitize historical maps, see Combes et al. (2020).

7 Conclusion

Urban economics research has just begun to exploit the potential for building heights data to serve as a proxy for economic activity. Compared to lights at night (Florida et al., 2008; Mellander et al., 2015; Henderson et al., 2012) or social media data (Chauvet et al., 2016; Glaeser et al., 2018; Indaco, 2020), building heights have the advantage of being available for historical periods, enabling the study of the long-run evolution of spatial economies at a micro-geographic scale. Perhaps more than with other sources of big data, however, it is important that researchers are wary of the institutional context. Especially the most monumental tall structures may be the result of national vigour or the attempt to spur local economic development, in which case they provide little information about local demand. In the presence of
height caps, building heights data need to be interpreted as top-coded.

There may also be some scope for creative use of tall buildings as source of exogenous variation in density, e.g., to identify agglomeration effects. Geologic features have been proposed as instruments since they presumably impact the construction costs of tall buildings (Rosenthal and Strange, 2008; Combes et al., 2011). Tall historic buildings such as cathedrals may provide an additional source of identifying variation as heights in their surroundings are often regulated to protect views (an intriguing idea we owe to Sascha Becker).

In this short piece, we have sketched the potential of tall historical buildings to serve as indicators of the location and strength of concentration of economic activity within cities in history using readily accessible data. The true potential, however, lies in comprehensive historical maps, such as fire insurance maps, that provide full coverage at a given point in time. Since similar maps are available for virtually all cities in the U.S., if not in the developed world, they hold the potential to move the literature analysing the spatial history of cities beyond the case-study context.
Figure 4: Heights and land prices in 1920 Manhattan and Chicago

Note: Heights are proportionate to the extrusion of vertical bars and correspond to the tallest building within a 500m x 500m grid cell built up until 1921 in New York City and 1919 in Chicago. Colour coding is proportionate to the land price. Height data in Chicago limited to buildings covered by Emporis. Land price data in New York limited to Manhattan. See Figure 3 notes for details on the data.
Figure 5: Land value vs. height rank

Note: Unit of observation is 500m x 500m grid cell. Land value is the maximum land value within the grid cell at the beginning of the decade. Field rank (lowest rank to highest value) is computed for the maximum height of buildings completed within a grid cell within a decade. See Figure 3 notes for details on the data.
References


Hoyt, Homer, *100 years of land values in Chicago: The relationship of the growth of Chicago to the rise in its land values, 1830–1933*, University of Chicago Press, 1933.


