

Curci, Federico

Conference Paper

The taller the better? Agglomeration determinants and urban structure

55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Curci, Federico (2015) : The taller the better? Agglomeration determinants and urban structure, 55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/124742>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

The taller the better?
Agglomeration determinants and urban structure

Federico Curci¹

5th European Meeting of the UEA: Young Scientists 2015

Preliminary working paper. Please do not cite

¹Universidad Carlos III de Madrid. Departamento de Economía. Calle Madrid 126, 28903 Getafe (Spain). Email: fcurci@eco.uc3m.es. I thank Ricardo Mora Villarrubia and Adelheid Holl for the excellent supervision. I thank Klaus Desmet, Natalia Fabra, Federico Masera, Jan Stuhler, and Ulrich Wagner and the participants to 2015 SERC Annual conference for valuable comments.

Abstract

This paper explores how urban structure and building height can play an important role for agglomeration and the consequent productivity advantages, looking at the role of skyscraper in influencing the concentration of establishments in U.S. cities. In addition to productivity advantages associated to this extreme form of density, skyscraper can be a particularly attractive location for firms because of the associated gains in prestige from being located in a tall landmark building. Geological and technological instruments are used to determine the effect of skyscraper on firms' location, exploiting a panel of 14,114 ZIPs 147 U. S. Metropolitan Areas from 2000 to 2012.

One of the most important results is that the effect of newly completed skyscraper on agglomeration differs between sectors. The attraction of establishments on the ZIP codes where tall buildings will be completed has an important anticipatory component. Evidence of small congestion effects cannot be rejected. Exploiting the variation of firm's density produced by tall buildings, my estimation suggests that agglomeration economies provided by tall buildings might provide an additional 20 percent increase in productivity.

1. Introduction

Agglomeration economies refer to the fact that both firms and workers are more productive in urban areas. A growing part of the literature is addressing the quantification of the elasticity of wages and productivity with respect to urban density. Moreover, different mechanisms have been proposed by the literature to explain agglomeration economies. The literature started with Marshall (1920) which recognizes that input sharing, labour market pooling and knowledge spillovers are responsible for higher productivity in more dense areas.

Skyscrapers can be seen as an extreme form to increase urban density. The construction of tall buildings has been also used for urban requalification and renewal. For instance, the construction of the World Trade Center in New York had as objective the revival of Lower Manhattan (Helsley and Strange, 2008). Little work has been done on the analysis of skyscrapers impact on urban economic development. Koster et al. (2014) have assessed the existence of a building height premium. Firms might be willing to pay higher rents in floors at higher floors because of within-building agglomeration and landmark reputation. Therefore, skyscrapers can make a particular location more attractive because of both productivity gains and prestige effects from being in the tallest area of a city or a country.

This paper aims at assessing the importance of urban structure and building height, establishing the effect of skyscrapers on firms' agglomeration. An important contribution of this paper is to establish the sector heterogeneity in the effects of skyscrapers in order to understand which sectors are more agglomerated in cities with more skyscrapers. This will be done performing different estimation of the effect of skyscraper on agglomeration for each NAICS sector, for both manufacturing and services sectors. Moreover, dynamic and spatial effects will be introduced in our specification in order to further prove the existence of an attraction of firms in areas close to tall buildings. It will be also assessed the impact of the completion of new skyscrapers on ZIP codes productivity. Finally, we can exploit the variation in density given by the construction of a new tall buildings in order to give a quantification of agglomeration economies, seen as the elasticity of productivity to firms' density, at ZIP code level. Differently from previous papers time variation will also be exploited to estimate agglomeration economies.

The empirical analysis is conducted using a rich database including all NAICS sector, for 14114 ZIP codes for 147 Metropolitan Statistical Areas (MSA) in U.S. from 2000 to 2012. This database have been personally built combining information on geographic establishments location from the U.S. Census Bureau with data on skyscrapers construction. The estimation of the effect of the completion of new tall buildings have been conducted using instrumental variable fixed effects techniques. In order to obtain exogeneous variation, the completion of new skyscrapers have been instrumented using the interaction between the distance to bedrocks in one ZIP area with the Global steel price.

Results suggest that the completion of new tall buildings have an agglomeration effect that depends on the sector under consideration. In particular I find a positive and significant effect on agglomeration of the following sectors: Finance and Insurance, Real Estate, Information, Management of Companies and Enterprises, Educational Services, and Accommodation and Food Services.

The agglomeration effect given by new tall buildings begins even before the construction of the structure and tall buildings have long-lasting agglomeration effects. Anticipatory effects have been estimated using an exponential discount model. Moreover, I find evidence of small diseconomies of scale and firms' relocation outside the area where a skyscrapers have been completed.

The last section of this paper gives a quantification of agglomeration economies, that is the correlation between productivity and city density. This estimation has been done exploiting the fact that tall buildings provide a clear shock in density at lower level of geographical aggregation. Moreover, identification of this elasticity has been done exploiting time variation in addition to cross-sectional variation. It has been found that the completion of tall buildings contribute to an additional 20 percent elasticity of productivity

with respect to firms density. Looking at shocks at more disaggregated level of geography, agglomeration economies has been found to be higher than previous estimates.

Finally, this paper contributes to the literature analysing the agglomeration effect from an increase in urban density related to an increase in height of building. This approach will allow to suggest that other factors different from the classical determinants of agglomeration, such as landmark reputation and prestige of particular areas, cannot be discarded from the analysis of agglomeration economies. This implies that urban structure cannot be neglected when considering firm location.

The rest of the paper is structured as follows. Section 2 discusses the related literature. Section 3 and 4 illustrate the identification strategy adopted and the data used. In Section 5 I have reported results and evidence found of the relationship between agglomeration and completion of tall buildings. Section 6 presents the estimation of agglomeration economies obtained. Section 7 provides concluding remarks.

2. Related literature

The increase in urbanization and the economic advantages of cities have attracted the attention of many scholars. Several works have investigated the sources of agglomeration economies both at a theoretical and empirical level (see Rosenthal and Strange, 2004, Duranton and Puga, 2004, and Puga, 2010, for a complete review). The microfoundations of urban increasing returns trace back to the work of Marshall (1920), who argues that input sharing, labour market pooling, and knowledge spillover are sources of agglomeration economies. Firms would decide to locate in cities in order to be closer to their input suppliers and due to economies of scale in input production. The presence of a large and better formed labour supply in cities would allow firms to obtain a better and quicker labour matching. Moreover, being close to other firms would help the emergence and the spillover of ideas and innovation. This sources of agglomeration allow cities to have higher productivities and therefore to attract more firms.

Rosenthal and Strange (2004) summarize other determinants of agglomeration economies discussed in the literature, such as home market effects, consumption possibilities and natural advantages. Home market effects derives from the fact that concentration of employment in large factories due to increasing returns creates a large market which induce other firms to choose the same location. Furthermore, consumption and amenity possibilities would make cities more attractive. Natural advantages have been also indicated as possible source of urban increasing returns.

From an empirical perspective several works have tried to assess the relative importance of the determinants of agglomeration. This studies have investigated what are the conditions for which some industries tend to cluster more than others. This task has

been accomplished by regressing an agglomeration measure for a specific sector on a set of sector-specific proxies of the different sources. Audretsch and Feldman (1996) and Rosenthal and Strange (2001) have performed this analysis at a static level. On the other hand, Dumais et al. (2002) analyses the dynamic nature of agglomeration, decomposing the change in the agglomeration measure in the part due to mean reversion, randomness, new firm or plants births, and plant closure. Evidence of labour market pooling and some of knowledge spillovers have been found by the authors. These works have benefited from the discussion on the characteristics of agglomeration measures done by Ellison and Glaeser (1997). More recently Ellison et al. (2010) have investigated the determinants of coagglomeration between industries.

The econometric strategy used in previous works cannot account for city specific characteristics that influence agglomeration. Exploiting variation at both ZIP and metropolitan area I can assess what are the conditions of cities in which there is higher agglomeration, and in particular if the construction of tall buildings has any agglomeration effect.

The role of urban structure has received a more limited attention as a source of agglomeration economies. While cities have higher productivity with respect to rural areas, there are also differences between cities' productivity due to transport infrastructures and the compactness of the city (Cervero, 2001). Harari (2015) shows that cities which are more compact are characterized by larger populations and that there exist welfare costs related to city shape.

Central business districts are characterized by tall buildings and Koster et al. (2014) have assessed the existence of a building height premium, given by within-building agglomeration and landmark reputation. This may suggest that workers are more productive in skyscrapers, because tall buildings can provide high density, opportunities of face-to-face contacts and possibility of specialization. Furthermore, the high density in the areas characterized by the presence of skyscrapers can contribute to agglomeration economies in the all neighbouring area and not just in the single tall buildings.

In this paper I aim at assessing that agglomeration economies are also created because of the height of the buildings in a particular area. In addition to productivity benefits given by the high density, firms might be attracted to particular areas because of the presence of landmark buildings which increase their prestige. Moreover, firms might locate there because other firms expected them to be located there. Skyscrapers provide an example of a situation in which agglomeration might be also caused by causes different from productivity advantages.

Another important branch of the literature have tried to quantify the magnitude of these agglomeration economies (see Combes and Gobillon, 2015 for a complete review). The magnitude of agglomeration economies have been computed as the elasticity between

population or employment density and a measure of productivity, which can be either wages or TFP. Ciccone and Hall (1996) is the first paper in finding a rigorous estimation of the correlation between income and density, by instrumenting density with historical population in 1880. However, as demonstrated by Combes et al. (2008) and Combes et al. (2010), this estimation can be biased by worker heterogeneity, since more productive workers live in more productive areas. In order to deal with endogeneous local determinants and sorting Combes et al. (2010) estimate this elasticity using worker fixed effects and instrumenting population density by historical and geological variables, such as historical populations and soil information. Estimated elasticities in the literature are around 0.02 and 0.04 depending on controlling for individual endogeneity (see Combes and Gobillon, 2015, and Melo et al., 2009).

In this paper I aim at providing further evidence in the estimation of agglomeration economies by looking at density shocks that happen at very disaggregated level within city. Agglomeration economies estimation presented by the previous literature relies on cross-sectional variation. I will exploit the fact that tall buildings provide a clear case in which it is possible to know the time in which density increase. In this way I can estimate the increase in productivity following an increase in establishment density. Moreover, I can exploit variation at lower level of geographical disaggregation (ZIP code) to assess if the density elasticity might be higher inside a metropolitan area.

3. Empirical strategy

3.1. Static empirical strategy

This paper aims at establishing the role of tall buildings on agglomeration using an empirical approach. My estimations face several econometric challenges: within-cluster correlation of the errors, reverse causality, omitted variable bias and time persistency. In order to control for part of the within-cluster correlation of the error, I have performed a cluster-specific fixed effects estimation. A Metropolitan Statistical Area is a proxy of a local labour market. ZIPs in the same MSA share the same pool of labour and other inputs. Therefore, my empirical model contains τ_t and μ_m , the dummy variables for the year and MSA of the observation, and the additional controls for the classical agglomeration mechanism are measured at MSA level. Moreover, standard errors are clustered at MSA level. In order to estimate empirically the effect of skyscrapers on agglomeration the following model 1 is estimated for each different j sector:

$$y_{zjt} = \alpha + \tau_t + \mu_m + \beta D_{zt} + \gamma X_{mt} + \varepsilon_{zjt} \quad (1)$$

where z and m are the geographic units of interest (ZIP codes and MSA respectively),

D_{zt} is the number of new skyscrapers completed in one ZIP code in the previous years ², and X_{mt} includes a number of controls that proxy the different agglomeration determinants: input sharing, labour market pooling, knowledge spillovers and natural advantages. The dependent variables y_{zjt} used are the log number of establishments of sector j in ZIP area z and the log productivity measure. Section 4 discusses how I have measured these variables.

It is difficult to claim that the completion of new tall buildings is an exogenous variable. An important threat to identification comes from reverse causality, and this can arise if the increase in agglomeration in one city leads to demand pressure for more tall buildings. Moreover, omitted variable bias can also be present if the construction of skyscrapers happen in places where land value is lower or when zoning rules have been changed in order to increase the number of commercial activities. In order to control for these endogeneities, I have instrumented the number of completion of new tall buildings using geological and technological variables. In particular, I have used the interaction between depth to bedrocks and the third lag of the Global steel price. The advantages of this instrument is that exploiting this interaction it is possible to obtain both cross-sectional and time variation.

The relevance of this instrument is given by technological condition of the construction of skyscrapers since tall buildings are predominantly built with steel and they need to be anchored to bedrocks in order to prevent uneven settling (Barr et al., 2010), implying that construction costs are higher in cities with more distant bedrocks from the surface. For the same reasons the distance to bedrocks have been used as instruments in other studies that tries to estimate the attenuation of human capital spillovers (see Rosenthal and Strange, 2008) and the magnitude of agglomeration economies (see Combes et al., 2010) Exogeneity of the instrument is guaranteed by the random assignment of bedrocks and by the fact that each other local reasons that determines ZIP establishments are not influenced by the past Global steel price. Moreover, following Combes et al. (2010) I have also controlled for natural advantage and I have drop observations from agriculture and mining sectors since bedrocks distance might be correlated with the historical natural advantages that leads to early development of U.S. MSA.

The decision of using the third lag of the steel price is motivated by the fact that on average skyscrapers construction lasts for 2.5 years and therefore steel price should affect the decision of construction before the foundation have been constructed. Further evidence is presented in Section 5.1 Moreover, according to Dunn (1993) where the distance of

²since my dependent variables are stock variables, the treatment variable will be considered as stock measuring the number of new skyscrapers completed in one ZIP code in the previous years, and not the number of new skyscrapers completed in one ZIP code in that give year only

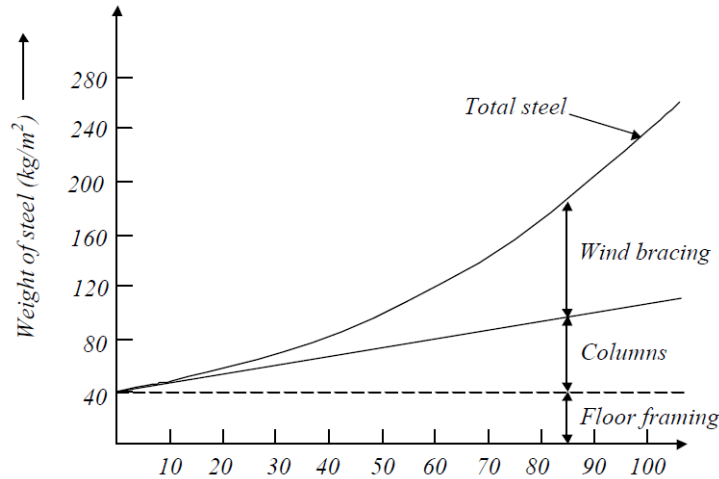


Figure 1: The premium for height: amount of steel needed as function of number of storeys

bedrock is higher the amount of steel needed is likely to be higher. In fact, “if the bedrock lies very deep, [...] one technique involves driving steel piles into place by repeatedly dropping a heavy weight on their tops”³.

A threat to my identification strategy realizes if past steel price influence the construction of other type of buildings which can explain the location of establishments in one particular area. Exogeneity of my instrument is assured by the principle developed by Khan (1969) called “premium for height”. This concept can be explained looking at Figure 1, which has been taken from INSDAG (2013). In fact, according to Ali and Moon (2007) and INSDAG (2013) wind loading and earthquakes put at risk the structure of a tall building since they “act over a very large building surface, with greater intensity at the greater heights”. Therefore, there is a non-linear relationship between additional steel for wind resistance and height. Steel price influences particularly construction costs of tall buildings because of the additional need of this material in order to provide structure resistance. Further evidence is provide in Section 5.1.

The agglomeration control proxies (input sharing, labour pooling and knowledge spillover) have been instrumented using Bartik instruments (Bartik, 1991). The idea of this instrumental strategy is to exploit the different level of aggregation present in our data and estimation. In fact, we will instrument one variable measured at MSA level with the corresponding value at a higher agglomeration, in this case using the U.S. Census Regions⁴. Relevance of this instrument is provided by the correlation of these proxies existing

³More information at <http://www.madehow.com/Volume-6/Skyscraper.html>

⁴The following U.S. Census Regions have been used. *New England*: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont; *Middle Atlantic*: New Jersey, New York, Pennsylvania; *East North Central*: Indiana, Illinois, Michigan, Ohio, Wisconsin; *West North Central*: Iowa, Nebraska, Kansas, North Dakota, Minnesota, South Dakota, Missouri; *South Atlantic*: Delaware, District

between neighbouring areas. Exogeneity is given by the fact that the regional level of one of those variables is given by factors different from local shocks.

3.2. Dynamic and spatial empirical strategy

By introducing the completion of skyscrapers in ZIP codes at several km radius distances from the current ZIP code in consideration as a new control in equation 1 it is possible to shed further light on the spillover effects of tall buildings on neighbouring areas. I will consider several radius distances: between 0 to 5 km, 5 to 10, 10 to 25, 25 to 50, and 50 to 100. A ZIP is considered to be in one particular radius if its centroid is not distant more than the considered km from the centroid of the ZIP code under consideration. This analysis will allow me to have some insights about the existence of diseconomies of scale and possible congestion effects in the area where tall buildings are constructed.

Model 1 can be extended in order to include dynamic effects. Introducing leads and lags of the completion of tall buildings it will be possible to confirm that we have an attraction of firms in area where tall buildings will be constructed even prior to their completion. However, introducing leads and lags of the treatment variable in Model 1 has one important drawback. In particular, this procedure assumes that the number of periods for which anticipation effects occur is known. This can lead to results that are no robust to different specification of the models. In order to overcome this problem and confirm the existence of anticipatory effects, I have estimated an exponential discounting model using the estimation strategy proposed by Malani and Reif (2015).

Ignoring ex-ante anticipatory effects the estimation of the ex-post effect of tall buildings in Model 1 is biased. Therefore, assuming that firms are forward-looking in their location decision I can write a simplified version of Model 1 that also includes the completion of future tall buildings:

$$y_{zjt} = \beta D_{zt} + \beta \sum_{s=1}^{\infty} \theta^s E_t [D_{t+s}] + \varepsilon_{zjt} \quad (2)$$

In Model 2 the ex-post effect of the completion of a new tall building is given by $\beta (1 + \theta F + \theta^2 F^2 + \theta^3 F^3 + \dots) = \beta [1 - \theta (F)]^{-1} = \frac{\beta}{1-\theta}$, where θ is the discounting factor, assuming expectation decay at an exponential rate, and F is the lead operator.

Following Malani and Reif (2015) I can obtain a simplified version of Model 2 that includes only one term for the treatment variable assuming rational expectations. Let's

of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia; *East South Central*: Alabama, Kentucky, Mississippi, Tennessee; *West South Central*: Arkansas, Louisiana, Oklahoma, Texas; *Mountain*: Arizona, Colorado, Idaho, New Mexico, Montana, Utah, Nevada, Wyoming; *Pacific*: California, Oregon, Washington. Information about Hawaii and Alaska have been dropped

define $v_{t,t+s}$ the forecast error done at time t about the completion of a new tall building at $t+s$. Because of rational expectations agents can compute the expectations of the completion of future skyscrapers as the sum of the real construction and a forecast error, that is $E_t[D_{t+s}] = D_{t+s} + v_{t,t+s}$. Adding and subtracting θy_{t+1} from equation 2, we can arrive to this new model:

$$y_{zjt} = \theta y_{zjt+1} + \beta D_{zt} + u_{zjt} \quad (3)$$

where u_{zjt} depends on the three components: the time difference in the original error ($\varepsilon_t - \theta \varepsilon_{t+1}$), the forecast error done at time t about time $t+1$ ($\beta \theta v_{t,t+1}$), and change in forecasts ($\beta \sum_{s=2}^{\infty} (v_{t,t+s} - v_{t+1,t+s})$).

The coefficients I am interested to estimate are the ex-post effect of the completion of new tall buildings ($\frac{\hat{\beta}}{1-\hat{\theta}}$), and the ex-ante effect. In order to test the presence of anticipatory effects I will test the null hypothesis that $\beta \theta = 0$. It will also be possible to compute the ex-ante effects one year before the completion of the tall building ($\hat{\beta} \sum_{s=1}^{\infty} \hat{\theta} = \hat{\beta} \frac{\hat{\theta}}{1-\hat{\theta}}$), two years before ($\hat{\beta} \sum_{s=2}^{\infty} \hat{\theta} = \hat{\beta} \frac{\hat{\theta}^2}{1-\hat{\theta}}$), and so on.

Equation 3 can be estimated using as instruments lags of the dependent variable as suggested in Arellano and Bond (1991). Moreover, from equation 2 is evident that y_{zjt} depends on leads of the treatment variable and then y_{zjt} will be correlated with leads of the dependent variable. Therefore, as suggested in Malani and Reif (2015) I will use leads of the treatment and outcome variables as instruments. I will use leads from the second order further. Exogeneity restriction of these instruments is given by the fact that agglomeration is related to future tall buildings only through y_{t+1} . When estimating Model 3 I will also control for natural advantages and the other determinants of agglomeration.

There are several assumptions imposed to obtain identification. It is possible to have small order autocorrelation, but there should not be autocorrelation higher than order 1 in ε_t . It is also needed that the error in the estimation of agglomeration determinants, ε_t , is orthogonal to the ahead forecast errors. Moreover, change in forecast should be uncorrelated with the actual level of the forecast and independent information must be used to update the forecast.

3.3. Estimation of the agglomeration economies

The estimation that can be performed using the Models in Sections 3.1 and 3.2 are useful in order to understand if the completion of new tall buildings has an effect in increasing agglomeration in the area. However, agglomeration economies refer to the fact that workers are more productive in more dense areas. Therefore, it is interesting to understand if the increase of density due to the completion of new skyscraper have any

effect in increasing the productivity of the firms in that area. This can be realized if tall buildings make possible, using the terminology elaborated by Duranton and Puga (2004), to share, match and learn in a more efficient way.

Following Ciccone and Hall (1996) and Combes et al. (2010) the quantification of the agglomeration economies will be reached by regressing a measure for productivity, in my case the ratio between annual payroll and total employment, on local density. I will use log firms' density as measure of density, referring to the log number of establishments divided by the ZIP area. Local density will be instrumented by historical population density in 1900 in the same MSA.

In addition to this separate density term I will also consider an interaction between density and the completion of tall buildings. This last term will allow me to quantify the additional increase in agglomeration economies provided by skyscrapers. As showed in the previous sections, this interaction is instrumented using the interaction between bedrock distance and the third lag of Global steel price.

Moreover, I can exploit the exponential discounting model presented in Section 3.2 to obtain a quantification of agglomeration economies that properly account for the anticipatory agglomeration effects given by the construction of new tall buildings. Therefore, the model I estimate is the following:

$$\log(P_{zt}) = \theta \log(P_{zt+1}) + \beta_1 \log\left(\frac{est_{zt}}{size_z}\right) + \beta_2 \log\left(\frac{est_{zt}}{size_z} D_{zt}\right) + \gamma X_{mt} + \varepsilon_{zjt}$$

Equation 4 has been estimated with Arellano-Bond using as additional instruments lags and leads of $\log(P_{zt+1})$ and leads of D_{zt} . Since my database does not provide individual information about the firms in the area my estimation might suffer a bias given by sorting. That is, the completion of new tall buildings might attract more productive firm. In order to reduce the bias produced by sorting and the endogeneity in firm's density I also control for X_{mt} which includes natural advantages and the other determinants of agglomeration. MSA fixed effects will also partially control for sorting.

The main advantages of my procedure with respect to the previous literature is that I can estimate agglomeration economies exploiting cross-sectional but also time variation in density. Moreover, previous literature usually regress productivity on density using MSA level information, while I can obtain an intra-city estimation.

4. Database

The database used for the empirical analysis have been personally constructed using different sources. The number of establishments for ZIP code and NAICS sector has been collected from the County Business Patterns (CBP) Database of the U.S. Census Bureau. Another dependent variable that will be used is a productivity proxy given by the ratio of the total annual payroll and the number of employees in one ZIP area. This measure is only present for the whole ZIP and it is not disaggregated by NAICS sector. This productivity measure has been also constructed using the CBP database ⁵.

The number of completed tall buildings has been derived from the CTBUH (Council on Tall Buildings and Urban Habitat) Global Tall Building Database. A building is defined as tall if it exhibits one or more of the following categories: height relative to the context, proportion and building technologies⁶. Proportion is measured using size and floor area, while some of the particular technologies required for being considered as tall buildings are specific vertical transport technologies. In general, a building with 14 or more stories or over 50 meters tall and where at least 50 percent of its height is occupied by usable floor area can be considered as tall. It has been considered only tall buildings which are not for residential use only.

In addition to the number of tall building completed in one ZIP code our estimation also controls for the classical determinants of agglomeration: input sharing, labour pooling and knowledge spillover. Since these mechanisms are expected to take place at a metropolitan level, the relative measurements have been computed at MSA level. For each sector j input sharing have been measured summing the number of establishments of other sectors k weighted by the proportion of inputs by the sector j required (directly and indirectly) in order to deliver one dollar of industry output to final users (denoted as W). This is a measure similar to the one used by Jofre-Monseny et al. (2011). The weighting matrix W comes from the Bureau of Economic Analysis Input-Output Accounts. Hence, denoting est as the number of establishments in one MSA, input sharing has been computed as follows:

⁵It is important to notice the existence of missing data for productivity because of confidentiality reasons. However, if it exists this might create an downward bias in my estimation. In fact, productivity data are missing in ZIPs with a higher number of establishments with more employees. Larger establishments are usually more productive than smaller establishment. The ratio of the mean number of establishments that have more than a 1000 employees to the mean number of establishments that have between 1 and 4 employees is 1:590 and 1:526 in the sample with ZIPs with productivity data and not, respectively.

⁶additional information can be found at <http://www.ctbuh.org/HighRiseInfo/TallestDatabase/Criteria/tabid/446/language/en-US/Default.aspx>

$$I_{jt} = \sum_{j \neq k} W_{j,k} \times est_k \quad (4)$$

Labour pooling and knowledge spillovers have been measured using the proportion of population with at least a BA and the proportion of population in Management, professional, and related occupations. These data are collected from the American Community Survey. Moreover, our database also contains the number of patents for each MSA published by the United States Patent and Trademark Office. These variables are usual proxies for labour pooling and knowledge spillovers, as it is described in Rosenthal and Strange (2004). Ellison and Glaeser (1999) have underlined the importance of natural advantages for agglomeration. Therefore, we have controlled for natural advantages using a dummy if the MSA is either coastal or on the Great Lake.

Proxies for input sharing, labour pooling and knowledge spillovers are measured at MSA level. Since in 2003 there has been a revision of the MSA definition, for years between 2003 and 2012 I have matched ZIPs with MSA definition in use between 2000 and 2003. Primary Metropolitan Statistical Areas has been used in presence of Combined Metropolitan Statistical Areas. Moreover, the American Community Survey did not publish information about education for some counties before 2005, therefore ZIP codes in counties with no education information have been dropped.

The completion of tall buildings have been instrumented by the distance from bedrocks and the Global steel price. I have constructed a variable containing the average depth to bedrock for each ZIP code in U.S. using the information provided by Miller and White (1998)⁷. The Global steel price indicator has been extracted from the CRU Steel Price Indicators. For estimating agglomeration economies, an additional instrument used is historical population density in 1900 at MSA level, which is provided by NHGIS (2011)⁸.

5. Evidence

5.1. Summary statistics

Table 1 presents summary statistics for the variables used in this study. Table 2 reports the difference in mean of the most relevant variables between cities that have completed at least one tall building in the period 2000 and 2012 and the one that have

⁷For almost 4,000 ZIPs no information of distance to bedrock was provided. I have computed this information as the mean value of its closest neighbours: neighbours at 0, 5 or 10 km

⁸Since counties in 1900 does not correspond to counties in 2015, every ZIP have been associated to the counties that it would have belonged in 1900. Population density in 1900 have been computed for all the counties. A current MSA could have belonged to different counties in 1900, therefore for these MSAs I have used the maximum value of density in 1900 between the counties to which it belongs.

not constructed any skyscraper. As it is possible to evince, cities that have constructed new tall buildings tend to have a higher number of establishments, education levels and people in management occupations. Moreover, firms share more inputs and they tend to produce more patents. These cities happen to be also more likely to be on the coast or on the Great Lakes region.

Table 1: Summary statistics

	Mean	Standard deviation	Min	Max
Number of establishments, log	2.074	1.702	0	8.929
Productivity, log	3.471	0.3938008	0.447	7.660
Input sharing	1429.239	557.850	496.651	3231.144
Education	27.707	7.453	10.5	59.1
High skills	34.668	5.247	21.372	54.309
Patent	479.995	916.334	1	11490
Natural advantage	0.333	0.471	0	1
New tall buildings	0.122	0.661	0	13
Distance bedrock	126.424	17.247	53.061	152
Steel price	140.830	38.364	68.9	186.8

Table 2: Difference in mean between cities that built and not at least one tall buildings between 2000 and 2012

	Difference	
Number of establishments, log	0.235***	(125.04)
Input sharing	541.9***	(19.71)
Education	3.302***	(8.20)
High skills	1.885***	(6.61)
Patent	666.9***	(14.07)
Natural advantage	0.102***	(4.00)

t statistics in parenthesis.

*, **, ***: statistically significant at 10, 5 and 1 percent respectively.

The United States have been the birth place of modern skyscrapers and, despite the massive construction of tall buildings in other part of the world, mainly Asia, they still have the highest number of tall structures. From Figure 2 it is observable the contemporaneous increase in the construction of tall buildings in the U.S.. The construction of skyscrapers have followed a cycle around the history of the U.S.. The biggest boom of construction have been in coincidence with the 30s, 70s, 80s and the 00s. However, the biggest increase in construction have been only recently and the financial crisis had a dramatic impact in reducing the tall buildings construction. Contextually, the increase in the height of the skyscrapers shows a positive trend during time (see Figure 3).

Table 3 reports some statistics about the current construction of tall buildings in the U.S.. In fact, in the period 2000-2012 546 new tall buildings have been completed. Out of this 546 buildings, 228 are structures which use is not residential only. That means that

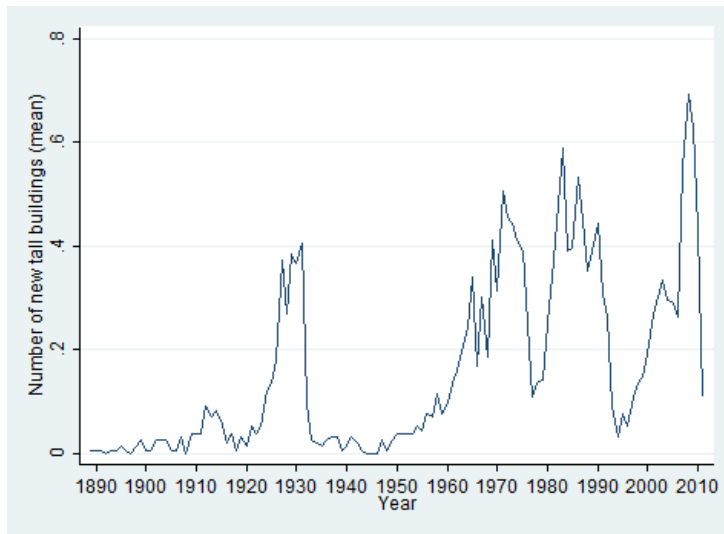


Figure 2: Mean number of new tall buildings completed by city

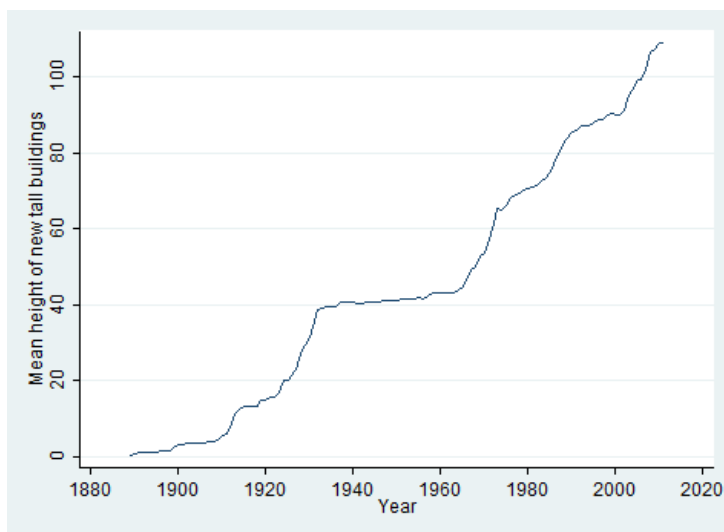


Figure 3: Mean height of new tall buildings completed by city

at 2012 there are a total of 2625 tall buildings, out of which 1713 are not used uniquely for residential reasons. Table 4 reports the cities in which the construction of tall buildings have been the highest. New York leads this particular ranking, followed by Houston and Chicago. Pittsburgh, Detroit and Rochester have been added to this list because they had built many tall buildings in the past but they are not doing it nowadays at the same rate.

Table 3: Number of new completed tall buildings in 2000-2012 and total stock

	Value
New tall buildings, all	546
New tall buildings, no residential	228
Total stock of tall buildings, all	2625
Total stock of tall buildings, no residential	1713

Table 4: Number of new completed tall buildings in 2000-2012 and total stock at 2012, by city

	New tall buildings	Stock of buildings
New York, NY PMSA	62	574
Houston, TX PMSA	30	144
Chicago, IL PMSA	24	178
Miami, FL PMSA	16	26
Dallas, TX PMSA	13	86
Atlanta, GA PMSA	12	46
San Francisco, CA PMSA	8	43
Seattle-Beelevue-Everett, WA PMSA	8	30
Boston, MA-NH PMSA	6	33
Austin-San Marcos, TX MSA	5	21
Jersey City, NJ PMSA	5	6
Fort Worth-Arlington, TX PMSA	3	21
Los Angeles-Long Beach, CA PMSA	3	60
Philadelphia, PA-NJ PMSA	3	79
Raleigh-Durham-Chapel Hill, NC MSA	3	10
San Antonio, TX MSA	3	20
Indianapolis, IN MSA	2	16
...		
Pittsburgh, PA MSA	1	18
Detroit, MI PMSA	0	17
Rochester, NY MSA	0	15

The identification strategy used in this paper relies in both cross-section and time variation. The instrument used is the interaction between bedrock distance and the third lag of steel price. As it is possible to see from the example of Figure 4 bedrock distance assures to provide cross-sectional variation at a very low level of geographical disaggregation. On the other hand, Figure 5 reports the time variation that is guaranteed by the use of the Global Steel Price.

The decision of using the third lag of steel price is not arbitrary. Using information from a limited sub-sample of my database I have estimated that the average year of

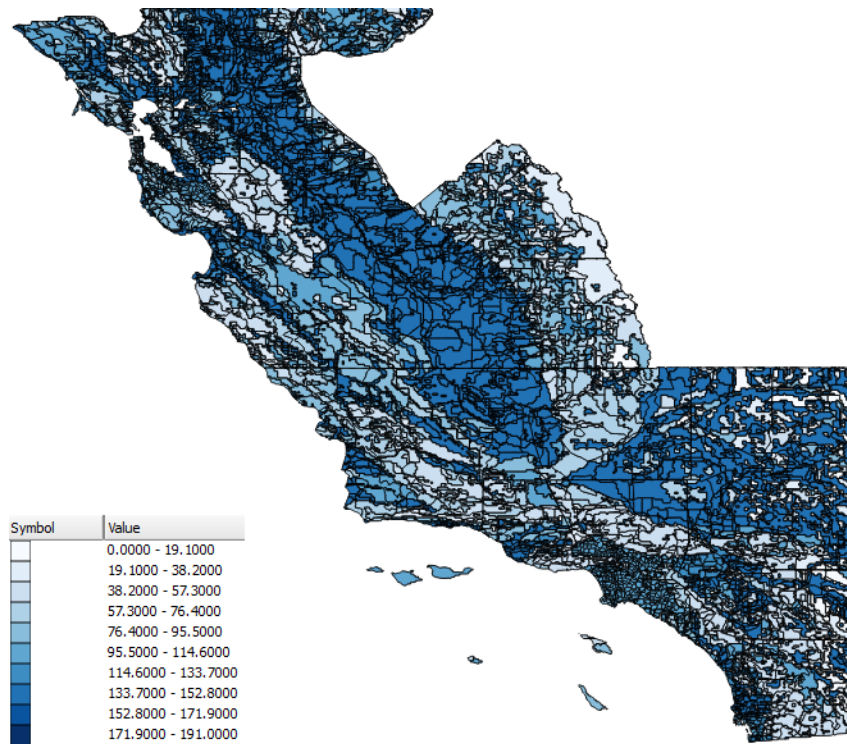


Figure 4: Distance to bedrock in California



Figure 5: Global steel price time series

proposal of a skyscraper is 5.2 years before its completion, while and the average year of beginning of construction is 2.5 years before. The Council of Tall Buildings and Urban Habit consider a building as proposed “when it fulfills all of the following criteria:

1. Has a specific site with ownership interests within the building development team
2. Has a full professional design team progressing the design beyond the conceptual stage
3. Has obtained, or is in the process of obtaining, formal planning consent/legal permission for construction
4. Has a full intention to progress the building to construction and completion

Only buildings that have been announced publicly by the client and fulfill all the above criteria are included in the CTBUH ”proposed” building listings. The source of the announcement must also be credible”. While a tall building has started construction “once site clearing has been completed and foundation/piling work has begun”⁹. Therefore, steel price should affect construction decision between the proposal and the actual construction, that is at least 3 years before the completion of the building. Moreover, this result is also confirmed by the regression in Table 5. The fact that the the source of announcement must be credible provides evidence in favour of the rational expectation assumption for which we can separate the expected number of completed buildings as the sum of the realization and the forecast error.

As it is possible to see steel price at lag 2 and 3 are the only statistically significant coefficients in explaining the number of completed tall buildings. I will use the third lag also in order to avoid any possible endogeneity of steel price, which can happen if a Global shock in steel price might affect differently local labour markets. By using the third lag I can argue that this potential bias can be neglected assuming that past steel price shocks does not affect differently current local labour markets. However, my results are robust to the inclusion of the percentage of jobs in manufacturing in 1999 (before my time sample begins) as proxy of local labour structure.

In section 3.1 I have stressed that for exogeneity of the instrument to be validated it is necessary that steel price might not influence any other factor that can influence agglomeration in one particular area. I have explained how the concept of “premium for height” insures on a theoretical basis that steel price does not influence the construction

⁹Quotation comes from <http://www.ctbuh.org/HighRiseInfo/TallestDatabase/Criteria/tabid/446/language/en-US/Default.aspx>

Table 5: First stage estimations with interaction of bedrock with different lags of Global steel price

Variables	New tall buildings
Distance bedrock X steel price (Lag 0)	-1.87E-07 (4.10E-07)
Distance bedrock X steel price (Lag 1)	3.88E-09 (1.92E-07)
Distance bedrock X steel price (Lag 2)	-1.30e-06** (5.71E-07)
Distance bedrock X steel price (Lag 3)	-5.10e-07* (2.79E-07)
Distance bedrock X steel price (Lag 4)	-6.05E-07 (4.26E-07)
Observations	178,095
R-squared	0.013
MSA FE	YES
Year FE	YES
Natural advantage	YES
Controls	YES
Estimation	OLS

Clustered standard errors in parenthesis.

*, **, ***: statistically significant at 10, 5 and 1 percent respectively.

of other buildings as it does for tall buildings. Using data from the U.S. Census Bureau Construction Spending Historical Database from 1994 to 2014 I have regressed construction value for different usage of the buildings on past values of steel price. Results are shown in Tables 6 and 7. In this preliminary evidence, steel price is not statistically determining construction spending for the big majority of buildings. Possible concerns can arise from the construction of transport infrastructures. However, the fixed effects present in my estimation will capture for time-invariant characteristics of city transport. Moreover, results are robust to dropping the Transport sector.

Table 6: Construction spending determinants by buildings usage

Variables	Non residential	Lodging	Office	Commercial	Healthcare	Educational	Public safety
Steel price (lag 0)	-224.0 (445.4)	-9.116 (69.03)	-40.36 (105.6)	58.39 (118.9)	20.90 (27.89)	-14.41 (18.78)	0.740 (1.547)
Steel price (lag 1)	618.6 (512.4)	88.36 (75.77)	130.5 (101.7)	175.5 (137.9)	56.06 (30.16)	13.76 (23.00)	2.916** (1.191)
Steel price (lag 2)	512.1 (495.9)	89.11 (64.05)	118.9 (114.2)	72.60 (114.3)	48.31 (36.63)	1.780 (23.63)	1.475 (1.102)
Steel price (lag 3)	480.0 (579.9)	74.26 (86.70)	103.3 (103.6)	66.31 (138.9)	7.533 (37.85)	-2.364 (27.03)	2.464** (0.886)
Steel price (lag 4)	733.9 (476.7)	123.5* (61.91)	146.5 (110.3)	95.56 (133.8)	5.745 (35.33)	11.87 (21.49)	1.669 (1.003)
Steel price (lag 5)	273.9 (790.4)	35.89 (116.3)	-41.56 (138.8)	-125.9 (218.0)	5.462 (48.03)	18.70 (33.65)	0.536 (1.393)
Observations	15	15	15	15	15	15	15
R-squared	0.306	0.311	0.320	0.248	0.253	0.197	0.454

Table 7: Construction spending determinants by buildings usage

Variables	Amusement	Transportation	Communication	Power	Sewage	Water supply	Manufacturing
Steel price (lag 0)	13.82 (10.75)	-1.959 (5.142)	20.49 (27.83)	-142.1* (74.27)	2.001 (1.101)	1.362 (1.646)	-130.6 (83.98)
Steel price (lag 1)	5.580 (12.07)	-0.930 (5.818)	54.86 (34.54)	0.316 (89.12)	2.735* (1.251)	-1.198 (1.614)	94.20 (71.87)
Steel price (lag 2)	19.13* (10.07)	13.43** (5.393)	46.33 (28.72)	-23.05 (88.31)	1.520 (1.182)	-0.731 (2.307)	127.6 (71.44)
Steel price (lag 3)	10.67 (11.52)	7.034 (5.695)	70.17 (44.88)	72.10 (109.7)	-0.401 (1.000)	2.594 (1.685)	72.35 (72.29)
Steel price (lag 4)	24.82** (9.465)	13.82* (7.104)	34.81 (30.11)	124.1 (99.32)	3.396** (1.210)	2.005 (1.922)	151.6** (64.88)
Steel price (lag 5)	-5.165 (17.73)	-8.109 (8.951)	-25.45 (52.03)	271.2* (134.0)	1.362 (0.987)	-4.112** (1.459)	163.0* (73.75)
Observations	15	15	15	15	15	15	15
R-squared	0.402	0.449	0.385	0.505	0.705	0.585	0.673

Estimation performed in first difference to avoid unit root and unobserved heterogeneity
Robust standard errors in parenthesis.

*, **, ***: statistically significant at 10, 5 and 1 percent respectively.

5.2. Static regressions

In order to identify the effect of the completion of new tall buildings on agglomeration I rely on an instrumental variable strategy using the interaction between distance to bedrocks and steel price as instruments. Table 8 presents the results for the relevant elements of the first stage estimation of the previous model. As it is possible to see, relevance condition is satisfied for all the instrumented variables with the expected signs. In particular, ZIP codes with a higher depth to bedrocks in years with higher steel prices are associated with a lower probability of completing a new tall building. The Bartik instruments have the expected signs in the first stage. In fact, whenever the variable of one of these controls is higher at a more regional level also the MSA level is higher. The F-test presented suggest that I should not worry about weak instruments.

Before turning to the evidence at ZIP level, I present the results of the estimations of equation 1 taking the average at MSA level of each variable. This will allow me to partially understand the importance of the determinants of agglomeration. Hence, it will be necessarily to control for these determinants even in the ZIP level regression for determining the role of new tall buildings. As it is possible to observe from Table 9 input sharing seems to be the most important determinant of agglomeration exploiting uniquely variation between MSAs. However, I cannot reject that the other determinants are not jointly significant ¹⁰.

I have subsequently exploited variation between ZIPs in the same MSA and in different MSA in order to understand whether tall buildings can increase the location of firms in particular areas of a city. At first, estimating the model presented in equation 1 pooling all sectors together it is not possible to observe a statistically significant effect of the

¹⁰The p-Value of the associated F-test is 0.0002

Table 8: First stage estimations

Variables	New tall buildings	Input	Education	High skills	Patent
Distance bedrock X steel price (Lag 3)	-2.61e-06*** (1.65E-07)				
Regional input sharing		0.448*** (0.0026)			
Regional education			0.749*** (0.0056)		
Regional high skills				0.663*** (0.00775)	
Regional patents					0.975*** (0.00487)
Observations	179,298	179,298	174,942	174,876	179,298
R-squared	0.012	0.994	0.979	0.951	0.961
MSA FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Natural Advantage	YES	YES	YES	YES	YES
Estimation	OLS	OLS	OLS	OLS	OLS
F	19.02	9130.59	5730.81	7036.15	10325.59

Clustered standard errors in parenthesis.

Estimation includes distance bedrock X steel price, regional input sharing, education, high skills, patents, and natural advantages.

F-statistics obtained estimating the first stage by OLS with only the excluded instruments

*, **, ***: statistically significant at 10, 5 and 1 percent respectively.

Table 9: MSA level estimation for the all NAICS sectors

MSA level	
Variables	Log establishment
New tall buildings	-0.00122 (0.00255)
Input sharing	0.000746*** (4.92E-05)
Education	0.00652 (0.00556)
High skills	-0.00589 (0.0055)
Patent	1.21E-05 (1.69E-05)
Observations	1,849
R-squared	0.999
MSA FE	YES
Year FE	YES
Natural advantage	YES
Estimation	IV

Clustered standard errors in parenthesis.

*, **, ***: statistically significant at 10, 5 and 1 percent respectively.

completion of new tall buildings on the number of establishments (see first row in the left panel of Table 10). However, this result might be driven by the possible existence of heterogeneous effects between sectors and the possible presence of dynamic effects. The result is not driven by endogeneity in the instruments as suggested by an exogeneity test ¹¹.

It has been possible to disentangle the effect of new tall buildings on agglomeration looking at the effect for each different sector. The heterogeneity of sector responses can be seen in Table 10 in which it is presented the coefficient for new tall buildings for separate estimations of Model 1 for the different NAICS sectors. As it can be noted industry sectors are associated with a no statistically significant effect of new tall buildings on their agglomeration. In particular, I have not found a significant effect for the following sectors: Utilities, Construction, Manufacturing, Wholesale trade, Retail trade, Transportation and Warehousing, Health Care and Social Assistance, Arts, and Administrative and Support and Waste Management and Remediation Services.

On the other hand, service sectors which are more likely to make use of higher levels of human capital and high skills are associated with a positive and statistically significant of the completion of new skyscrapers. In particular, the completion of a new tall buildings is associated with an increasing of 9 and 10 percent of the number of establishments in the Finance and Insurance and Real Estate sectors in the same ZIP code. The other sectors that present a positive coefficient are Information, Management of Companies and Enterprises, Educational Services, and Accommodation and Food Services.

Despite the existence of agglomeration effect caused by new tall buildings for particular sectors, this does not automatically imply that agglomeration economies might be existing, that is whether there is an increase in productivity caused by higher density. In Table 11 I have estimated the previous regression in order to assess the whether the completion of a new tall building is associated with an increase in productivity. As it is observable a new tall building leads to an increase in productivity in the same ZIP code of almost 2 percent. This increase in productivity could either come from increased agglomeration economies or from firms sorting.

It is important to notice the existence of missing data for productivity because of confidentiality reasons. However, this might create an downward bias in my estimation. In fact, productivity data are missing in ZIPs with a higher number of establishments with more employees. Larger establishments are usually more productive than smaller

¹¹The test has been carried out obtaining residuals of the second stage estimation and regressing them on the instruments. The p-value of the joint significance of the excluded instruments is 0.9639

Table 10: Sectoral estimations

Sector	Log establishment	Sector	Log establishment
All sectors	8.691		
	-5.963		
Utilities	0.822	Information	9.662*
	(0.733)		(5.258)
Construction	-3.677	Finance and insurance	10.34*
	(2.838)		(5.824)
Manufacturing	2.577	Real estate	14.42*
	(3.334)		(7.555)
Wholesale trade	4.154	Professional	12.56
	(3.898)		(8.044)
Retail trade	5.813	Management	6.401*
	(4.559)		(3.393)
Transportation	-2.218	Administrative	5.053
	(2.588)		(4.232)
Health care	9.95	Educational	9.593*
	(6.803)		(5.498)
Arts	8.29	Accommodation	10.01*
	(5.184)		(5.971)
MSA FE	YES	MSA FE	YES
Year FE	YES	Year FE	YES
Natural Advantages	YES	Natural Advantages	YES
Controls	YES	Controls	YES
Estimation	IV	Estimation	IV

Clustered standard errors in parenthesis.

Every row represents at different estimation for a different sector

*, **, ***: statistically significant at 10, 5 and 1 percent respectively.

Exogeneity test of the instruments performed by regressing the residuals of the regressions for all sectors on the excluded instruments.

Test for joint significance of excluded instruments on residuals: p-value=0.9639

establishment¹². Further work will provide estimation of the same model using Heckman techniques to confirm the established result.

5.2.1. Dynamic regressions

In the previous section it has been possible to observe that the completion of new tall buildings have the effect of increasing firms' agglomeration and to increase the general level of productivity of the area. In order to obtain identification of the effect of the completion of tall buildings I have made use of instrumental variable strategy. This estimation is important in order to limit one of the main possible biases of my estimation: reverse causality. Reverse causality might occur if the completion of tall buildings happens in places where firms' demand for building and firms agglomeration is higher. A first look at Figure 6 might suggest that reverse causality could be present. However, this dynamic can be driven by anticipation effects instead of reverse causality. In fact, if firms' location depends on future level of productivity of the area firms might be willing to locate in an area where tall buildings will be completed in order to take advantages of future

¹²Ratio of the mean number of establishments that have more than a 1000 employees to the mean number of establishments that have between 1 and 4 employees is 1:590 and 1:526 in the sample with ZIPs with productivity data and not, respectively

Table 11: Productivity estimation

ZIP level	
Variables	Log productivity
New tall buildings	2.711** (1.276)
Input sharing	-0.000133 (0.000261)
Education	-0.0112 (0.0167)
High skills	0.0141 (0.0173)
Patent	3.25E-05 (2.95E-05)
Natural advantage	0.0691 (0.159)
Observations	152,131
MSA FE	YES
Year FE	YES
Estimation	IV

Clustered standard errors in parenthesis.

*, **, ***: statistically significant at 10, 5 and 1 percent respectively.

agglomeration economies.

The Model 3 explained in section 3.2 will allow to provide a clear test for the presence of anticipation effects. Another contribution of this model is to prove that the previous agglomeration effect is not just driven by firms filling new tall buildings. In fact, I can show that firms are locating in the area where the skyscrapers will be erected even before its actual completion.

Results of my dynamic estimation are presented in Table 12. It is possible to obtain a significant coefficient of the completion of tall buildings on agglomeration looking also at all sectors together. Therefore, previous results might have been biased by dynamic terms. In fact, considering for dynamic effects, the completion of one tall building have the effect of increasing the number of establishment after its completion by 0.6 percent. Moreover, we cannot reject the presence of anticipation effects. This effect is lower in magnitude than the ex-post effects but it is present in all the years after the skyscrapers have been proposed (usually 5.3 years before its completion).

Estimation of the discount term, θ , suggests that time-persistence is important in this estimation. In fact, we obtain a discount term equals to 0.93. However, an AR(2) test on the errors in first difference suggests that persistency is solved.

5.3. Spatial regressions

In the previous discussion it has been argued that the increase in building height in one ZIP code has the effect of attracting firms from particular sectors and induce an overall increase in productivity for the area. However, this positive effect can be counterbalanced

Table 12: Estimation of exponential discount model

Variables	Log establishment
New tall buildings ($\hat{\beta}$)	0.0400*** (0.00788)
F.log establishment ($\hat{\theta}$)	0.936*** (0.00333)
Calculated effects:	
Ex-post effect ($\frac{\hat{\beta}}{1-\hat{\theta}}$)	0.625*** (0.123)
Ex-ante effect (t-1) ($\frac{\hat{\beta}\hat{\theta}}{1-\hat{\theta}}$)	0.585*** (0.115)
Ex-ante effect (t-2) ($\frac{\hat{\beta}\hat{\theta}^2}{1-\hat{\theta}}$)	0.548*** (0.108)
Ex-ante effect (t-3) ($\frac{\hat{\beta}\hat{\theta}^3}{1-\hat{\theta}}$)	0.513*** (0.101)
Ex-ante effect (t-4) ($\frac{\hat{\beta}\hat{\theta}^4}{1-\hat{\theta}}$)	0.480*** (0.095)
Ex-ante effect (t-5) ($\frac{\hat{\beta}\hat{\theta}^5}{1-\hat{\theta}}$)	0.449*** (0.090)
Test anticipation effects:	
$\hat{\beta}\hat{\theta}$	0.037*** (0.00736)
Observations	160,279
MSA FE	NO
Year FE	NO
Natural advantage	YES
Controls	YES
Estimation	AB
Test AR(2)	0.074

Clustered standard errors in parenthesis.

*, **, ***: statistically significant at 10, 5 and 1 percent respectively.

Instruments used: interaction distance bedrock and steel price (third lag), regional level of input sharing, education, high skills and patents, leads of log establishments and tall buildings (from 2nd to 5th), lags of log establishments (from 1st to 5th)

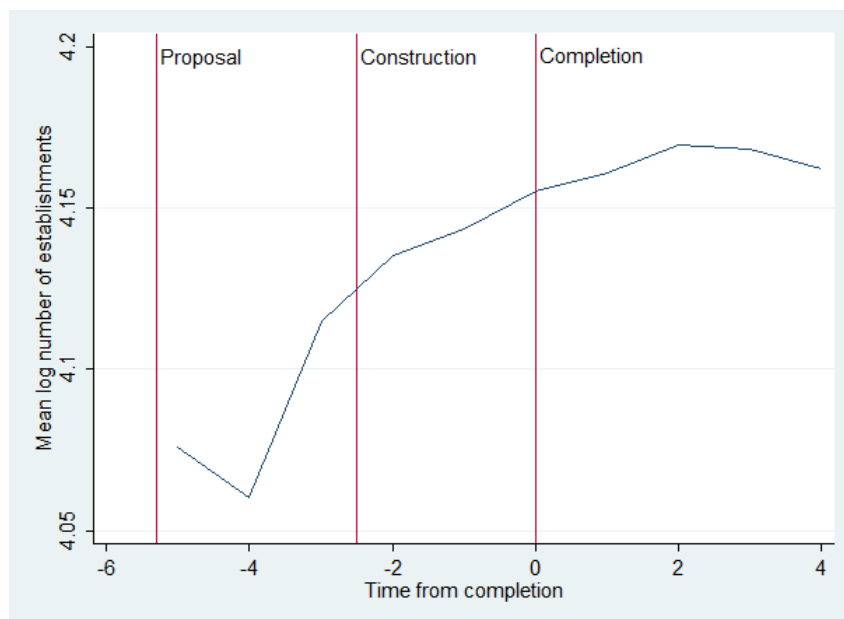


Figure 6: Mean log number of establishments pre and post the completion of a tall building in the same ZIP code

by increased congestion which could be leading to diseconomies of scale in the surrounding areas. In Table 13 I have presented the results of the estimation of our econometric model with the addition of variables representing the number of completed tall buildings in the same year at several radius of distances: between 0 to 5 km, 5 to 10, 10 to 25, 25 to 50, and 50 to 100. This model has been estimated for all NAICS sector and results are presented also for the the more representative sectors.

Generally, the presence of congestion effects and diseconomies of scale cannot be rejected. For some sectors I have found a negative and significant effect of skyscrapers in the closest areas to where tall buildings are completed (between 0 and 5 km distance) and a possible relocation of establishments in more distant areas possibly in the same MSA, between 10 and 50 km from tall buildings. However, this diseconomies seems to be small with respect to the agglomeration effect present in same ZIP code where skyscrapers have been completed. In fact, for the Real estate sector the creation of new tall buildings leads to an increase of almost 20 percent of establishment in the same ZIP code where the structure have been constructed while almost a 0.03 percent increase in the neighbouring ares between 10 and 50 km. Results for the Real estate sector are similar to those obtained for Information, Education, Accommodation and Arts sectors.

Similarly, sectors in which I have not found any statistically significant effect of tall buildings in the previous section presents some evidence of increased location either between 10 and 25 km, as Manufacturing, or between 10 and 50 km, as Retail trade. The results obtained for the Manufacturing sector are mimic by the Construction and Transportation sectors, while the conclusion for Retail trade are also found for Wholesale trade,

Health Care and Administrative service sectors.

A particular case is the one represented by the Finance and Insurance sector. In fact, controlling for spatial lag of the completion of tall buildings I do not encounter anymore a significant effect in the same ZIP code. The same is true for Professional services and Management of companies sectors.

It is important to note that the results presented represent a net agglomeration effect, discounted by the possible negative congestion effect. This estimation does not clearly disentangle for agglomeration and congestion. Assuming that congestion will have a negative effect on firms' location and agglomeration economies a positive one I am estimating in this equation the net effect of the two. That is, it seems that agglomeration economies are stronger in the same ZIP code for Real estate, Information, Education, Accommodation and Arts sectors. The increase in location between 20 and 50 km away from the skyscraper could be either by caused by relocations given by congestion in the ZIP code where the building have been completed or increase location of new firms in those surrounding areas for complementary and sharing reasons.

Table 13: Spatial estimations

Distance	Log establishment				
	All sectors	Manufacturing	Retail trade	Real estate	Finance and insurance
Same ZIP	12.43 (9.538)	-4.616 (5.21)	8.033 (7.435)	20.94* (11.74)	14.68 (9.27)
<5 km	-0.0747 (0.051)	-0.00568 (0.0287)	-0.059 (0.0398)	-0.111* (0.0613)	-0.0722 (0.0491)
between 5 and 10 km	0.0133 (0.0179)	0.0204** (0.00896)	0.0201 (0.0142)	0.00617 (0.0228)	-0.00283 (0.0163)
between 10 and 25 km	0.0329*** (0.00961)	0.0113* (0.00618)	0.0302*** (0.0072)	0.0395*** (0.0118)	0.0242*** (0.00925)
between 25 and 50 km	0.0207*** (0.00704)	0.00787 (0.00536)	0.0139*** (0.00511)	0.0229*** (0.00791)	0.0211*** (0.00598)
between 50 and 100 km	0.00524 (0.00644)	0.00171 (0.00441)	0.00104 (0.00554)	0.00486 (0.0071)	0.00167 (0.00633)
Observations	174,876	174,876	174,876	174,876	174,876
MSA FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Natural advantage	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES
Estimation	IV	IV	IV	IV	IV

Clustered standard errors in parenthesis.
*, **, ***: statistically significant at 10, 5 and 1 percent respectively.

6. Estimation of agglomeration economies

The nature of my database allows to obtain a quantification of the agglomeration economies at a lower level of geographical disaggregation than the rest of the literature. In fact, despite higher density in one city or region leads to an overall increase in productivity, this effect can be different between different zones of the same areas. Because of diseconomies of scale and general equilibrium effects within a city, agglomeration economies might be higher or lower considering ZIP code level data. A second

contribution of my database is to allow me to estimate agglomeration economies using temporal variation. In fact, the completion of a tall building provides a shock in density in a clear moment of time. Previous estimations of agglomeration economies where uniquely exploit cross-sectional differences while in my study I can analyse a framework with important dynamic components. The exponential discount model will be used to consider this dynamic effects.

Differently from previous works, I can analyse agglomeration economies that are not caused by an increase in population or employment density but firms' density. Therefore, it would be possible to disentangle which part of the agglomeration economies are driven exclusively by an increase in the number of establishments instead of the total employment in the area.

The creation of new tall buildings are the key for my estimation of the magnitude of agglomeration economies in one area. In fact, tall buildings will create a variation in density through an increase in the height of the buildings in the area. In order to obtain exogenous variation I will use the same instrumental variable approach used in this paper. That is, the quantification of agglomeration economies will be provided by comparing differences in firms' productivity given by difference in firms' density originated from the completion of new tall buildings accrued to different soils and in years with different steel price.

In my estimations I will consider both a term referring to firms density and an additional term composed by the interaction between firms density and the completion of tall buildings. This second term will allow me to assess the additional increase in agglomeration economies provided by an increase in height of the buildings of a particular area.

Results of the estimation of agglomeration economies considering a dynamic model are presented in Table 14. The ex-post elasticity of log establishment density that is not passing through tall buildings is found to be 0.04, a result which is in line with the estimations found in the literature. This coefficient is not individually significant but it is jointly significant considering also its interaction with the construction of skyscrapers. The individual no significance of this coefficient could be explained by the fact that our specification that does not guarantee enough cross-sectional variation to estimate density effect on agglomeration alone.

Separating density that is coming via tall buildings and accounting for the anticipatory effects reduce significantly the estimation of agglomeration economies given by tall buildings with respect to a static framework. In fact, the ex-post effect of the completion of new tall building is about 0.017. However, since the elasticity of productivity to firms density is 0.04 and the elasticity considering tall buildings is 0.01, it is possible to say that

new tall buildings add a 20 percent higher increase in productivity on top of the increase given by firms density. Importantly, this additional productivity increase is statistically significant.

With this calculation is possible to have a first evidence that shocks of the density at more disaggregated level can be higher than what is found using MSA level variation, and in this case corresponds to 0.057. One important limitation of my estimation is that I cannot properly control for firms' sorting and this estimation can be smaller. However, part of this sorting is controlled by the fixed effects and by the agglomeration channels included in the estimation: input sharing characteristics, percentage of graduates and workers in managerial occupations, and patents.

Anticipatory agglomeration economies seem to be important even if they are smaller in magnitude than ex-post effects. That is, the additional increase in productivity to that one due to density that is accrued to tall buildings is present even before the completion of these structures. This is given by the existence of persistence in this estimations, which can be seen by an estimated discount term of 0.5.

Table 14: Dynamic elasticity estimation

Variables	Log productivity
F. log productivity	0.506*** (0.0171)
New tall buildings X log establishment density	0.00841** (0.00377)
Log establishment density	0.0236 (0.0229)
Calculated ex-post effects:	
New tall buildings X log establishment density	0.01702** (0.00765)
Log establishment density	0.04777 (0.04632)
Test anticipation effects:	
New tall buildings X log establishment density	0.00426** (0.00191)
Log establishment density	0.01194 (0.01158)
Observations	136,218
MSA FE	NO
Year FE	NO
Natural advantage	YES
Controls	YES
Estimation	AB
p-value Joint significance	0.000

Clustered standard errors in parenthesis
 *, **, ***: statistically significant at 10, 5 and 1 percent respectively.
 Instruments used: interaction distance bedrock and steel price (third lag),
 regional level of input sharing, education, high skills and patents,
 population density in 1900, leads of log productivity (from 2nd to 5th)
 lags of log productivity (from 1st to 5th),
 leads of interaction log establishment density and tall buildings (from 2nd to 5th).

7. Concluding remarks

The objective of this paper is to stress how urban structure, and in particular height of buildings, can act as a mechanism for agglomeration of firms' establishments. Controlling

for the classical agglomeration determinants, input sharing, labour pooling and knowledge spillover, firms might be attracted to areas in which tall buildings are constructed because of the productivity gains associated with this extreme form of density and the prestige associated with landmark buildings. Therefore, the contribution of this paper is to quantify the agglomeration impact of new tall buildings using a panel of more than 14,000 ZIP codes in U.S. from 2000 to 2012.

The empirical strategy in order to identify the effect of new tall buildings on agglomeration exploits the exogenous variation provided by geological and technological instruments. In particular, the completion of new skyscrapers have been instrumented using the interaction between the average depth to bedrock and the third lag Global steel price. Dynamic and spatial effects have been successively added in order to enrich our econometric model.

One of the most important results is that the effect of newly completed skyscraper on agglomeration differs between sectors. Sectors which are more related to the production of goods, such as Manufacturing or Construction, are not affected by the construction of new tall buildings. However, estimations for service sectors which are more likely to employ more human capital and high skill labour, such as Finance and Insurance, present a positive and significant effect of the completion of new tall buildings on agglomeration.

It has been possible to confirm that the agglomeration effect is not only driven by firms filling tall buildings but also the attraction of firms relates generally to all ZIP codes with the presence of skyscrapers. In fact, using an exponential discount model I encounter an anticipatory agglomeration effect of firms, which happen before the actual completion of the building.

New tall buildings possess an agglomeration effect for several sectors and this over-concentration of firms leads to an overall increase in productivity in the same ZIP code. However, introducing spatial elements of the construction of tall buildings reveals the presence of small diseconomies of scale and congestion effects in the surrounding areas.

Using the construction of tall buildings and the proposed instrumental variable procedure I provide a quantification of agglomeration economies at a low level of geographical aggregation that exploits time variation for identification. Controlling for agglomeration economies that are not passing through tall buildings and dynamic effects I find an elasticity of log establishment density given by the completion of tall buildings of 0.017 percent. That is, the magnitude agglomeration economies given by tall buildings add an extra 20 percent to the elasticity of productivity to firms' density.

Finally, one of the limitations of my estimations is related to the difficulty in distinguishing between congestion and agglomeration effects. The estimated effect is just the net increase of establishments in one area that can be given by a positive agglomeration

effects and negative congestion. Similarly, I cannot distinguish the agglomeration mechanisms driven by tall buildings, either if agglomeration is driven by present and future productivity increase or prestige of landmark buildings. Moreover, the increase in productivity following a construction of tall building can be given by the increase density or by firms sorting, that is more productive firms locating in that particular area. My estimation partially control for firm sorting and the presence of additional agglomeration economies given by tall buildings cannot be rejected. Despite the limitations of my work, the presented results already point out that urban structure cannot be neglected while studying firms location choice and that building height has important consequences for the attraction of establishments.

References

- Ali, M. M. and K. S. Moon (2007). Structural developments in tall buildings: Current trends and future prospects. *Architectural Science Review* 50(3), 205–223.
- Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The review of economic studies* 58(2), 277–297.
- Audretsch, D. B. and M. P. Feldman (1996). R&d spillovers and the geography of innovation and production. *The American economic review*, 630–640.
- Barr, J., T. Tassier, R. Trendafilov, et al. (2010). Bedrock depth and the formation of the manhattan skyline, 1890-1915. Technical report, Fordham University, Department of Economics.
- Bartik, T. J. (1991). Who benefits from state and local economic development policies? *Books from Upjohn Press*.
- Cervero, R. (2001). Efficient urbanisation: economic performance and the shape of the metropolis. *Urban Studies* 38(10), 1651–1671.
- Ciccone, A. and R. E. Hall (1996). Productivity and the Density of Economic Activity. *American Economic Review* 86(1), 54–70.
- Combes, P.-P., G. Duranton, and L. Gobillon (2008, March). Spatial wage disparities: Sorting matters! *Journal of Urban Economics* 63(2), 723–742.
- Combes, P.-P., G. Duranton, L. Gobillon, and S. Roux (2010, July). Estimating Agglomeration Economies with History, Geology, and Worker Effects. In *Agglomeration Economics*, NBER Chapters, pp. 15–66. National Bureau of Economic Research, Inc.
- Combes, P.-P. and L. Gobillon (2015). The empirics of agglomeration economies. *Handbook of regional and urban economics* 5.
- Dumais, G., G. Ellison, and E. L. Glaeser (2002). Geographic concentration as a dynamic process. *Review of Economics and Statistics* 84(2), 193–204.
- Dunn, A. (1993). *Structures: Skyscrapers*. New York: Thomson Learning.
- Duranton, G. and D. Puga (2004). Micro-foundations of urban agglomeration economies. *Handbook of regional and urban economics* 4, 2063–2117.
- Ellison, G. and E. L. Glaeser (1997). Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy* 105(5), 889–927.

- Ellison, G. and E. L. Glaeser (1999). The geographic concentration of industry: does natural advantage explain agglomeration? *American Economic Review*, 311–316.
- Ellison, G., E. L. Glaeser, and W. R. Kerr (2010). What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. *American Economic Review* 100(3), 1195–1213.
- Harari, M. (2015). Cities in bad shape: Urban geometry in india. Job market paper, Massachusetts Institute of Technology.
- Helsley, R. W. and W. C. Strange (2008). A game-theoretic analysis of skyscrapers. *Journal of Urban Economics* 64(1), 49–64.
- INSDAG (2013). Multi-storey buildings. Technical report, Institute for Steel Development and Growth.
- Jofre-Monseny, J., R. Marín-López, and E. Viladecans-Marsal (2011). The mechanisms of agglomeration: Evidence from the effect of inter-industry relations on the location of new firms. *Journal of Urban Economics* 70(2), 61–74.
- Khan, F. R. (1969). Recent structural systems in steel for high-rise buildings. In *Proceedings of the British Constructional Steelwork Association Conference on Steel in Architecture*. British Constructional Steelwork Association.
- Koster, H. R., J. van Ommeren, and P. Rietveld (2014). Is the sky the limit? high-rise buildings and office rents. *Journal of Economic Geography* 14(1), 125–153.
- Malani, A. and J. Reif (2015). Interpreting pre-trends as anticipation: Impact on estimated treatment effects from tort reform. *Journal of Public Economics* 124(C), 1–17.
- Marshall, A. (1920). *Principles of Economics*. London: Macmillan.
- Melo, P. C., D. J. Graham, and R. B. Noland (2009). A meta-analysis of estimates of urban agglomeration economies. *Regional science and urban Economics* 39(3), 332–342.
- Miller, D. and R. White (1998). A conterminous united states multi-layer soil characteristics data set for regional climate and hydrology modeling. *Earth Interactions* 2.
- NHGIS (2011). Minnesota population center’s national historical geographic information system: Version 2.0. Database, University of Minnesota.
- Puga, D. (2010). The magnitude and causes of agglomeration economies. *Journal of Regional Science* 50(1), 203–219.

Rosenthal, S. S. and W. C. Strange (2001). The determinants of agglomeration. *Journal of Urban Economics* 50(2), 191–229.

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

Rosenthal, S. S. and W. C. Strange (2008). The attenuation of human capital spillovers. *Journal of Urban Economics* 64(2), 373–389.