The taller the better?
Agglomeration determinants and urban structure

Universidad Carlos III de Madrid

Federico Curci

Submission for the XVIII Applied Economics Meeting

Work in progress

Abstract

The causes of higher productivity in denser areas have received important attention in the urban economics literature. Input sharing, labour pooling and knowledge spillover are considered as the classical determinants of agglomeration economies. In addition to these forces, firms may be attracted to particular locations because of the prestige associated to particular areas. 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 and employment in U.S. cities. In fact, 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 14114 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. Sectors which are characterized by higher use of human capital and high labour skills are associated with a positive and significant coefficient of new buildings completion on firms agglomeration, such as for the Finance and Insurance sector. The attraction of establishments on the ZIP codes where tall buildings will be completed has an important anticipatory component. At a ZIP level the completion of a new skyscraper is associated with a positive increase in productivity of 20 percent.

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. This implies that urban structure cannot be neglected when considering firm location. Moreover, different mechanisms have been proposed by the literature to explain agglomeration economies. The literature started with **?** 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 (?). Little work has been done on the analysis of skyscrapers impact on urban economic development. ? have assessed the existence of a building height premium. Firms might be willing to pay higher rents in floors at higher stories 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 firms' prestige for agglomeration, 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. Finally, it will be assessed the impact of the completion of new skyscrapers on ZIP codes productivity.

The empirical analysis is conducted using a rich database including all 6 digits NAICS sector, for 14114 ZIP codes for 147 Metropolitan Areas 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 North-American steel price.

2. Database

The database used for the empirical analysis have been 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 constructed too using the CBP database.

The number of completed tall buildings for non-residential 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 technologies[1]. 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.

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 **?**. 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:

$$I_{jt} = \sum_{j \neq k} W_{j,k} \times est_k \tag{1}$$

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 cointains 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 **?**. **?** 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.

The completion of tall buildings have been instrumented by the distance from bedrocks and the North-American steel value. We have constructed a variable cointaining the average depth to bedrock for each ZIP code in U.S. using the information provided by **?**. The North-American steel price indicator has been extracted from the CRU Steel Price Indicators.

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

3. Empirical strategy

This paper aims at establishing the role of tall buildings on agglomeration using an empirical approach. Our estimations face several econometric challenges: within-cluster correlation of the errors, reverse causality, omitted variable bias and time persistency. A Metropolitan Statistical Area is a proxy of a local labour market. Therefore, ZIPs in the same MSA share the same pool of labour and other inputs. In order to control for part of the within-cluster correlation of the error, I have performed a cluster-specific fixed effects estimation. 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. Therefore, in order to estimate empirically the effect of skyscrapers on agglomeration the following model 2 is estimated for each different j sector:

$$y_{zjt} = \tau_t + \mu_m + \beta D_{zt} + \gamma X_{mt} + \varepsilon_{mjt} \tag{2}$$

where z and m are the geographic units of interest (ZIP codes and MSA respectively), D_{zt} is the number of skyscrapers completed in one ZIP code in one given year, and X_{mt} is a number of control 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.

It is difficult to claim that the completion of new tall buildings is an exogeneous 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. 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 North-American steel price.

The relevance of these instruments are 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 **?**, implying that construction costs are higher in cities with more distant bedrocks from the surface. Exogeneity of the instrument is guaranteed by the random assignment of bedrocks and by the fact that each ZIP establishments cannot influence the North-American steel price. Moreveor, following **?** 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 agglomeration control proxies (input sharing, laobur pooling and knowledge spillover) have been instrumented using Bartik instruments (**?**). 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 Regions [2]. Relevance of this instrument is provided by the correlation existing between these proxies 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.

The previous model can be extended in order to include dynamic and spatial effects. In addition to understanding whether skyscrapers are attractive to firms, this analysis will allow to determine if being located in areas close to skyscrapers can increase firms' prestige. In fact, 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. This new model will be estimated using Arellano-Bond two-step, using both lags value of the dependent and the treatment variable, in order to consider the possible presence of time persistency.

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 2 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. 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.

4. Evidence

In this section I provide the main empirical evidence of the effect of tall buildings on agglomeration. Estimating the model presented in equation 2 without differentiating for each sector it is not possible to observe a statistically significant effect of the completion of new tall buildings on the number of establishments (see Table 1). This suggests the

[2]The following U.S. Census Regions have been used. *New England*: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont; *Middle Atlatic*: 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 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

possible existence of heterogeneous effects between sectors and the possible presence of dynamic effects.

Table 1: ZIP level estimation for the all NAICS sectors

	ZIP level
Variable	Log establishment
New tall buildings	57.93
	(44.57)
Input sharing	-0.0019*
	(.00109)
Education	-0.032
	(0.038)
High skills	0.0325
	(0.0433)
Patents	-0.000053
	(0.000098)
Natural advantage	1.606***
	(0.6000)
Observations	162,874
MSA FE	YES
Year FE	YES
Estimation	IV

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

Looking at the results for the additional controls included it is possible to evince that input sharing have a negative effect on agglomeration and natural advantage a positive effect. The proxy for input sharing used is a measure of co-agglomeration between industries since it includes information about the use of inputs from different sectors without considering own-sector input use. Therefore, a negative coefficient may suggest that at a ZIP level firms locate close to firms of the same sector and not of different sectors. The weak statistical significance of the controls is explained by the fact that our estimation is exploiting variation between ZIPs, while our controls are measured at a MSA level. However, a F-test suggests that it is important to introduce input sharing, labour pooling and spillover proxies in order to avoid within-cluster endogeneity [3].

Table 2 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 F-test of the first stage suggests that the instrument for the completion of new tall buildings might be weak. However, all the estimations provided in this section are robust to the use of Fuller estimation procedure in order to solve the weak instrument problem.

[3]The p-Value of the associated F-test is 0.0002

Table 2: First stage estimations

	First stage				
Variable	New building	Input	Education	High skill	Spillover
Distance bedrock X steel price	-2.91e-07***				
	(3.61e-08)				
Regional input sharing		0.4270***			
		(0.00268)			
Regional education			0.754***		
			(0.00758)		
Regional high skills				0.089***	
				(0.0064)	
Regional patents					0.97***
					(0.0051)
Observations	167,340	167,340	162,941	162,874	167,340
MSA FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Estimation	OLS	OLS	OLS	OLS	OLS
F	4.90	.	46,171.13	19,030.58	33,731.83

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

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 3 in which it is presented the coefficient for new tall buildings for separate estimations of Model 2 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, Transportation and Warehousing, Management of Companies and Enterprises, 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 80 and 88 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, Professional, Scientific, and Technical Services, Retail trade, Educational Services, Health Care and Social Assistance, Arts, Entertainment, and Recreation, and Accommodation and Food Services.

My database does not allow to separate the location of establishments inside and outside the new skyscrapers in one ZIP code. Therefore, in order to prove that the previous agglomeration effect is not just driven by firms filling new tall buildings I have estimated Model 2 with the inclusion of leads and lags of the completion of new tall buildings. As it is possible to evince from Figure 1 the agglomeration effect generated by new skyscrapers begins before the actual completion of the building. In the Figure it is presented the

Table 3: Sectoral estimations

Sectoral analysis		Sectoral analysis	
Sector	Log establishment	Sector	Log establishment
Utilities	5.343	Information	50.58*
	(4.119)		(26.13)
Construction	10.51	Finance and insurance	80.28**
	(19.98)		(40.47)
Manufacturing	47.54	Real estate	88.12**
	(30.14)		(42.87)
Wholesale trade	50.34	Professional	88.62*
	(31.03)		(50.61)
Transportation	13.63	Retail trade	61.78*
	(15.77)		(35.61)
Management	18.46	Educational	50.55*
	(11.52)		(26.33)
Administrative	39.03	Health care	80.83*
	(26.80)		(45.44)
Arts	46.77*	Accommodation	83.23**
	(25.46)		(41.16)
MSA FE	YES	MSA FE	YES
Year FE	YES	Year FE	YES
Estimation	IV	Estimation	IV

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

average year of proposal of a skyscraper (5.2 years before its completion) and the average year of beginning of construction (2.5 years before). The agglomeration effect given by new tall buildings begins even before the real construction of the structure but the coefficient associated to the new building increases considerably with its construction beginning. Moreover, agglomeration caused by tall buildings has long-lasting effect. Indeed, I have found a positive and significant coefficient until 5 years of its construction.

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 4 I have estimated the previous regression in order to assess the presence of agglomeration economies caused by the increase in density caused by the increase in height in the ZIP code. As it is observable the agglomeration effect leads to an increase in productivity in the same ZIP code of almost 20 percent. One important weakness of this estimation is connected with the existence of missing data for productivity because of confidentiality reasons. However, the estimation of the same model using Heckman techniques confirms the established result.

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 by increased congestion which could be leading to diseconomies of scale in the surrounding

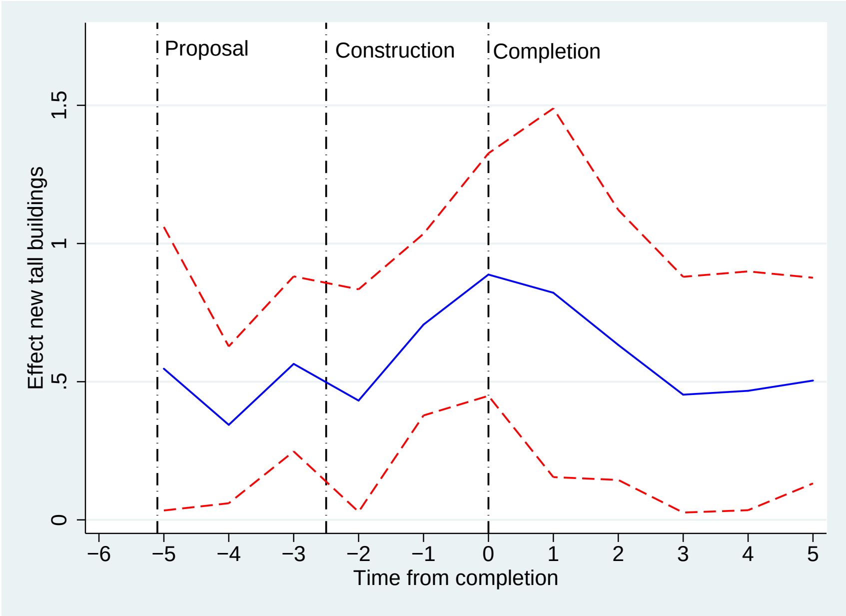

Figure 1: Dynamic effects of construction new tall buildings
In red the coefficient and in the blue the upper and lower confidence intervals. Coefficients estimates with Arellano-Bond two steps

Table 4: Productivity estimation

	ZIP level
Variable	Log Productivity
New tall buildings	20.73***
	(9.2395)
Input sharing	-0.00014
	(0.00026)
Education	-0.00859
	(0.00997)
High skills	0.00130
	(0.01202)
Patents	-2.12e-06
	(0.0000233)
Natural advantage	-0.03337
	(0.13343)
Observations	141,267
MSA FE	YES
Year FE	YES
Estimation	IV

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

areas. In Table 5 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 Manufacturing and the Finance sectors.

Generally, the presence of congestion effects and diseconomies of scale cannot be rejected. 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 5 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 Finance sector the creation of new tall buildings leads to an increase of almost 100 percent of establishment in the same ZIP code where the structure have been constructed while almost a 0.1 percent increase in the neighbouring ares between 5 and 50 km.

Table 5: Spatial estimations

	Log establishment		
Distance	All sectors	Manufacturing	Finance and insurance
Same ZIP	84.37	57.73	104.2*
	(61.19)	(42.86)	(55.90)
0 to 5 km	-0.419*	-0.409	-0.463
	(0.230)	(0.343)	(0.284)
5 to 10 km	0.109*	0.196***	0.0954*
	(0.0579)	(0.0672)	(0.0577)
10 to 25 km	0.181***	0.149**	0.169**
	(0.0679)	(0.0668)	(0.0703)
25 to 50 km	0.103**	0.0704	0.141***
	(0.0493)	(0.0443)	(0.0480)
50 to 100 km	0.0229	0.0162	0.0302
	(0.0400)	(0.0386)	(0.0450)
Observations	162,874	121,685	123,396
MSA FE	YES	YES	YES
Year FE	YES	YES	YES
Estimation	IV	IV	IV

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

5. 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 been 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 has exploited the exogeneous 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 North-American 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 the agglomeration effect is not only driven by firms filling tall buildings but that the attraction of firms relates generally to same ZIP code where skyscrapers have been constructed. In fact, I have encountered an anticipatory agglomeration effect of firms, which happen before the actual completion of the building. Moreover, the effect of tall buildings on agglomeration has been found to last for 5 years.

It has been encountered that new tall buildings possess an agglomeration effect for several sectors and that 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 it has been assessed the presence of small diseconomies of scale and congestion effects in the surrounding areas.

Finally, further work needs to be carried over in order to strengthen the robustness of the presented results. In particular, effort should be devoted in order to distinguish the presence of an agglomeration effect caused by the prestige and reputation of tall buildings. Nevertheless, 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

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*.

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.

Ellison, G. and E. L. Glaeser (1999). The geographic concentration of industry: does natural advantage explain agglomeration? *American Economic Review*, 311–316.

Helsley, R. W. and W. C. Strange (2008). A game-theoretic analysis of skyscrapers. *Journal of Urban Economics 64*(1), 49–64.

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.

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.

Marshall, A. (1920). *Principles of Economics*. London: Macmillan.

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*.

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.