Three essays in urban economics

Federico Curci

en cumplimiento parcial de los requisitos para el grado de Doctor en

Economía

Universidad Carlos III de Madrid

Directores:

Ricardo Mora Villarrubia
Adelheid Holl

Año Académico 2018/2019

AGRADECIMIENTOS

Writing the acknowledgment is probably one the hardest parts of my thesis, since I am indebted to so many people. First of all, I am extremely grateful to Ricardo Mora, my main supervisor, for all his devotion, the time spent with me, the continuous support and his optimism. His invaluable help, both professionally and personally, has been fundamental for the development of my thesis and my professional career. Similarly, I want to thank Adelheid Holl for the supervision of my Master's thesis, which finally converted in one of the chapter of this thesis, and her comments about the other chapters. I also thank Klaus Desmet for his supervision during my visit at Southern Methodist University. He was a great host, always accessible and eager to comment my work.

I also thank the department of economics at Universidad Carlos III de Madrid and in particular the applied economics group. I have enormously benefited from the accessibility and willingness to help of the faculty. This thesis has improved thanks to the comments and suggestions of many people, in particular Julio Caceres, Irma Clots, Jesus Fernandez-Huertas, Matthias Kredler, Luigi Minale, Jaime Millan, and Jan Stuhler. I want to thank Andres Erosa, Luisa Fuster and Marco Celentani for their continuous efforts in supporting the Master and PhD programs. Similarly, thanks to Aranzazu Alonso and Angelica Aparicio for their administrative help.

It has been easier to write this thesis in a wonderful environment surrounded by exciting, bright and supportive friends and PhD colleagues. I have met wonderful people who made me enjoying the writing of this thesis and my stay at Universidad Carlos III de Madrid. First, I am indebted to my Master's friends, Andrea Cintolesi, Carmen Garcia, Clara Martinez-Toledano, Cleiton Guollo, Derek Lougee, Emiliano Tealde, Ivan Zilic, Jose Montalban, Luigi Briglia, and Mattia Ricci, for our common support, solidarity and authentic friendship during some tough time. I thank Andres Garcia-Suaza, Christos Mavridis, Iacopo Morchio, Marco Serena, Mehdi Hamidisahneh, Meng Li, and Sebastian Panthofer for welcome me in the PhD programme after the leave of many Master's friend. I also thank the new generations in the PhD, Tomas Rodriguez, Ursula Mattioli, Javier Rodriguez, Michelangelo Rossi, Ruben Veiga, Onursal Bagirgan, Fabrizio Leone, Carla Varona and Francisco (Pancho) Pareschi, for joining with enthusiasm the group and contributing to keep great the time in Madrid. I thank all the Dinamo Cuore's member, our heroic football team, which year after year continuously qualifies for the university's playoffs.

I do not know how to thank Beatriz Gonzalez, Federico Masera, Hasin Yousaf, Ismael Galvez, Jose Carbo, Michele Rosenberg, Salvatore Lo Bello, for always being there, cheering me up, supporting me, listening me, giving me advices and creating friendships that will last forever. In particular, I thank Ismael and Salvatore for devoting so much efforts in organizing activities and transmitting me their love for Real Murcia and AS Roma. I thank Beatriz for always being sensible to every problem and having a "majo" for everybody. I thank Federico and Michele for their contagious passion about research, science, life and having a discussion for any possible topic that a person can imagine. I thank Hasin and Jose for being the best office mates one can desire and being supportive like family to each other.

My parents, Giovanni and Eleonora, and my sister Valentina have always being determinant for every success in my life. They always support me unconditionally, encourage me, believe in me, and remember me what are the priorities in life.

I dedicate this thesis to my loves, Lucia and Matilde. Matilde, anche se sei arrivata alla fine di questa esperienza è come se ci fossi sempre stata. Lucia, non so come avrei fatto senza di te e senza il tuo continuo sostegno. Come sempre, io e te tre metri sopra la tesi.

CONTENIDOS PUBLICADOS Y PRESENTADOS

Parte de esta tesis ha sido publicada como working paper. En particular,

- Una versión anterior del capítulo 1 ("The Taller the Better? Agglomeration Determinants and Urban Structure") se puede encontrar en el siguiente working paper:
    - Federico Curci, 2015. "The taller the better? Agglomeration determinants and urban structure," *ERSA conference papers ersa15p991*, European Regional Science Association. URL: https://ideas.repec.org/p/wiw/wiwrsa/ersa15p991.html

- El capítulo 3 de la tesis ("Flight from Urban Blight: Lead Poisoning, Crime and Suburbanization") ha sido publicado en el siguiente working paper:
    - Federico Curci & Federico Masera, 2018. "Flight from urban blight: lead poisoning, crime and suburbanization," *Working Papers 2018/09*, Institut d'Economia de Barcelona (IEB). Available at SSRN: https://ssrn.com/abstract=3245090 or http://dx.doi.org/10.2139/ssrn.3245090

El material de esta fuente incluido en la tesis no está señalado por medios tipográficos ni referencias.

Abstract

Cities are central to the live of the majority of the human population. While the proportion of people moving to cities has increased steadily, the population living in city centers has decreased in many developing and developed countries. My thesis focuses in understanding what are the causes and consequences of different urban developments.

A question I tackle in my research is what are the effects of firm and residential density. In the first chapter of my thesis (*The Taller the Better? Agglomeration Determinants and Urban Structure*), I estimate the productivity gains of an extreme form of urban density: skyscrapers in cities. In the second chapter of my thesis (*Vertical and Horizontal Cities: in Which Direction Should Cities Grow?*), I combine both data and theory in order to structurally estimate the effects of different city development on productivity and amenities. I show that cities that are more vertically developed have higher level of amenities.

A second strand of my thesis focuses on understanding how people and firms decide to locate inside cities. In particular, in the third chapter of my thesis (*Flight from Urban Blight: Lead Poisoning, Crime and Suburbanization*), joint with Federico Masera, we provide causal evidence that crime has been an important reason to explain suburbanization of U.S. cities.

1 The Taller the Better? Agglomeration Determinants and Urban Structure

This paper explores how urban structure and building height play an important role for agglomeration and the consequent productivity advantages. I do this by looking at the role of skyscrapers in influencing the concentration of establishments in U.S. cities. In addition to productivity advantages associated to this extreme form of density, skyscrapers are an attractive location for firms because of the associated gains in prestige from being located in a tall landmark building. The agglomeration effects of tall buildings have been identified instrumenting the completion of new skyscrapers by the interaction between the distance to bedrocks in one ZIP area with the Global steel price.

Results suggest that tall buildings have an effect on the location of firms inside a city. Tall buildings increase both agglomeration of firms in the surrounding area and their productivity. The effect of newly completed skyscrapers on agglomeration differs between sectors. The attraction of establishments to ZIP codes where tall buildings will be completed has an important anticipatory component. Exploiting the variation of firm's density produced by tall buildings, I find that agglomeration economies caused by tall buildings provide an additional 20 percent increase in productivity. That is, I estimate that firms' productivity elasticity to establishment density is 0.05, while the additional productivity elasticity if the firm locate in a skyscraper is 0.01.

2 Vertical and Horizontal Cities: in Which Direction Should Cities Grow?

This research establishes the different consequences of taller or more spread out cities using a strategy that combines reduced form estimation with a more structural approach. In order to achieve this goal I build a general spatial equilibrium model which includes both within-city and between-cities spatial equilibrium concepts. Results from IV regressions suggest that both vertical and horizontal increase of a city, measured as an increase in the floors of buildings and in total lot size occupied by buildings in a city, are associated with a positive increase in house prices, and no statistically significant effect on wage. However, increasing a city vertically would lead to a higher effect on house prices with respect to increasing it horizontally. These reduced form estimates are consistent with a calibration of the model in which building height and city size have a similar positive effect on city-specific productivity, while height has stronger positive effect on city-specific amenities than city size.

3 Flight from Urban Blight: Lead Poisoning, Crime and Suburbanization

In the post World War II period, most U.S. cities experienced large movements of population from the city centers to the suburbs. In this paper we provide causal evidence that this process of suburbanization can be explained by the rise of violent crime in city centers. We do so by proposing a new instrument to exogenously predict violent crime. This instrument uses as time variation the U.S. national levels of lead poisoning. Cross-sectional variation comes from a proxy for soil quality, which explains the fate of lead in soil and its subsequent bioavailability. Using data for more than 300 U.S. cities, results show that the increase in violent crime from the level in 1960 to its maximum in 1991 decreased the proportion of people living in city centers by 15 percentage points. This increase in crime moved almost 25 million people to the suburbs. As a result of suburbanization, we find that people remaining in the city center are more likely to be black people, consistent with the “white flight" phenomenon. We then demonstrate that this suburbanization process had aggregate effects on the city. Exploiting a spatial equilibrium model, we determine that violent crime had externalities on productivity and amenities.

Table of contents: Chapter 1
The taller the better? Agglomeration determinants and urban structure

Table of contents: Chapter 2
Vertical and horizontal cities: in which direction should cities grow?

Table of contents: Chapter 3
Flight from Urban Blight: Lead Poisoning, Crime and Suburbanization

Table of contents: Chapter 3
Flight from Urban Blight: Lead Poisoning, Crime and Suburbanization
Online Appendix

The taller the better?
Agglomeration determinants and urban structure

Federico Curci [*]

Abstract

This paper explores how urban structure and building height play an important role for agglomeration and the consequent productivity advantages. I do this by looking at the role of skyscrapers in influencing the concentration of establishments in U.S. cities. In addition to productivity advantages associated to this extreme form of density, skyscrapers are an attractive location for firms because of the associated gains in prestige from being located in a tall landmark building. The agglomeration effects of tall buildings have been identified instrumenting the completion of new skyscrapers by the interaction between the distance to bedrocks in one ZIP area with the Global steel price.

Results suggest that tall buildings have an effect on the location of firms inside a city. Tall buildings increase both agglomeration of firms in the surrounding area and their productivity. The effect of newly completed skyscrapers on agglomeration differs between sectors. The attraction of establishments to ZIP codes where tall buildings will be completed has an important anticipatory component. Exploiting the variation of firm's density produced by tall buildings, I find that agglomeration economies caused by tall buildings provide an additional 20 percent increase in productivity. That is, I estimate that firms' productivity elasticity to establishment density is 0.05, while the additional productivity elasticity if the firm locate in a skyscraper is 0.01.

[*]Universidad Carlos III de Madrid. Departamento de Economia. 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, Jesus Fernandez-Huertas, Federico Masera, Luigi Minale, Marco Pagliero, Jan Stuhler, Ulrich Wagner, and the participants to 3rd Madrid Microeconomics Graduate Workshop, UC3M PhD Workshop, 20th Spring Meeting of Young Economists, 2015 SERC Annual conference, 5th European Urban Economics Association meeting, Infrastructure and the Spatial Organization of Economic Activity Workshop at BBVA Foundation, 2015 Econometric Society European Winter Meeting, and 8th Italian Doctoral Workshop in Empirical Economics for valuable comments. Financial support from the Spanish Ministry of Economics and Competitiveness (ECO2012-33053) is acknowledged.

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 or size. Studies comparing different cities in one point in time found an elasticity of productivity to city size of 0.04. In this paper I estimate agglomeration economies exploiting shocks in firms' density within a city given by the construction of skyscrapers. This procedure has three main advantages. First, it allows me to obtain an estimation of the benefits of agglomeration in terms of productivity looking at very fine level geographical observation inside a city. Second, I am able to exploit the dynamics of the increase of density to estimate agglomeration economies. Third, I can estimate separately agglomeration economies coming from either an horizontal or a vertical increase in density, that is through the increase in building height.

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 on how skyscrapers, and more generally vertical density, 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. Similarly, Liu et al. (2018) estimate vertical agglomeration economies by looking at the vertical rent gradient, and how rents change inside a building by increase floors. 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. The empirical analysis is conducted using a rich database including all NAICS sector, for 14,114 ZIP codes in 147 Metropolitan Statical Areas (MSA) in U.S. from 2000 to 2012. This database has 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 past Global steel price. Since tall buildings are particularly vulnerable to earthquakes and wind, they need to be anchored to bedrocks. As a result if a bedrock is closer to the surface, construction costs of tall buildings decrease. Moreover, the instrument exploits the fact that the construction of tall buildings require a large use of steel to sustain the structure, in particular if the bedrock lies far away from the surface. One of the main advantages of this instrument is that it varies both in time and geography, allowing the execution of fixed effects techniques.

I show that tall buildings increase agglomeration of firms in the surrounding area. Moreover, they also increase the number of establishments in the all city. As a result

I estimate that the increase in vertical density of firms establishment has a positive effect on ZIP code and MSA productivity. I found that vertical density increases input sharing between firms and the creation of new patent at city levels, which consecutively increase productivity. The effect of tall buildings on productivity is present even controlling for proxies of the classical determinants of agglomeration economies: input sharing, labour pooling and knowledge spillovers. This result suggests 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.

An additional result of the paper is that the construction of tall buildings in one specific ZIP code has an effect on the location of firms in the all city. I found that tall building increases agglomeration of firms even in neighbour ZIP codes. I show the completion of new tall buildings have an agglomeration effect that depends on the sector under consideration. While firms in the majority of service sectors are attracted by the area in which tall buildings are constructed, firms in the manufacturing and wholesale trade decrease their presence in those parts of the city. As a result, I show that distribution of firms inside a city does not become more concentrated in the ZIP codes where tall buildings are built. This is consistent with the idea that tall building can potentially influence the formation of different employment centers in the city. Since firms from different sectors agglomerate close to tall buildings, I do not find that ZIP codes becomes more specialized. However, I find that cities become more specialized.

The agglomeration effect given by new tall buildings begins even before the construction of the structure. Anticipatory effects have been estimated using an exponential discount model. Using this model I have also obtained agglomeration estimates that are not biased to the increase in the number of establishment previous to the completion of tall buildings. I find that after the completion of tall buildings the overall number of establishments in one ZIP code increases by 78%. Furthermore, I estimate that tall buildings have long-lasting agglomeration effects.

The last section of this paper gives a quantification of agglomeration economies, that is the elasticity of productivity with respect to firms' density. Differently from the literature, this estimation has been done exploiting the fact that tall buildings provide a clear shock in firms' density at lower level of geographical aggregation. 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. That is, I estimate that firms' productivity elasticity to establishment density is 0.05, while the additional productivity elasticity if the firm locate in a skyscraper is 0.01. Looking at shocks at more disaggregated level of geography, overall agglomeration economies has been found to be higher than previous estimates in the 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, Puga, 2010, and Combes and Gobillon, 2015 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. This sources of agglomeration allow cities to have higher productivities and therefore to attract more firms.

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. Liu et al. (2018) estimate the vertical rent gradient inside skyscrapers. They find that the rent gradient is non-monotonic in height, and that high productivity companies locate in higher floors. They conclude that within-building employment has larger impact on rent than nearby employment outside the skyscrapers. Similarly, Danton and Himbert (2018) estimate the vertical rent gradient for residential buildings.

Other works on skyscrapers have focused on why they are built and their welfare gains. Ahlfeldt and McMillen (2017) considers that vertical development of a city is the result of high land prices, that is, developers respond to increasing land prices by increasing density through building taller. Bertaud and Brueckner (2005) develop theoretical predictions of building height regulations, while Brueckner and Sridhar (2012), and Brueckner et al. (2015) estimate the welfare gains in terms of commuting costs in the Indian and Chinese cases, respectively. Borck (2014) provides a theoretical framework to see the impact of skyscrapers in reducing pollution.

In this paper I assess 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 provide further evidence in the estimation of agglomeration economies by looking at static and dynamic density shocks that happen at very disaggregated level within city.

The rest of the paper is structured as follows. I describe the identification strategy exploited and the data used in Section 2. I provide the main static results about the relationship between agglomeration and completion of tall buildings in Section 3. I discuss the dynamic empirical strategy and the dynamic results in Section 4. Section 5 presents the estimation of agglomeration economies. Section 6 provides concluding remarks.

2 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 persistence. In order to control for part of the within-cluster correlation of the error, I have performed a cluster-specific fixed effects estimation. Standard errors are clustered at ZIP level. In order to control for time-invariant unobserved heterogeneity which can explain the location of tall buildings in one ZIP code I control for ZIP code fixed effects, τ_z.

In order to estimate empirically the effect of skyscrapers on agglomeration the following model 1 is estimated for each different sector j:

$$y_{zjt} = \tau + \tau_z + \tau_t + \beta D_{zt} + \varepsilon_{zjt} \tag{1}$$

where z and m are the geographic units of interest (ZIP codes and MSA respectively), D_{zt} is the stock of new skyscrapers completed in one ZIP code in the previous years [1]. The dependent variables y_{zjt} used are the number of establishments of sector j in ZIP area z and the log productivity measure. Section 2.2 discusses how I have measured these variables.

2.1 Instrumental variable

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. Evidence for the existence of this reverse causality has been found by Ahlfeldt and McMillen (2017). Moreover, omitted variable bias can also be present if

[1] Since 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.

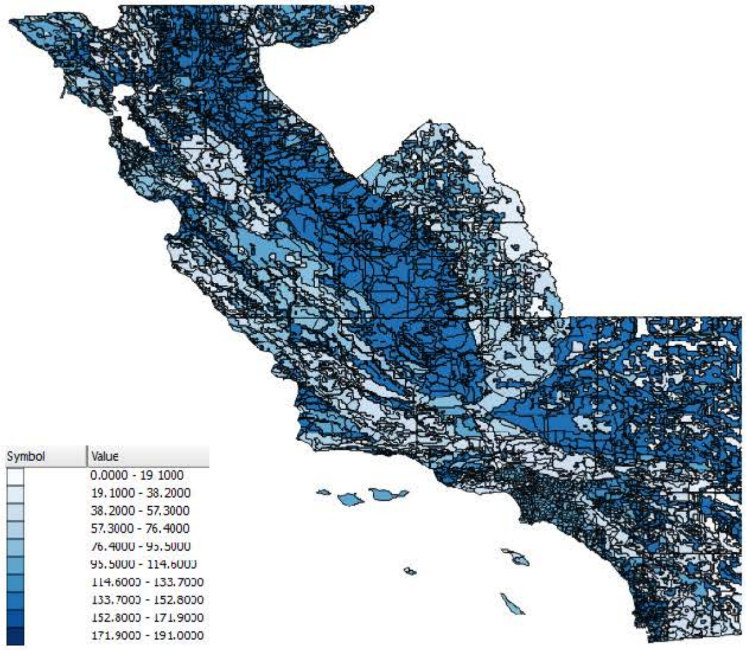


Figure 1: Distance to bedrock (in centimeters) in California

the construction of skyscrapers happens 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 time-variant endogeneities, I have instrumented the number of completed 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 it provides 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). This implies 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). As it is possible to see from the example of California in Figure 1 bedrock distance provide cross-sectional variation at a very low level of geographical disaggregation.

Steel is particularly important for the construction of tall buildings because of the principle developed by Khan (1969) called “premium for height”. 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. This particular building technology

Figure 2: Global steel price time series

reassures us that steel price mainly influences the costs of tall buildings and not the construction costs of any building. Appendix A gives further empirical evidence to this claim.

The impact of bedrocks on the construction of tall buildings depend on the availability of steel. According to Dunn (1993) where the distance of 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"[2]. Therefore, in years when steel is more expansive, the negative effect of bedrocks on construction costs should be bigger. My instrument interacts distance to bedrocks with past global steel price. The resulting instrument changes both within time and space, allowing the inclusion of geographical and time fixed effects.

My instrument obtains time variation from the third lag of global steel price. Using global steel price, instead of local prices, I exploit variation in steel price that are not related to local construction markets. Figure 2 reports the time variation of the Global Steel Price. According to Economist (2012) the main reason for the increase in global steel price between 2000 and 2003 is the increase of steel imports from China. In fact, China's imports increased by three times, in a period when investments in the mining industry were low.

Table 1 shows the first stage results. From Columns (1) to (4) it is possible to see that further away is a bedrock the lower the construction of tall buildings in one ZIP code. This result, using recent information for the all U.S., contradicts the finding of Barr et al. (2010) for the city of New York only at the beginning of the 20th century. Having a bedrock at 10 meters away reduces the construction of tall buildings by 0.25 units. The negative impact of bedrock depths to the construction of tall buildings is robust to the inclusion of MSA fixed effects, MSA specific trends and clustering of the standard errors.

[2] More information at http://www.madehow.com/Volume-6/Skyscraper.html

Table 1: First stage results

VARIABLES	(1) Tall build.	(2) Tall build.	(3) Tall build.	(4) Tall build.	(5) Tall build.	(6) Tall build.	(7) Tall build.	(8) Tall build.
Dist. bedrock (10 m)	-0.332***	-0.332***	-0.250***	-0.250***				
	(0.0150)	(0.103)	(0.0795)	(0.0799)				
Bedrock x Steel price (lag 3)					-0.562***	-0.955***	-0.704***	-0.704**
					(0.169)	(0.257)	(0.201)	(0.354)
Observations	181,120	181,120	181,120	181,120	179,298	179,214	179,214	179,214
R-squared	0.004	0.004	0.012	0.015	0.004	0.723	0.726	0.726
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
ZIP FE	NO	NO	NO	NO	NO	YES	YES	YES
MSA FE	NO	NO	YES	YES	NO	NO	NO	NO
MSA x Year FE	NO	NO	NO	YES	NO	NO	YES	YES
Estimation	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Cluster s.e.	NO	ZIP	ZIP	ZIP	ZIP	ZIP	ZIP	MSA
F	491.80	10.40	9.90	9.80	11.10	13.80	12.30	4.00

Tall build.: New stock of tall buildings (not uniquely residential) in the ZIP code. *Dist. Bedrock (10 m)* / *Bedrock*: distance from earth surface to bedrock in 10 meters. *Steel price (lag 3)*: third lag of global steel price, normalized by the maximum value of steel price between 1997 and 2012. *Year FE*: year fixed effects. *ZIP FE*: ZIP code fixed effects. *MSA FE*: MSA fixed effects. *MSA x Year FE*: MSA times year fixed effects. *Cluster s.e.*: level of clustering of standard errors. *F*: F-statistics on the excluded instruments. *** p<0.01, ** p<0.05, * p<0.1

Columns (5) to (8) shows that the interaction between bedrock depth and global past steel price is negative, meaning that after an increase in steel price ZIP codes where the bedrock is further away construct less tall buildings. This result is robust to the inclusion of ZIP fixed effects but also the inclusion of MSA specific trends to allow the comparison of ZIP codes uniquely inside the same MSA. Moreover, the decision of the level of clustering of the standard errors does not affect this result. Looking at the F-statistics, clustering standard errors at MSA level reduces the relevance of the instrument leading to the potential of weak instrument. In the baseline specification (column 6) I control for ZIP and year fixed effects and cluster standard errors at ZIP code to allow correlation between time in the same ZIP code. Using this specification we can conclude that increasing steel prices from the level of 1997 to the level of 2011 (when it was maximum), leads to a reduction of 0.5 tall buildings built for every 10 meters of distance to a bedrock[3]. The F-statistics on the excluded instrument in this specification is 13.8, suggesting that my estimation does not suffer of weak instrument problems.

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

[3]This is given the fact that steel price is normalized in the regression and takes 1 when steel price is maximum. Since in 1997 global steel price index was 99.1 and in 2011 it was 186.8, the effect of the interaction between bedrock distance (in 10 meters) and steel price is $\frac{99.1}{186.8} * (-0.955)$=-0.507, with a standard error of 0.136

Table 2: First stage results using different lags of steel price

VARIABLES	(1) Tall build.	(2) Tall build.	(3) Tall build.	(4) Tall build.
Bedrock x Steel price (lag 1)	-0.862***			
	(0.234)			
Bedrock x Steel price (lag 2)		-0.897***		
		(0.242)		
Bedrock x Steel price (lag 3)			-0.955***	
			(0.257)	
Bedrock x Steel price (lag 4)				-0.951***
				(0.259)
Observations	179,986	179,493	179,214	179,015
R-squared	0.723	0.723	0.723	0.723
Year FE	YES	YES	YES	YES
ZIP FE	YES	YES	YES	YES
Estimation	OLS	OLS	OLS	OLS
Cluster s.e.	MSA	MSA	MSA	MSA
F	13.50	13.70	13.80	13.40

Tall build.: New stock of tall buildings (not uniquely residential) in the ZIP code. *Bedrock*: distance from earth surface to bedrock in 10 meters. *Steel price (lag 3)*: third lag of global steel price, normalized by the maximum value of steel price between 1997 and 2012. *Year FE*: year fixed effects. *ZIP FE*: ZIP code fixed effects. *Cluster s.e.*: level of clustering of standard errors. *F*: F-statistics on the excluded instruments. *** p<0.01, ** p<0.05, * p<0.1

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". A tall building has started construction "once site clearing has been completed and foundation/piling work has begun" [4]. 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. This result is also confirmed by the fact the use of the third lag of steel price provides the highest F-statistics and the strongest first stage effect (see Table 2). All the results reported in this paper do not depend on the particular use of the third lag of steel price.

To sum up, identification comes from the comparison of ZIP codes with different bedrock distance before and after the increase in global steel price given by China import. The identification assumption is that the interaction between distance to bedrock and steel price is affecting agglomeration and productivity in one ZIP code only through its effect on the construction of tall buildings. Exogeneity of the instrument is guaranteed by the random assignment of bedrocks. Moreover, exogeneity is

[4] Quotation comes from http://www.ctbuh.org/HighRiseInfo/TallestDatabase/Criteria/tabid/446/language/en-US/Default.aspx

met if local determinants of establishment location at ZIP levels are not influenced by the past Global steel price. Ellison and Glaeser (1999) have underlined the importance of natural advantages for agglomeration. ZIP fixed effects controls for possible natural advantages. I have also 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.

2.2 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. According to the CTBUH, 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[6]. Proportion is measured using size and floor area. Building technologies refer to the use of specific vertical transport technologies. In general, a building with 14 or more stories or over 50 meters tall where at least 50 percent of its height is occupied by usable floor area can be considered as tall. I have dropped all tall buildings which are uniquely for residential use.

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.

In order to understand the mechanisms behind the effect of tall building on agglomeration I have obtained data 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[8]. For each sector j input sharing have been measured sum-

[5] It is important to notice the existence of missing data for productivity because of confidentiality reasons. If any, this can potentially create a downward bias in my estimation. In fact, productivity data are missing in ZIPs with a higher number of establishments and 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.

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

[7] For 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

[8] Proxies for input sharing, labour pooling and knowledge spillovers are measured at MSA level.

ming 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:

$$I_{jt} = \sum_{j \neq k} W_{j,k} \times est_k \tag{2}$$

Labour pooling and knowledge spillovers have been measured using the proportion of population with at least a Bachelors' degree and the proportion of population in Management, professional, and related occupations[9]. These data are collected from the American Community Survey. Moreover, my 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 Rosenthal and Strange (2004). Finally, I have collected data for natural advantages using a dummy if the MSA is either coastal or on the Great Lake.

2.3 Summary statistics

Table 3 presents summary statistics for the variables used in this study. It also 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 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 also happen to be on the coast or on the Great Lakes region.

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 3 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 4).

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 2000. Primary Metropolitan Statistical Areas has been used in presence of Combined Metropolitan Stastical Areas.

[9]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

Table 3: Summary statistics

Variable	Mean	s.d.	Min	Max	Diff.	Obs.
Number of establishments, ZIP	317.946	467.87	1	7549	.	185539
Number of establishments, MSA	30869.350	39775.34	4863	270846	53057.47***	1911
Productivity, ZIP	34.950	18.93	2	2121	.	160093
Productivity, MSA	37.160	8.81	19	90	5.51***	1911
New tall buildings (no res.), ZIP	0.001	0.04	0	4	.	185539
New tall buildings (no res.), MSA	0.119	0.65	0	13	0.52***	1911
Stock tall buildings (no res.), ZIP	0.113	1.62	0	80	.	185539
Stock tall buildings (no res.), MSA	10.939	48.63	0	574	38.68***	1911
Distance to bedrock (in cm), ZIP	120.522	28.76	0	176	.	181120
Distance to bedrock (in cm), MSA	126.470	17.12	53	152	-13.82***	1911
Input sharing, MSA	1424.489	555.67	497	3231	541.86***	1911
Education (more BA), MSA	27.699	7.45	11	59	3.30***	1857
High skill workers, MSA	34.668	5.25	21	54	1.89***	1855
Patents, MSA	476.962	917.39	1	11490	666.85***	1911
Natural advantage, MSA	0.333	0.47	0	1	0.10***	1911
Steel price, yearly	139.885	39.78	68.90	186.80	.	13

s.d.: Standard deviation. *Diff.*: Difference in mean of the variable between cities hat have completed at least one tall building in the period 2000 and 2012 and those that did not. *Diff.* obtained regressing variable under interest on dummy for construction of a tall building in the city between 2000 and 2012. For variables that change over ZIP *Diff.* has not been computed. *Obs.*: number of observations. *ZIP*: variable measured at ZIP level. *MSA*: variable measured at MSA level. variable measured at ZIP level. *no res.*: no residential. *More BA*: proportion of people with education achievement higher than BA. Measurement of variables described in Section 2.2.

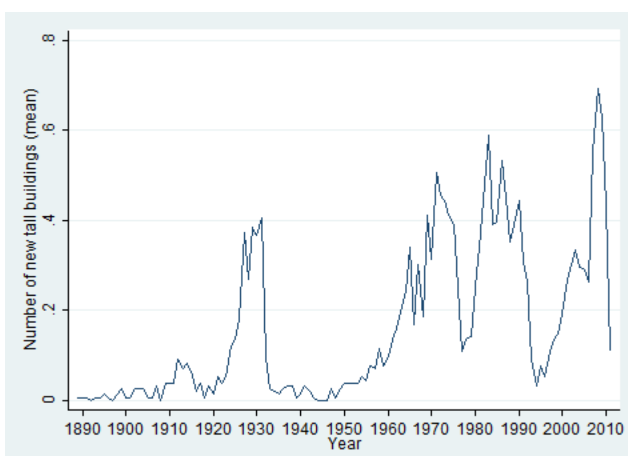


Figure 3: Mean number of new tall buildings completed by city

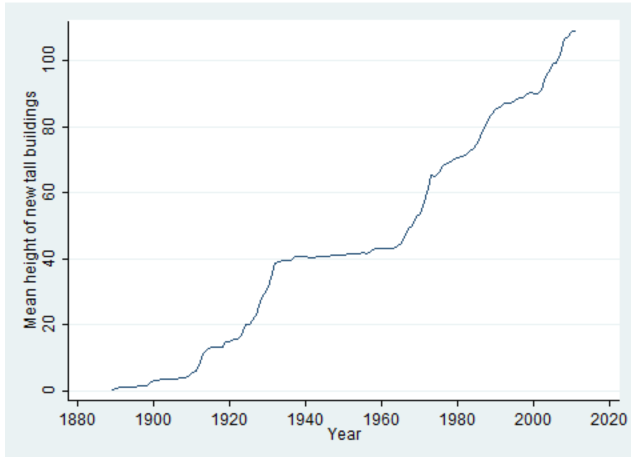


Figure 4: Mean height of new tall buildings completed by city, in meters

Table 4 reports some statistics about the current U.S. construction of tall buildings. In the period 2000-2012, 546 new tall buildings have been completed. Out of this 546 buildings, 228 are structures which use is uniquely not residential (office or hotels). That means that in 2012 there were a total of 2625 tall buildings, out of which 1713 are not used uniquely for residential reasons. Table 5 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 4: 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

3 Static analysis

3.1 Static regressions

I have exploited variation between ZIPs in the same MSA and in different MSAs in order to understand whether tall buildings can increase the location of firms in particular areas of a city. Results are reported in Table 6. Column (1) shows that the OLS estimation produce a non-significant effect of tall buildings on ZIP code agglomeration, measured as the total number of establishments. However, Column (2) shows

Table 5: 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
Seatlle-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

that after the increase in international steel price ZIP codes with bedrocks lying far away from the surface have experienced a decrease in firms' agglomeration. I use the interaction between steel price and bedrock depth to obtain IV estimates of the effect of tall buildings on firms' agglomeration. Column (3) reports the IV estimation. Building a new tall building in a ZIP code increases the overall number of establishments in one ZIP code by 194 units. The effect of tall buildings is statistically and economically significant, since on average a ZIP code has 318 establishments.

Despite the existence of agglomeration effect caused by new tall buildings, this does not automatically imply that agglomeration economies might be existing, that is whether there is an increase in productivity caused by higher density. Estimating the effect of tall buildings on productivity, it is possible to obtain an estimation of the effect of vertical density on agglomeration economies. In Column (4) of Table 6 I have estimated 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 23.5 percent. This increase in productivity could either come from increased agglomeration economies or from firms sorting in that ZIP code.

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 Tables 7 and 8 in which the coefficient for new tall buildings for separate estimations of Model 1 for the different NAICS sectors are reported. Table 7 shows the NAICS sectors that experience an increase in agglomeration in the ZIP codes in which tall buildings are built. A positive and significant effect of tall buildings on agglomeration has been found for the following sectors: *Construction, Real Estate and Rental and Leasing, Professional, Scientific, and Technical Services, Educational Services, Health Care and Social Assistance, Arts, Entertainment, and Recreation, Ac-*

Table 6: Baseline results

VARIABLES	(1) Establishments	(2) Establishments	(3) Establishments	(4) Log productivity
Tall buildings	6.265		194.0**	0.235*
	(8.231)		(86.22)	(0.141)
Bedrock x Steel price (lag 3)		-185.3***		
		(64.92)		
Observations	185,432	179,214	179,214	155,944
R-squared	0.992	0.992		
Year FE	YES	YES	YES	YES
ZIP FE	YES	YES	YES	YES
Estimation	OLS	OLS	IV	IV
Cluster s.e.	ZIP	ZIP	ZIP	ZIP
F	.	.	13.82	13.73

Tall buildings: New stock of tall buildings (not uniquely residential) in the ZIP code. *Bedrock*: distance from earth surface to bedrock in 10 meters. *Steel price (lag 3)*: third lag of global steel price, normalized by the maximum value of steel price between 1997 and 2012. *Establishments*: number of total establishments (in any NAICS sector) in the ZIP code. *Log productivity*: log average (between any NAICS sector) annual payroll per employee in the ZIP code. *Year FE*: year fixed effects. *ZIP FE*: ZIP code fixed effects. *Cluster s.e.*: level of clustering of standard errors. *F*: Kleibergen-Paap rk Wald F-statistics on the excluded instruments. *** p<0.01, ** p<0.05, * p<0.1

commodation and Food Services, and *Other Services*

Table 8 shows the sectors for which it has no found an impact of tall buildings on agglomeration and those in which the number of establishment decreases. In particular, the coefficient of tall buildings on agglomeration is not statistically significant for the following sectors: *Utilities*, *Retail trade*, *Transportation and Warehousing*, *Information*, *Finance and Insurance*, *Management of Companies and Enterprises*. On the other hand, there are some sectors which decide to move away their establishments from ZIP codes in which tall buildings are built. Those sectors are: *Manufacturing*, *Wholesale trade*, and *Administrative and Support and Waste Management and Remediation Services*.

Table 7: Effect of tall buildings on number of establishments, by sector

VARIABLES	(1) All	(2) Constr.	(3) Real est.	(4) Profes.	(5) Educat.	(6) Health	(7) Entert.	(8) Accomod.	(9) Other
Tall buildings	194.0**	80.40***	26.80***	71.05***	18.22***	56.60***	39.54***	76.40***	28.01***
	(86.22)	(22.45)	(8.160)	(22.63)	(4.845)	(17.63)	(11.90)	(21.83)	(10.23)
Observations	179,214	179,214	179,214	179,214	179,214	179,214	179,214	179,214	179,214
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
ZIP FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Estimation	IV	IV	IV	IV	IV	IV	IV	IV	IV
Cluster s.e.	ZIP	ZIP	ZIP	ZIP	ZIP	ZIP	ZIP	ZIP	ZIP

For table notes see 8

Table 8: Effect of tall buildings on number of establishments, by sector

VARIABLES	(1) Utilit.	(2) Retail	(3) Transp.	(4) Informat.	(5) Finance	(6) Managem.	(7) Manuf.	(8) Whole.	(9) Adminis.
Tall buildings	0.815	9.797	-1.563	6.899	-34.17	-6.229	-76.98***	-44.92***	-10.27*
	(0.575)	(12.53)	(3.611)	(4.602)	(22.60)	(3.794)	(18.53)	(12.92)	(5.687)
Observations	179,214	179,214	179,214	179,214	179,214	179,214	179,214	179,214	179,214
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
ZIP FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Estimation	IV	IV	IV	IV	IV	IV	IV	IV	IV
Cluster s.e.	ZIP	ZIP	ZIP	ZIP	ZIP	ZIP	ZIP	ZIP	ZIP

Dependent variable: number of total establishments in a NAICS sector in the ZIP code. *Tall buildings*: New stock of tall buildings (not uniquely residential) in the ZIP code. Instrument: distance from earth surface to bedrock in 10 meters interacted with third lag of global steel price, normalized by the maximum value of steel price between 1997 and 2012. NAICS sectors. *All*: All NAICS sectors, not separated. *Utilit.*: Utilities.*Constr.*: Construction. *Manuf.*: Manufacturing. *Whole.*: Wholesale Trade. *Retail*: Retail Trade. *Transp.*: Transportation and Warehousing. *Informat.*: Information. *Finance*: Finance and Insurance. *Real est.*: Real Estate and Rental and Leasing. *Profes.*: Professional, Scientific, and Technical Services. *Managem.*: Management of Companies and Enterprises. *Adminis.*: Administrative and Support and Waste Management and Remediation Services.*Educat.*: Educational Services. *Health*: Health Care and Social Assistance. *Entert.*: Arts, Entertainment, and Recreation. *Accomod.*: Accommodation and Food Services. *Other*: Other Services (except Public Administration). *Year FE*: year fixed effects. *ZIP FE*: ZIP code fixed effects. *Cluster s.e.*: level of clustering of standard errors. *F*: Kleibergen-Paap rk Wald F-statistics on the excluded instruments. *** p<0.01, ** p<0.05, * p<0.1

3.2 Spatial analysis

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. The positive effect of tall buildings on local agglomeration can be potentially explained mechanically by firms filling the new office spaces. In order to prove that tall buildings have an agglomeration effect, which do not pass from this filling effect, in this section I show that tall buildings have an effect not only in the ZIP in which they are built but also on neighbouring areas and the overall city.

By introducing in equation 1 the completion of skyscrapers in ZIP codes at several km radius distances from the ZIP code in consideration 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, and 25 to 50. 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 economies of scale and possible congestion effects in the area around where tall buildings are constructed.

Defining j as the distance from a zip code, we can estimate the static model to estimate the effect of the construction of skyscrapers in neighbouring zip codes as it is shown in Equation 3. For every radius considered, this model can be estimated using as instrumental variable the lagged international steel price and the average distance to bedrock in the corresponding radius.

$$y_{zjt} = \tau + \tau_z + \tau_t + \beta_j D_{z+i,z+j,t} + \varepsilon_{zjt} \quad \text{for i=\{0, 5, 10, 25\}, and j=\{5, 10, 25, 50\}} \tag{3}$$

In Table 9 I have presented the results of the estimation of the effect of completing a tall building at several radius of distances: between 0 to 5 km, 5 to 10, 10 to 25, and 25 to 50. I have reported the results for all NAICS and for the manufacturing sector, since in this sector tall buildings has been found to have a negative agglomeration effect. The overall agglomeration effect is present even outside the ZIP code in which the tall building is built. The spillover effects are considerably lower than the effect in the ZIP code where it is build but they are statistically significant for all the radii considered. Relative to the effect in the central ZIP code, the spillover effects are between 1 and 3% of the effect in the ZIP code of the construction of the tall buildings. Therefore, we can derive two results. First, the estimation in Table 9 implies that the construction of tall buildings is generating an increase in firms' agglomeration in all the city. Second, the magnitude of this agglomeration effect dissipates quickly with space.

Table 9: Spatial effect of tall buildings on number of establishments

VARIABLES	(1) All	(2) Manuf.	(3) All	(4) Manuf.	(5) All	(6) Manuf.	(7) All	(8) Manuf.	(9) All	(10) Manuf.
Same ZIP	194.0**	-76.98***								
	(86.22)	(18.53)								
0-5 km			2.576***	-0.876***						
			(0.824)	(0.125)						
5-10 km					4.715***	-1.370***				
					(1.205)	(0.179)				
10-25 km							2.927***	-0.705***		
							(0.544)	(0.0798)		
25-50 km									4.157***	-0.655***
									(0.631)	(0.0874)
Observations	179,214	179,214	172,019	172,019	166,183	166,183	181,469	181,469	182,266	182,266
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
ZIP FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Estimation	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV
Cluster s.e.	ZIP	ZIP	ZIP	ZIP	ZIP	ZIP	ZIP	ZIP	ZIP	ZIP
F	13.82	13.82	178.85	178.85	138.52	138.52	312.79	312.79	337.89	337.89

Dependent variable: number of total establishments in a NAICS sector in the ZIP code. *Same ZIP*: New stock of tall buildings (not uniquely residential) in the same ZIP code. *0-5 km*: tall buildings in the ZIP codes in a radius of 0 to 5 km from construction. *5-10 km*: tall buildings in the ZIP codes in a radius of 5 to 10 km from construction. *10-25 km*: tall buildings in the ZIP codes in a radius of 10 to 25 km from construction. *25-50 km*: tall buildings in the ZIP codes in a radius of 25 to 50 km from construction. Instrument: distance from earth surface to bedrock in 10 meters in the radius considered interacted with third lag of global steel price, normalized by the maximum value of steel price between 1997 and 2012. *All*: All NAICS sectors, not separated. *Manuf.*: Manufacturing. *Year FE*: year fixed effects. *ZIP FE*: ZIP code fixed effects. *Cluster s.e.*: level of clustering of standard errors. *F*: Kleibergen-Paap rk Wald F-statistics on the excluded instruments. *** p<0.01, ** p<0.05, * p<0.1

In the case of manufacturing, the presence of congestion effects and diseconomies of scale cannot be rejected. Table 9 reports for manufacturing a negative and significant effect of skyscrapers in the closest areas to where tall buildings are completed. However, this diseconomies seems are small with respect to the effect present in same ZIP code where skyscrapers have been completed.

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 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.

We can further understand the effect of tall buildings on the economic activity of one city looking on how tall buildings change the overall location of establishments in one city. First, I investigate whether tall buildings influence the concentration of establishments in one particular location of the city. Let's define x_{zt} as the proportion of MSA m establishments that are in ZIP code z, and s_{jzt} as proportion of ZIP code z establishments that are in sector j. To understand whether establishments are concentrated in one ZIP code in the city we can use the following concentration index: $G_{jmt} = \sum_{z \in m} \left(x_{zt} - s_{jzt}\right)^2 = \sum_{z \in m} \left(\frac{\sum_j est_{jzt}}{\sum_{z \in m} \sum_j est_{jzt}} - \frac{est_{jzt}}{\sum_j est_{jzt}} \right)^2$. If an industry is allocated across space in same way as total establishments, $x_{zt} = s_{jzt}$, we have no concentration and then $G_{jmt} = 0$. From Table 10, Column (1), we can see that the completion of tall buildings in one city does not increase the concentration of establishments in one ZIP code. This is consistent with the fact that we found a heterogeneous impact of tall buildings on local establishments location depending on the sectors. That is, while some sectors increase agglomeration in the ZIP code where tall buildings are built, some others (like manufacturing) actually decrease their presence.

Table 10: Effect of tall buildings on concentration and specialization economic activity

	(1)	(2)	(3)
	Concentration	Specialization	Specialization
VARIABLES	MSA	ZIP	MSA
Tall buildings, MSA	0.0211		0.000321***
	(0.0142)		(8.14e-05)
Tall buildings, ZIP		0.0187	
		(0.0634)	
Observations	32,385	179,214	1,905
Level Obs.	MSA-NAICS	ZIP	MSA
Year FE	YES	YES	YES
ZIP FE	NO	YES	NO
MSA FE	YES	NO	YES
NAICS FE	YES	NO	NO
Estimation	IV	IV	IV
s.e.	Robust	cluster ZIP	Robust
F	332.77	13.82	18.03

Tall buildings, ZIP: New stock of tall buildings (not uniquely residential) in the same ZIP code. *Tall buildings, MSA*: New stock of tall buildings (not uniquely residential) in one MSA. Instrument: distance from earth surface to bedrock in km at ZIP or MSA level interacted with third lag of global steel price, normalized by the maximum value of steel price between 1997 and 2012. *Level Obs.*: geographical or sectoral level at which observations vary. *Year FE*: year fixed effects. *ZIP FE*: ZIP code fixed effects. *MSA FE*: MSA fixed effects. *NAICS FE*: NAICS sector fixed effects. *s.e.*: standard errors. *F*: Kleibergen-Paap rk Wald F-statistics on the excluded instruments. *** p<0.01, ** p<0.05, * p<0.1

Second, I have checked whether one ZIP code become more specialized in some NAICS sector after the completion of a tall building. I have measured the specialization at ZIP code level using the following Herfindahl index: $H_{zt} = \sum_j \left(s_{jzt}\right)^2 = \sum_j \left(\frac{est_{jzt}}{\sum_j est_{jzt}} \right)^2$. From Table 10, Column (2), we can conclude that tall buildings do not increase the specialization of a ZIP code, leading to agglomeration of different sectors.

Third, I have tested whether MSA specialization in some sector increase. In fact, tall buildings can also attract firms from other MSAs. Similarly, the negative effect of tall buildings on some sectors can lead these establishments to reallocate to some other city. I have measured the specialization at MSA level using the following Herfindahl index: $H_{mt} = \sum_j (s_{jmt})^2 = \sum_j \left(\frac{\sum_{z \in m} est_{jzt}}{\sum_j \sum_{z \in m} est_{jzt}} \right)^2$. From 10, Column (3), we can see that the construction of tall buildings in one city creates an increase in the overall sectoral specialization of that city. Therefore, we cannot exclude that tall buildings can attract new establishments to a city and/or leading some establishments to leave that city.

3.3 City level analysis

I present the results of the estimations of equation 1 taking the average at MSA level of each variable. I also augment the model including a number of controls that proxy the different agglomeration determinants: input sharing, labour market pooling, knowledge spillovers and natural advantages. This estimation has two goals. First, this will allow me to partially understand the importance of the determinants of agglomeration and the mechanism of the effect of tall buildings on agglomeration. Second, in the previous Section I have shown that tall buildings have an effect of overall location of economic activity in one city. Estimating the agglomeration model at MSA level is possible to to obtain the effect of tall buildings that do not depend on spillover effects. Moreover, since some sector benefit and some are disadvantaged by the construction of tall buildings, I can understand if the overall number of establishments in one city increases or decreases after an increase in vertical density in one ZIP code.

Results are presented in Table 11. Column (1) shows the first stage effect of our instrument on the construction of tall buildings using MSAs (and not ZIPs) as unit of observation. The interaction of past international steel price and bedrock depth in one MSA has a negative and statistically significant effect on the total construction of tall buildings in one city. That is, relevance of our instrument is met also at higher level of geographical aggregation. Exploiting this instrument, I estimate a positive and significant effect of tall buildings on the total number of establishments in one city. Overall the construction of tall buildings increases economic activity in one city and then the net migration of establishments to that city is positive. Building a new tall building increases the total number of establishments in one city by 609 establishments, that is, 2% of the average number of establishments in one city[10].

Controlling for the classical determinants of agglomeration, I can understand what are the mechanisms behind the agglomeration effect of vertical density. In columns (3) and (4) of Table 11 I control for proxies of natural advantages (MSA fixed effects), input sharing, labour pooling and knowledge spillover (overall level of education, proportion of high skill workers and number of patents in the city). The coefficient of input sharing is the only significant coefficient, and it shows that higher level of input sharing increases agglomeration in one city. Part of the mechanism of the effect of tall buildings on agglomeration is via its effect on input sharing and patents. In fact, from

[10]The average number of establishments in one MSA is 30869

Table 11: Effect of tall buildings on agglomeration and productivity at MSA level

VARIABLES	(1) Est	(2) Est	(3) Est	(4) Est	(5) Educ.	(6) High sk.	(7) Patents	(8) Input sh.	(9) Log prod.	(10) Log prod.
Tall buildings, MSA		608.8***	541.1***	541.1**	0.0549	-0.0319	60.22***	6.421***	0.00491***	0.00412**
		(103.5)	(107.3)	(257.6)	(0.0408)	(0.0445)	(16.11)	(2.197)	(0.00160)	(0.00168)
Education, MSA			-1.444	-1.444						9.35e-05
			(19.09)	(26.00)						(0.000685)
High skills, MSA			-3.439	-3.439						9.46e-05
			(15.82)	(17.11)						(0.000607)
Patents, MSA			-0.452*	-0.452						1.75e-06
			(0.241)	(0.439)						(6.27e-06)
Input sharing MSA			13.81***	13.81***						6.24e-05***
			(0.987)	(1.944)						(1.94e-05)
Bedrock MSA x Steel pr.	-0.00109***									
	(0.000256)									
Observations	1,905	1,905	1,849	1,849	1,851	1,849	1,905	1,905	1,905	1,849
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
MSA FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Estimation	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV
s.e.	Rob.	Rob.	Rob.	cl. MSA	Rob.	Rob.	Rob.	Rob.	Rob.	Rob.
F	.	18.03	16.51	2.27	17.98	17.99	18.03	18.03	18.03	16.51

All dependent variables measured at MSA level. *Tall buildings, MSA*: New stock of tall buildings (not uniquely residential) in one MSA. *Bedrock MSA x Steel pr.*: distance from earth surface to bedrock in 10 meters at MSA level interacted with third lag of global steel price. *Est.*: number of total establishments (in any NAICS sector) in the MSA. *Log prod.*: log average (between any NAICS sector) annual payroll per employee in the MSA. *Education, MSA* / *Educ.*: proportion of people with more than a bachelor's degree in the MSA. *High skills, MSA* / *High sk.*: proportion of workers working in Management, professional and related occupations. *Patents, MSA* / *Patents*: number of patents in the MSA. *Input sharing, MSA* / Input sh.: Average for all sectors and MSA of input sharing for each sector. 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. *Year FE*: year fixed effects. *MSA FE*: MSA fixed effects. *s.e.*: standard errors. *Rob.*: robust standard errors. *Cl. MSA*: standard errors clustered at MSA level. *F*: Kleibergen-Paap rk Wald F-statistics on the excluded instruments. *** p<0.01, ** p<0.05, * p<0.1

Columns (5) to (8) it is possible to evince that the construction of tall buildings has a positive and significant effect on input sharing between sectors and the completion of patents in one MSA.

The effect of tall buildings on agglomeration remains significant after controlling for the determinants of agglomeration. This result suggests that tall buildings and vertical density has an agglomeration effect irrespective of its effect on input sharing, labour pooling and knowledge spillover. This is possibly due to the fact that skyscraper can be a particularly attractive location for firms because of the associated gains in prestige from being located in a tall landmark building.

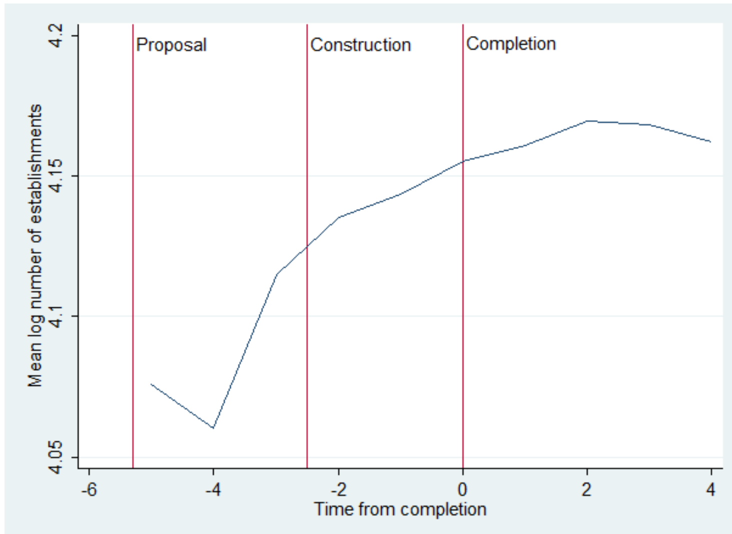
The completion of new tall buildings is increasing agglomeration of firms to that city but it also has a positive and significant effect on overall productivity of firms in that city (column (9) of Table 11). One new tall building in a city leads to an increase of 0.05% in overall MSA productivity. This effect is present even controlling for the potential mechanisms. In fact, controlling for the input sharing, labour pooling and knowledge spillover, one new tall building in one city leads to an increase of 0.04% in overall MSA productivity (see column (10) of Table 11). I have also found that input sharing is the only mechanism with a significant and positive effect on MSA productivity.

4 Dynamic analysis

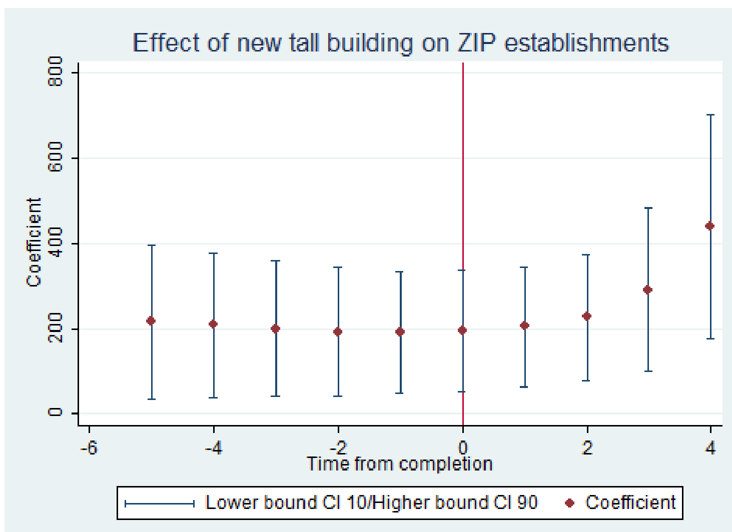
In the previous section it has been possible to observe that the completion of new tall buildings have the effect of increasing firms' agglomeration and to increase the general level of productivity of the area. In order to obtain identification of the effect of the completion of tall buildings I have made used 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 5(a) might suggest that reverse causality can be present. In fact, ZIP codes that complete a tall buildings where experiencing increases in the number of establishment prior to the construction of the structure. However, this dynamic can be driven by anticipation effects instead of reverse causality. 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 agglomeration economies.

I can provide a preliminary test of the presence of anticipation effects by running separate regressions of the current number of establishment on future and past completion of tall buildings. The result of these regressions are reported in Figure 5(b). As it is possible to see, the increase in firms' agglomeration begins even before a tall building is completed. This is evidence that firms are locating in the area where the skyscrapers will be erected even before its actual completion. Thus, the previous estimated agglomeration effects are not just driven by firms filling new tall buildings. Moreover, firms agglomeration continues for years after the completion of the building.

In the next section I introduce a new econometric model that will allow me to



(a) Trends



(b) Estimated coefficients

Figure 5: Dynamics of the effect of tall buildings on agglomeration, IV regressions

Panel a): Mean log number of establishments pre and post the completion of a tall building in the same ZIP code. Panel b): Effect of completion of new tall buildings on number of overall establishments in one ZIP code. Each point represents a different estimation. For each year y before or after the completion of the building, the coefficients have been estimated regressing the number of total establishments (in any NAICS sector) in one ZIP code in current year on new stock of tall buildings (not uniquely residential) in the same ZIP code in year y, instrumented by distance from earth surface to bedrock in 10 meters at ZIP level interacted with third lag of global steel price in year y. Estimated controlling for ZIP and year fixed effects. Standard errors clustered at ZIP level.

provide a clear test for the presence of anticipation effects, and to properly estimate the dynamic agglomeration effects.

4.1 Dynamic empirical strategy

Model 1 can be extended in order to include dynamic effects. One possibility would be to introduce leads and lags of the completion of tall buildings to confirm that there is 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 an augmented version of Model 1 that also includes the completion of future tall buildings:

$$y_{zjt} = \tau + \tau_z + \tau_t + \beta D_{zt} + \beta \sum_{s=1}^{\infty} \theta^s E_t \left[D_{t+s} \right] + \varepsilon_{zjt} \tag{4}$$

In Model 4 the ex-post effect of the completion of a new tall building is given by $\beta\left(1+\theta F+\theta^2 F^2+\theta^3 F^3+\cdots\right) = \beta\left[1-\theta\,(F)\right]^{-1} = \frac{\beta}{1-\theta}$, 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 4 that includes only one term for the treatment variable assuming rational expectations. Let's 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 $E_t\left[D_{t+s}\right] = D_{t+s} + v_{t,t+s}$. Adding and subtracting θy_{zjt+1} from equation 4,it is possible to arrive to this new model:

$$y_{zjt} = \mu + \mu_z + \mu_t + \theta y_{zjt+1} + \beta D_{zt} + u_{zjt} \tag{5}$$

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 ($\beta \sum_{s=2}^{\infty} \left(v_{t,t+s} - v_{t+1,t+s}\right)$)[11].

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 effects. 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. In order to test the presence of anticipatory effects I will test the null hypothesis that $\beta\theta = 0$.

[11] In fact, $u_{zt} = \varepsilon_t - \theta\varepsilon_{t+1} + \beta v_{t,t+1} + \beta\left[\sum_{s=2}^{\infty} \theta^s \left(v_{t,t+s} - v_{t+1,t+s}\right)\right]$

Equation 5 can be estimated in first difference using as instruments lags of the dependent variable as suggested in Arellano and Bond (1991). Moreover, from equation 4 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. Exogeneity restriction of these instruments is given by the fact that agglomeration is related to future tall buildings only through y_{t+1}.

There are several assumptions imposed to obtain identification, as it is being stressed by Malani and Reif (2015).

- Firstly, the outcome and its change should not be correlated with past disturbances. For this assumption to be met we need that $E\left[y_{z,T}u_{z,t}\right] = 0\,\forall z, \forall t \leq T-1$. Furthermore, in order to have assumption satisfied we also need that $E\left[\Delta y_{z,t+1}u_{z,t}\right] = 0$.

- It is possible to have small order autocorrelation, but there should not be autocorrelation higher than order 1 in ε_t. That is, $E\left[\varepsilon_t\varepsilon_{t+j}\right] = 0\,\forall j > 1$.

- It is also needed that the error in the estimation of agglomeration determinants, ε_t, is orthogonal to the ahead forecast errors. That is, $E\left[\varepsilon_t v_{t+j,t+k}\right] = 0\,\forall k > j, \forall j > 1$.

- Moreover, change in forecast should be uncorrelated with the actual level of the forecast. We can write this assumption as: $E\left[\left(v_{t,t+k} - v_{t+1,t+k}\right)v_{t+j,t+j+1}\right] = 0\,\forall k > 1, \forall j > 1$.

- The final assumption is that independent information must be used to update the forecast. Which means that $E\left[\left(v_{t,t+k} - v_{t+1,t+k}\right)\left(v_{t+j,t+m} - v_{t+j+1,t+m}\right)\right] = 0\,\forall k > 1, \forall m > j+1, \forall j > 1$.

4.2 Dynamic results

Results of my dynamic estimation pooling all sectors together are presented in Table 12. Column (2) shows the dynamic result using the number of establishments as dependent variable. The effect of tall buildings on agglomeration is positive and significant. The ex-post effect, that is the overall effect of tall buildings on agglomeration after the building is completed, is higher than the effect found in the static case (Column (1)). In fact, adding a new tall building has the effect of increasing the number of establishments in the same ZIP code for all the years after its completion by 376 units or 78% (see column (3)). While in the static case we estimated this effect to be 194 units.

Assuming a positive correlation between the current effects of tall buildings on agglomeration and future expectations, this result suggests that previous results were biased by dynamic terms downwards, and that anticipatory effects were lower than the ex-post effects. In fact, we cannot reject the presence of anticipation effects. Moreover, ex-ante effects are present in all the years after the skyscrapers have been proposed (usually 5.3 years before its completion) and they are lower in magnitude than

Table 12: Dynamic effect of tall buildings on agglomeration

VARIABLES	(1) Est.	(2) Est.	(3) Log est.
Tall buildings $\hat{\beta}$	194.0**	4.034**	0.0532***
	(86.22)	(1.999)	(0.00500)
Est. (lead 1) $\hat{\theta}$		0.989***	
		(0.000768)	
Log est. (lead 1) $\hat{\theta}$			0.932***
			(0.00125)
CALCULATED EFFECTS			
Ex post eff. $\frac{\hat{\beta}}{1-\hat{\theta}}$		376.396**	0.781***
s.e.		(185.602)	(0.071)
Ex ante eff. (t-1) $\frac{\hat{\beta}\hat{\theta}}{1-\hat{\theta}}$		372.362**	0.728***
s.e.		(183.625)	(0.066)
Ex ante eff. (t-2) $\frac{\hat{\beta}\hat{\theta}^2}{1-\hat{\theta}}$		368.371**	0.678***
s.e.		(181.669)	(0.061)
Ex ante eff. (t-3) $\frac{\hat{\beta}\hat{\theta}^3}{1-\hat{\theta}}$		364.422**	0.632***
s.e.		(179.733)	(0.057)
Ex ante eff. (t-4) $\frac{\hat{\beta}\hat{\theta}^4}{1-\hat{\theta}}$		360.516**	0.589***
s.e.		(177.820)	(0.053)
Ex ante eff. (t-5) $\frac{\hat{\beta}\hat{\theta}^5}{1-\hat{\theta}}$		356.652**	0.549***
s.e.		(175.927)	(0.049)
TEST ANTICIPATION			
$\hat{\beta}\hat{\theta}$		3.991**	0.050***
s.e.		(1.977)	(0.005)
Observations	179,214	164,692	164,692
Year FE	YES	YES	YES
ZIP FE	YES	FD	FD
Estimation	IV	AB	AB
Cluster s.e.	ZIP	ZIP	ZIP
p-value AR(2) test		0.549	0.138

Tall buildings: New stock of tall buildings (not uniquely residential) in the ZIP code. *(Log) est.*: (log) number of total establishments (in any NAICS sector) in the ZIP code. Log est.=log(est+1). *Year FE*: year fixed effects. *ZIP FE*: ZIP code fixed effects. *FD*: model estimated in first difference. *Cluster s.e.*: level of clustering of the standard errors. *AB*: Arellano Bond first difference estimation. Instruments used: distance from earth surface to bedrock in 10 meters at ZIP level interacted with third lag of global steel price, leads of (log) establishments and tall buildings (from 2nd to 5th), lags of (log) establishments (from 1st to 5th). *p-value AR(2) test*: p-value of test where null hypothesis is autocorrelation of order 2 in the errors. *** p<0.01, ** p<0.05, * p<0.1

the ex-post effects. Estimation of the discount term, θ, suggests that time-persistence is important in this estimation. I obtain a discount term between 0.932 and 0.989. However, an AR(2) test on the errors in first difference suggests that this persistence is not biasing our dynamic results.

5 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. 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 paper 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 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. The quantification of agglomeration economies is 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.

5.1 Empirical dynamic strategy

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 on local density. In this paper I use the ratio between annual payroll and total employment as proxy for productivity in one ZIP code. I use log firms' density as measure of density, referring to the log number of establishments divided by the ZIP area.

In addition to quantifying the elasticity of productivity to firms' density, I estimate the elasticity given by 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. Therefore, we can interpret this last term as the additional agglomeration economies given by vertical density. This interaction is instrumented using the interaction between bedrock distance and the third lag of Global steel price.

I exploit the exponential discounting model presented in Section 4.1 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:

$$log(P_{zt}) = \lambda + \lambda_z + \lambda_t + \theta log(P_{zt+1}) + \beta_1 log\left(\frac{est_{zt}}{size_z}\right) + \beta_2 log\left(\frac{est_{zt}}{size_z} D_{zt}\right) + \varepsilon_{zjt}$$

Equation 6 has been estimated in first difference with Arellano-Bond technique 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 from biases given by sorting. That is, the completion of new tall buildings might attract more productive firm. ZIP fixed effects will also partially control for time-invariant 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.

5.2 Agglomeration regressions

Results of the estimation of agglomeration economies considering a dynamic model are presented in Table 13. Column (1) presents the OLS estimation of the elasticity of productivity to firms' density. This estimate is 0.0262, an estimate similar to estimates of productivity with respect to employment density, which tend to be around 0.02 and 0.04 (see Combes and Gobillon, 2015 and Melo et al., 2009). Column (2) further controls for the effect of density passing through the construction of tall buildings. OLS estimate suggests that without any increase in vertical density, the elasticity of productivity to horizontal density is 0.0258. Then, if the increase comes from vertical density of firms' establishment the OLS estimate suggests an additional 0.003% increase in productivity for an increase in vertical density by 1%. Instrumenting the interaction between firms' establishment density and the completion of tall buildings the magnitude of this variable increase but become insignificant (see Column (3)).

Column (4) controls for the presence of anticipatory effects using the exponential discount model. The estimated static elasticities of horizontal density and its interaction with tall buildings are 0.058 and 0.004 respectively. The ex-post elasticity of log establishment density that is not passing through tall buildings is found to be 0.13. The presence of anticipatory effects cannot be rejected. However, a test in the autocorrelation of the error term suggests that our model is still biased. This is given by the fact that one identification assumption of the exponential discounting model is the absence of autocorrelation of order higher than 2 in the errors.

In order to solve the problem of time persistence, column (5) controls for additional leads of productivity. In this new model we can reject the presence of autocorrelation of order 2 in the errors. The estimated static elasticities are lower but still positive and statistically significant. The elasticity of productivity with respect to firms'

Table 13: Dynamic effect of tall buildings on agglomeration economies

VARIABLES	(1) Log prod.	(2) Log prod.	(3) Log prod.	(4) Log prod.	(5) Log prod.
Log est. density	0.0262***	0.0258***	0.0260**	0.0578***	0.0303***
	(0.00952)	(0.00951)	(0.0106)	(0.00267)	(0.00227)
Tall buildings x log est. density		0.00301***	0.0304	0.00367***	0.00652***
		(0.000754)	(0.0186)	(0.000920)	(0.00113)
Log. prod (lead 1)				0.544***	0.441***
				(0.00726)	(0.0155)
Log. prod (lead 2)					0.111***
					(0.0108)
Log. prod (lead 3)					0.0829***
					(0.00922)
Log. prod (lead 4)					0.0171* *
					(0.00884))
CALCULATED EX-POST EFFECT $\frac{\hat{\beta}}{1-\hat{\theta}}$					
Log est. density				0.127***	0.054***
s.e.				(0.005)	(0.004)
Tall buildings x log est. density				0.008***	0.012***
s.e.				(0.002)	(0.002)
TEST ANTICIPATION $\hat{\beta}\hat{\theta}$					
Log est. density				0.031***	0.013***
s.e.				(0.001)	(0.001)
Tall buildings x log est. density				0.002***	0.003***
s.e.				(0.0005)	(0.0005)
Observations	159,888	159,888	155,944	140,169	99,192
Year FE	YES	YES	YES	YES	YES
ZIP FE	YES	YES	YES	FD	FD
Estimation	OLS	OLS	IV	AB	AB
Cluster s.e.	ZIP	ZIP	ZIP	ZIP	ZIP
p-value AR(2) test				0.000	0.271

Log est.density: log number of total establishments (in any NAICS sector) in the ZIP code per squared km of the ZIP code. *Tall buildings*: New stock of tall buildings (not uniquely residential) in the ZIP code. *Log productivity*: log average (between any NAICS sector) annual payroll per employee in the ZIP code. *Year FE*: year fixed effects. *ZIP FE*: ZIP code fixed effects. *FD*: model estimated in first difference. *Cluster s.e.*: level of clustering of the standard errors. *AB*: Arellano Bond first difference estimation. Instruments used: distance from earth surface to bedrock in 10 meters at ZIP level interacted with third lag of global steel price, leads of log establishments and tall buildings (from 2nd to 5th), lags of log establishments (from 1st to 5th). *p-value AR(2) test* : p-value of test where null hypothesis is autocorrelation of order 2 in the errors. *** p<0.01, ** p<0.05, * p<0.1

establishment density is 0.03. In addition, increasing density via the construction of tall buildings increases the productivity elasticity of 0.007. I compute the ex-post elasticities to quantify the overall effect of density on productivity. Overall, increasing firms' establishment density by 1%, without a tall building, leads to the increase of productivity by 0.054%, a result which is in upper bound of the estimations found in the literature. The construction of a new tall building further rise productivity by 0.01%. Since the elasticity of productivity to firms density is around 0.05 and the elasticity considering tall buildings is 0.01, it is possible to say that new tall buildings add a 20 percent increase in productivity on top of the increase given by firms density. Importantly, this additional productivity increase is statistically significant.

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 accrued to tall buildings is present even before the completion of these structures. This is given by the existence of persistence in these estimations, which can be seen by an estimated discount term of 0.44.

6 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. Even 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. 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.

I found that the construction of tall buildings have a positive effect on the attraction of new firms in the ZIP codes in which they are built and on their productivity. This effect on agglomeration differs between sectors. Sectors which are more related to the production of goods or the use of land, such as manufacturing or wholesale trade, are actually pushed away by the construction of new tall buildings.

Tall buildings and the increase in vertical density of firms' location have also effects on the overall city distribution of economic activity. The completion of tall buildings increases agglomeration also in the ZIP codes closed to the where the structure is built. Tall buildings might function as a coordination device to increase concentration of firms in different parts of the city. Because of the heterogeneous effects of vertical density with respect to tall buildings, I find that a city becomes more specialized after an increase in the number of tall buildings.

Exploiting variation at MSA level, cities with higher vertical density are able to attract new firms. In terms of mechanisms, I found that the increase in vertical density at city level has a positive and significant effect on input sharing between sectors and the completion of patents in one MSA. The effect of tall buildings on agglomeration and productivity remains significant even after controlling for the determinants of agglomeration. That is, tall buildings might provide an agglomeration mechanism different from input sharing, labour pooling and knowledge spillovers.

It has been possible to confirm that the agglomeration effect is not only driven by firms filling tall buildings. I can conclude that there existsa nd attraction of firms to all ZIP codes close to skyscrapers. Furthermore, using an exponential discount model I encounter an anticipatory agglomeration effect of firms, which happen before the actual completion of the building.

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 0.01 percent. I also estimate an elasticity of productivity to firms' establishment density of 0.05. As a result, 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 in 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 alone. In particular I cannot distinguish whether agglomeration is driven by anticipation of present and future productivity increases or prestige of landmark buildings. Moreover, the increase in productivity following the 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

Ahlfeldt, G. M. and D. McMillen (2017). Tall buildings and land values: height and construction cost elasticities in chicago, 1870–2010. *Review of Economics and Statistics* (0).

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.

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.

Bertaud, A. and J. K. Brueckner (2005). Analyzing building-height restrictions: predicted impacts and welfare costs. *Regional Science and Urban Economics 35*(2), 109–125.

Borck, R. (2014). Will Skyscrapers Save the Planet? Building Height Limits and Urban Greenhouse Gas Emissions. CESifo Working Paper Series 4773, CESifo Group Munich.

Brueckner, J. K., S. Fu, Y. Gu, and J. Zhang (2015). Measuring the stringency of land-use regulation and its determinants: The case of china's building-height limits. Technical report, Working paper, May 19 2015.

Brueckner, J. K. and K. S. Sridhar (2012). Measuring welfare gains from relaxation of land-use restrictions: The case of india's building-height limits. *Regional Science and Urban Economics 42*(6), 1061–1067.

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*.

Danton, J. and A. Himbert (2018). Residential vertical rent curves. *Journal of Urban Economics 107*, 89–100.

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.

Economist (2012, October 12). The lore of ore. The Economist.

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

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.

Liu, C. H., S. S. Rosenthal, and W. C. Strange (2018). The vertical city: Rent gradients, spatial structure, and agglomeration economies. *Journal of Urban Economics*.

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 (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.

A Construction of other buildings

In section 2.1 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 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 14 and 15. 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 14: Construction spending determinants by buildings usage

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

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

Vertical and horizontal cities: in which direction should cities grow?

Federico Curci *

Abstract

This research establishes the different consequences of taller or more spread out cities using a strategy that combines reduced form estimation with a more structural approach. In order to achieve this goal I build a general spatial equilibrium model which includes both within-city and between-cities spatial equilibrium concepts. Results from IV regressions suggest that both vertical and horizontal increase of a city, measured as an increase in the floors of buildings and in total lot size occupied by buildings in a city, are associated with a positive increase in house prices, and no statistically significant effect on wage. However, increasing a city vertically would lead to a higher effect on house prices with respect to increasing it horizontally. These reduced form estimates are consistent with a calibration of the model in which building height and city size have a similar positive effect on city-specific productivity, while height has stronger positive effect on city-specific amenities than city size.

1 Introduction

Urbanization is one of the most important social phenomena in the current and last centuries. According to U.N. (2014) 30 percent of the world population lived in urban areas in 1950 and this percentage has increased to 54 percent in 2014. The massive inflow of new inhabitants to cities is shifting attention of urban planners and economists towards discovering efficient ways to distribute people across space. It is well know in the literature that bigger cities tend to have higher productivity level, due to agglomeration economies (see Combes and Gobillon, 2015 for a review). Size and density are concepts that are often used interchangeably in explaining why bigger cities are more productive. However, it still remains unanswered how the distribution of people between inside a city can explain agglomeration effects provided by cities.

*Universidad Carlos III de Madrid. Departamento de Economía. Calle Madrid 126, 28903 Getafe (Spain). Email: fcurci@eco.uc3m.es. I thank Klaus Desmet, Ricardo Mora, and Adelheid Holl for the excellent supervision. I thank Jesus Fernandez-Huertas, Luigi Minale and participants to the UC3M reading group, UC3M PhD workshop, 2016 IEB/UEA Summer School in Urban Economics, City Structure Workshop at Kraks Fond, 2017 ENTER Jamboree, and 7th European Meeting of the UEA. Financial support from the Spanish Ministry of Economics and Competitiveness (ECO2012-33053) is acknowledged. Part of this research was conducted while visiting the Department of Economics at Southern Methodist University, whose hospitality is gratefully acknowledged.

In this paper I analyse how city development can influence the externalities produced by cities in terms of productiviy and amenities. I separately consider the performances of cities that have expanded "vertically", that is in which the height of building is higher, versus cities that have expanded "horizontally", that is in which the spread of the city is bigger. This analysis is necessary in order to shed lights on how cities can work more efficiently, and to understand the factors that explain how people interact between cities.

The first goal of the paper is to derive different elasticities of wages and house prices with respect to the vertical and horizontal expansion of cities. The additional goal of this paper is to understand how tall and spread cities can influence other welfare variables, which are more difficult to measure, such as productivity and amenities. Since agglomeration economies are reflected in both wages and house prices, I build a spatial equilibrium model in order to separate the effects of different city expansions on productivity and amenities. I follow and improve the strategy proposed by Glaeser (2008). This allows me to map reduced form elasticities on wages and house prices to more general considerations about productivity and amenities. In particular, I disentangle productivity and amenity advantages given by an increase in population and understand the difference between size and density for agglomeration economies.

I develop a spatial equilibrium model with different cities and places inside a city. Utility should be equalized between people living in different areas. As in a Rosen-Roback model utility should be equalized between cities. Moreover, as in a Alonso-Mill-Muth model I also assume that cities are monocentric: people work in the center of the city and they locate at a particular distance from the Central Business District. The economy consists of three agents: consumers, firms in production sector, and firms in construction sector. This model produces reduced form equations of the effect of building height and horizontal spread of city to wages and house prices. The estimated coefficients depend on model parameters and this allows to infer conclusions about the effect of height and city size on productivity and amenities which are consistent with my model. In this way I will be able to understand the different externality effects of vertical and horizontal city development.

One of the most important contribution of this research is to provide causal evidence of a the effect of a vertical and horizontal city growth. Credible estimates of the effect of building height and horizontal spread of cities are obtained using instrumental variable techniques. In particular, I will instrument building height using earthquake risk and a proxy of the elevator technology available in each city at a certain point of time. Moreover, exogeneous variation in horizontal spread of cities will be obtained using geographical constraints to city development, such as presence of water and steepness, and a Bartik population shifter.

I focus my attention on the U.S. because of the existence of considerable heterogeneity in city development which allows me to compare cities that are taller with cities that are more spread in space. I assembled a database of observations at housing level for more than 460.000 houses in 55 Metropolitan Statistical Areas (MSA) from 1998 to 2013, with information about wages and house prices of people living in each house, number of floors of each building unit and horizontal spread of each city.

Results suggest that houses prices are positively related to both the increase in building height and a bigger horizontal spread of cities. However, the estimated elasticity of house prices with respect to building height is higher than in case of bigger spread of cities. The effect on wages of both building height and horizontal spread of cities is not significant. Theoretical predictions of my model are that an increase in productivity should increase both house prices and wages, while an increase in amenities should increase house prices and reduce wages. As a result I can combine the reduced form elasticities of city expansion to wages and house prices with the theoretical predictions from the model. I conclude that a vertical expansion of the city is likely to influences positively productivity and amenities, while a horizontal expansion have a positive effect only on productivity. These results are robust to several specification but they prove to be heterogeneous to several dimensions.

Once controlling for their effect on population, the elasticities of productivity with respect to both a vertical and a horizontal expansion of the city are similar. It has been estimated that the increase in productivity caused by increasing population adding an additional floor to every building in one city is equivalent to the increase caused by increasing city radius by 5.86 km. Nevertheless, my results show that the amenities respond differently to either a vertical or horizontal city development. In fact, the elasticity of amenities with respect to horizontal spread is 87% lower than what would happen if additional floors would have been added to the building of the city.

The importance of tall buildings, city shape and land use regulation for productivity, land prices, greenhouse limitations has been the topic of study of several recent studies. Koster et al. (2014) is one of the first study to analyze the agglomeration effect of tall buildings. This paper has 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. Ahlfeldt and McMillen (2017) considers that vertical development of a city is the result of high land prices, that is, developers respond to increasing land prices by increasing density through building taller. Liu et al. (2018) assess that tall buildings has important effect for the city spatial structure by influencing the rent gradient. Bertaud and Brueckner (2005) develop theoretical predictions of building height regulations, while Brueckner and Sridhar (2012), and Brueckner et al. (2015) estimate the welfare gains in terms of commuting costs in the Indian and Chinese cases, respectively. Borck (2014) provides a theoretical framework to see the impact of skyscrapers in reducing pollution. Finally, Duranton and Turner (2018) find small effects of urban shape to individual driving behaviours.

My work also relate to the literature which tries to estimate why productivity and amenities are higher in urban areas. Agglomeration economies, the positive correlation between city density or size and productivity, has been estimated since the work of Ciccone and Hall (1996). In this work the authors instrument density by historical population. In order to control for sorting of more productive workers in cities Combes et al. (2010) estimate this elasticity using worker fixed effects and an instrumental variable strategy. Estimated elasticities of productivity with respect to density are around 0.02 and 0.04 (see Combes and Gobillon, 2015 and Melo et al., 2009). In my work I contribute to the literature by separating the effect of size and density and understand if their elasticities differ.

Bigger urban centers are associated with higher levels of productivity but they also can generate better amenities. Because of the unobservability of amenities several works have tried to estimate the effect of any variable of interest on amenities using more structural approaches. Ahlfeldt et al. (2015) use the shock of the fall of the Berlin wall to estimate how productivity and amenities reacted to an increase of population. Diamond (2016) also uses a structural approach to estimate the college wage gap in presence of endogeneous level of amenities. Harari (2015) is the work which is more similar to mine. In fact, she estimates the welfare consequences of having a city with better shape, more compactness, in the Indian case using the strategy proposed by Glaeser (2008). However, in my work I propose a strategy to disentangle the vertical and the horizontal dimensions of cities, irrespectively of their land shape. Curci (2018) estimates that vertical density produces additional agglomeration economies, in terms of increased productivity, from horizontal density.

An additional contribution of my work is to improve on the strategy proposed by Glaeser (2008). In particular, I do not consider city horizontal size taken as given but I endogeneize this variable by allowing people to decide where to locate. Moreover, Glaeser (2008) derives the structural equations of house prices, wages, and population and then he obtains reduced form equations by assuming that amenities, productivity and city size depends on the exogeneous variable under consideration. In my work the variables I am interested to study are already present in my model and my reduced forms are internally consistent with my model. Finally, since height and city size have been specifically modeled I can also obtain first stage equations from my model which I will use to obtain exogeneous variation in my empirical part.

This paper is structured as follows. Section 2 presents the model that I will use to understand how building height and horizontal city spread influence city-specific outcomes and to transform reduced forms estimates to theoretical considerations about productivity and amenities. The data used and the identification strategy of my empirical estimations are described in Section 3. In Section 4 I show and comment the empirical results obtained. Finally, Section 5 concludes.

2 Rosen-Roback monocentric city model

The spatial equilibrium model under consideration consists of three agents: workers, production firms, and construction firms. No heterogeneity is present, each agent is identical. In this model each city is a competitive economy and free mobility of workers equalizes utility levels across cities, as it has been introduced by Rosen (1979) and Roback (1982). Cities will differ for the specific level of amenities and productivity they can offer.

Moreover, cities are assumed to be monocentric, all the employment is located in the Central Business District (CBD), the center of the circle, and individuals need to decide where to live in the city. Utility of individuals should be equalized at each distance from the CBD, consistent with the monocentric model proposed by Alonso (1964), Muth (1969), and Mills (1972).

2.1 Agent problems

2.1.1 Workers problem

Let's define C as consumption, H as housing, d as the distance from the CBD, and θ as city-specific amenity value. Moreover, w is the wage value and p_H is the price of housing. Let's also assume Cobb-Douglas utility function and inelastic supply of work by the workers. I assume that transport costs, $t(d)$, are a linear function of the distance from the CBD. Moreover, transport cost represents a portion τ of the wage income, and we can think of transport costs as opportunity cost of working. Then, the problem of the worker is to choose consumption, housing and distance from the CBD as follows:

$$\begin{aligned} \max_{C,H,d} \quad & \theta C^{1-\alpha} H^{\alpha} \\ \text{s.t.} \quad & C = w - t(d) - p_H(d)\, H \\ & t(d) = w\tau d \end{aligned}$$

Worker's problem solution gives raise to the following differential equation: $\frac{p'_H(d)}{p_H(d)} = -\frac{\tau}{\alpha(1-\tau d)}$, which can be solved as $p_H(d) = p_H(0)(1-\tau d)^{\frac{1}{\alpha}}$. Let's define $\underline{\text{U}}$ as the indirect utility, that should be equal across locations. Hence, its value at the center of the city, d=0, is $\underline{\text{U}} = \theta(1-\alpha)^{1-\alpha}\alpha^{\alpha} w p_H(0)^{-\alpha}$. Consequently, house prices at the center of city should be $p_H(0) = \left(\frac{\theta}{\underline{\text{U}}}\right)^{\frac{1}{\alpha}}(1-\alpha)^{\frac{1-\alpha}{\alpha}}\alpha w^{\frac{1}{\alpha}}$.

Combining the differential equation solution and the equation for the house prices at the city center I can obtain the spatial indifference condition, since utility should be equalized between cities and at each different location inside a city. The between-cities and within-city spatial indifference condition, equation 1, says that lower value of wages in a city should be compensated either by low house prices in that city or high value of city amenities. At the same time, high transport costs due to living further away from the city center should be compensated by lower house prices. This condition and the assumption of Cobb-Douglas utility determine a negative house prices gradient with respect to distance from the CBD and a positive gradient for housing consumption, as it is shown in Figure 1.

$$log\left(\underline{\text{U}}\right) = log(w) - \alpha log\left[p_H(d)\right] + log(\theta) + log(1-\tau d) + \alpha log(\alpha) + (1-\alpha)\, log(1-\alpha) \tag{1}$$

2.1.2 Firms in the production sector problem

Firms in the production sector produce the consumption good using labour (N), and two kind of capitals: traded capital (K) and non-traded capital (Z), which is supplied in a fixed quantity $\bar{Z}$ in each location. This assumption allows to have firms facing constant returns to scale but to have decreasing returns to scale at city level, and then the presence of a finite number of firms in each city (see Glaeser and Gottlieb, 2009).

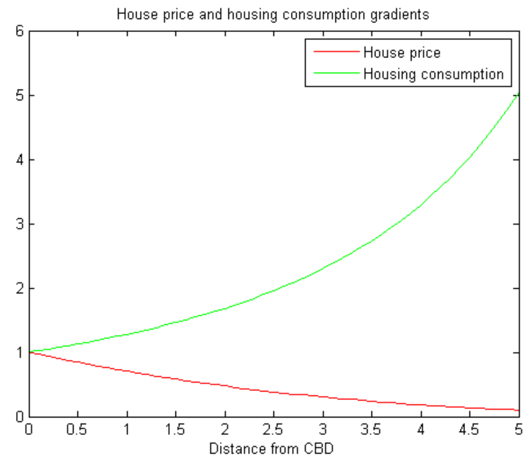


Figure 1: House price and housing consumption gradients

Obtained assuming $\tau = 0.1$, $p_H(0) = H(0) = 1$, $\alpha = 0.3$

Price equation: $p_H(d) = p_H(0)(1-\tau d)^{\frac{1}{\alpha}}$

Housing consumption equation: $H(d) = \frac{\alpha w^{-\frac{1-\alpha}{\alpha}} \underline{\mathrm{U}}^{\frac{1}{\alpha}}}{\theta^{\frac{1}{\alpha}} (1-\alpha)^{\frac{1-\alpha}{\alpha}} \alpha} (1-\tau d)^{-\frac{1-\alpha}{\alpha}} = H(0)(1-\tau d)^{-\frac{1-\alpha}{\alpha}}$

Traded capital is assumed to be priced 1. A is the parameter that reflects city-specific productivity. From the firm's problem that follows it is possible to obtain the labour demand, reported in Equation 2.

$$\max_{N,K} \quad AN^{\beta} K^{\gamma} \bar{Z}^{1-\beta-\gamma} - wN - K$$

$$log(w) = \frac{1}{1-\gamma} log(A) + log(\beta) + \frac{1-\beta-\gamma}{1-\gamma} \left[log(\bar{Z}) - log(N) \right] + \frac{\gamma}{1-\gamma} log(\gamma) \qquad (2)$$

2.1.3 Firms in the construction sector problem

The last agents in this model are the firms in the construction sector. They use a fixed quantity of land (L) which is available at every distance at a cost p_L, to build the housing units, which are assumed to be the product between height (h) and land area. In addition to the land cost, the cost of building, $C(H)$, depends on the height of the housing unit to build. The cost of building high is assumed to be convex: adding one more floor to a house leads to a more than proportional increase in construction costs, that is $\delta > 1$. This assumption rules out the possibility of a perfectly vertical city, where all the population lives in the CBD. While δ refers to the current technology to build higher, which is common across cities, c_0 refers to a city-specific factor that influence the cost of height. Finally, land availability at each distance depends positively on the distance from city center, and negatively on any other external factor that might prevent construction at a particular location. ϕ represents the part of land which is undevelopable at every location, and this will later model exogeneous variation of city spread.

$$
\begin{aligned}
\max_{H} \quad & p_H(d)\, H - C(H) \\
\text{s.t.} \quad & H = h(d) \times L(d) \\
& C(H) = c_0 h(d)^{\delta} L(d) + p_L(d)\, L(d) \\
& L(d) = \left(1-\phi\right) 2\pi d
\end{aligned}
$$

The problem of construction firm can be used to derive a height supply equation. From this equation I can derive the first stage equation for height of buildings, described in equation 3. In this first stage equation, δ is a constant and c_0 is the city-specific instrument. Exogeneous variation will be provided by the interaction of earthquake risk with a time-variant shift in the cost of height, i.e. proportion of houses with elevators in the same MSA. Exogeneous variation allows to obtain an exogeneous linear prediction of height. That is, $\widehat{\log(h(d))} = L\left[\log(h) \,|\log(\delta), \log(c_0)\right]$. u is the error in the first stage estimation. The estimation of the first stage equation will allow me to test the assumption that height construction costs are convex, since δ enters as a parameter in the coefficient obtained by regressing height on the instrument.

$$
\log[h(d)] = \underbrace{\frac{1}{1-\delta}\log(\delta) + \frac{1}{1-\delta}\log(c_0)}_{\widehat{\log[h(d)]}} + \underbrace{\frac{1}{\delta-1}\log\left[p_H(d)\right]}_{\text{u}} \tag{3}
$$

Figure 2 shows how height of building decreases as we move further away from the city center. In fact, at the city-center, where houses are more expansive, land is scarce, households are willing on living on fewer square meters, and developers have incentive to build taller buildings. On the other hand, further away from the center land becomes more available and houses become cheaper leading to a decrease in height of buildings. Moreover, the derivative of equation 3 with respect to δ is negative[1], meaning that higher the convexity of vertical construction costs, more costly is to build up vertically and then lower will be the height of buildings. With the particular calibration used in Figure 2 is possible to see how higher levels of δ decrease height at a given distance and flatten the height gradient.

2.2 Equilibrium

Housing market equilibrium is given by equating housing supply ($H^S(d)$) and demand ($H^D(d)$) at each distance, that is $h(d)\,L(d) = H^S(d) = H^D(d) = H(d)\,N(d)$, where $N(d)$ is the total population at the distance d. Combining housing market equilibrium with the decomposition of $\log(h(d))$ in an endogeneous and exogeneous component, equation 3, it is possible to derive the following housing price equation.

---

[1] since $h(d) = \left[\frac{p_H(d)}{c_0\delta}\right]^{\frac{1}{\delta-1}}$ then $\frac{\partial h(d)}{\partial \delta} = -\frac{\left[\frac{p_H(d)}{c_0\delta}\right]^{\frac{1}{\delta-1}}\left[\delta log\left[\frac{p_H(d)}{c_0\delta}\right]^{\frac{1}{\delta-1}} + \delta - 1\right]}{(\delta-1)^2\delta} < 0$

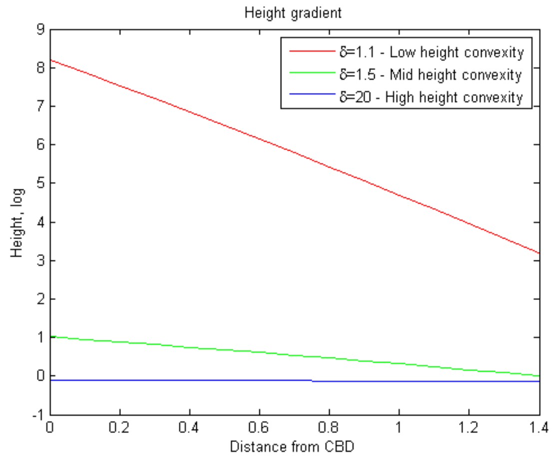


Figure 2: Height gradient

Obtained assuming $\tau = 0.1$, $p_H(0) = 1$, $c_0 = 0.4$, $\alpha = 0.3$

Height equation: $h(d) = \left[\frac{p_H(0)(1-\tau d)^{\frac{1}{\alpha}}}{c_0\delta}\right]^{\frac{1}{\delta-1}}$

$$log\left(p_H(d)\right) = \frac{\delta-1}{\delta}\left[log(w) + log(1-\tau d) + log[N(d)] - \widehat{log(h)} - log\left[\left(1-\phi\right)2\pi d\right] + log(\alpha)\right] \tag{4}$$

Moreover, using the housing market equilibrium condition is possible to derive the city population ($N(d)$) and density ($n(d)$) gradients, expressed in Equation 5 and 6 and reported in Figure 3. The density gradient is clearly a negative function of distance from the CBD. Two opposite forces explain the distribution of population with respect to distance from the CBD. On the one side, further away from the city center more land is available increasing houses built and lowering price. Population will increase with distance from the CBD up to a point in which higher transportation costs make people less willing to move further away.

$$N(d) = \frac{h^S(d)\,L^S(d)}{H} = \left[\underline{U}^{-\frac{1}{\alpha}}(1-\alpha)^{1-\alpha}\right]^{\frac{\delta}{\delta-1}} \alpha^{\frac{1}{\delta-1}}\theta^{\frac{\delta}{\alpha(\delta-1)}} w^{\frac{\delta-\alpha(\delta-1)}{\alpha(\delta-1)}}\,\hat{h}\left(1-\phi\right)2\pi d\,(1-\tau d)^{\frac{\delta-\alpha(\delta-1)}{\alpha(\delta-1)}} \tag{5}$$

$$n(d) = \frac{N(d)}{L(d)} = \left[\underline{U}^{-\frac{1}{\alpha}}(1-\alpha)^{1-\alpha}\right]^{\frac{\delta}{\delta-1}} \alpha^{\frac{1}{\delta-1}}\theta^{\frac{\delta}{\alpha(\delta-1)}} w^{\frac{\delta-\alpha(\delta-1)}{\alpha(\delta-1)}}\,\hat{h}\,(1-\tau d)^{\frac{\delta-\alpha(\delta-1)}{\alpha(\delta-1)}} \tag{6}$$

The total population of a given city can be computed using the fact that $N = \int_0^D N(d)\,\partial d$, where D is the maximal distance from the city center and it represents the horizontal spread of the city. Solving the integral it is possible to obtain the total city population, as it is shown in Equation 7. I define $f(\tau, D)$ as the part of city

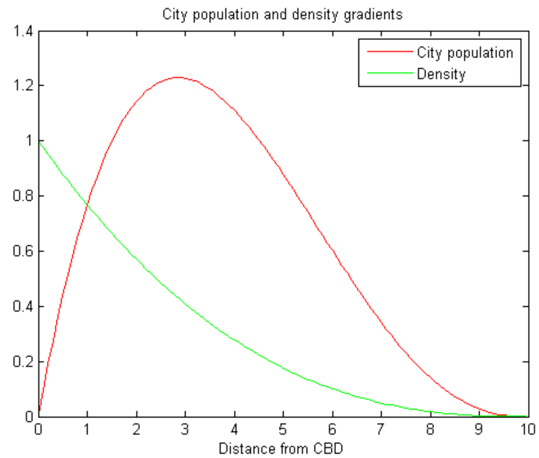


Figure 3: City population and density gradients

Obtained assuming $\left[\underline{\text{U}}^{-\frac{1}{\alpha}}(1-\alpha)^{1-\alpha}\right]^{\frac{\delta}{\delta-1}}\alpha^{\frac{1}{\delta-1}}\theta^{\frac{\delta}{\alpha(\delta-1)}}w^{\frac{\delta-\alpha(\delta-1)}{\alpha(\delta-1)}}\hat{h}\left(1-\phi\right)2\pi=1$, $\delta=20$, $\alpha=0.3$.
City population gradient: equation 5, density gradient: equation 6

The value of density at d=0 is not definite

population that depends non-linearly on the total city spread, D. The exact mathematical value of $f(\tau,D)$ is shown in equation 8. This function depends positively on D, that is higher the horizontal spread of the city higher will be the total population in that same city. Despite the fact that total population depends non-linearly on the horizontal city size, this relationship can be approximated linearly. In fact, looking at Figure 4 it is possible to see how the non-linear part of total population can be approximated as $log\left[f(\tau,D)\right]\approx\kappa+\nu log(D)$, where the approximation coefficient, ν, is roughly equal to 4. Moreover, from Figure 4 it seems that this approximation reasonably work at higher levels of D, which are the values of the horizontal city size in my data under analysis (see Panel right).

$$N=\int_0^D N(d)\,\partial d=\left[\underline{\text{U}}^{-\frac{1}{\alpha}}(1-\alpha)^{1-\alpha}\right]^{\frac{\delta}{\delta-1}}\alpha^{\frac{1}{\delta-1}}2\pi\theta^{\frac{\delta}{\alpha(\delta-1)}}w^{\frac{\delta-\alpha(\delta-1)}{\alpha(\delta-1)}}\hat{h}\left(1-\phi\right)\underbrace{\int_0^D(1-\tau d)^{\frac{\delta-\alpha(\delta-1)}{\alpha(\delta-1)}}\,d\partial d}_{f(\tau,D)} \tag{7}$$

$$f(\tau,D)=\frac{\alpha^2(\delta-1)^2-\alpha(\delta-1)(1-\tau D)^{\frac{\delta}{\alpha(\delta-1)}}\left[\alpha(\delta-1)+\delta\tau D\right]}{\tau^2\delta\left[\delta+\alpha(\delta-1)\right]} \tag{8}$$

Hence, total population in the city can be found using equation 9. Since the total horizontal spread of the city is multiplied by the proportion of the city which is developable we can interpret $\left(1-\phi\right)D$ as the exogenous part of total city spread, that does not depend on other economic variables involved in the model. I define $\hat{D}$ as the exogeneous part of the city spread, which equals $\left(1-\phi\right)D$, and this term is obtained

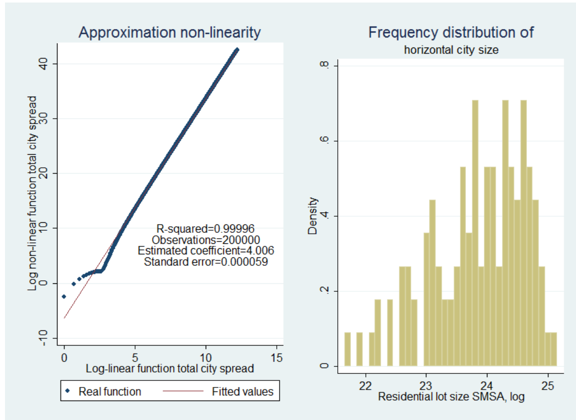


Figure 4: Approximation of $f(\tau, D)$ (panel left) and distribution horizontal city size (right)
Horizontal city size: weighted sum of the residential lot size of every house in one city-year.

Source: own calculation from American Housing Survey.

by estimating a first stage regression in an instrumental variable scenario where total city size is regressed on the interaction between the proportion of the MSA which is undevelopable, because of steepness or presence of water, and a population Bartik shifter (which will exogeneously pushes for an increase in city size). The first stage equation for total city size is reported in Equation 10, and $\widehat{log(D)} = \log\left[(1-\phi)D\right] = L\left[\log(D) \mid (1-\phi), \kappa_D\right]$, where κ_D is a constant term.

$$log(N) = log[N(d)] + \nu\widehat{log(D)} - \frac{\delta - \alpha(\delta-1)}{\alpha(\delta-1)} log(1-\tau d) - log\left[(1-\phi)d\right] \tag{9}$$

$$log(D) = \kappa_D + \xi \log(1-\phi) + \varepsilon \tag{10}$$

2.3 Structural equations

Using the spatial indifference (1), labour demand (2), the house price equations (4), and the population equation (9) it is possible to derive three structural equations that model the behaviour of house prices, wages and city population in function of the exogeneous part of height and horizontal spread and the two city-specific parameters: productivity (A) and amenities (θ). The three structural equations are Equations 11 to 13, where the variable K_i represents constant terms that influences house prices, wages and population respectively.

$$log\left[p_H(d)\right] = K_P + \frac{(\delta-1)\left[log(A) + \beta log(\theta)\right] - (\delta-1)(1-\beta-\gamma)\left[\widehat{log(h)} + \nu\widehat{log(D)}\right]}{\delta(1-\beta-\gamma) + \alpha\beta(\delta-1)} + \frac{1}{\alpha} log(1-\tau d) \tag{11}$$

Theoretical predictions:

$$\frac{\partial log\left[p_H(d)\right]}{\partial\widehat{log(h)}}<0 \quad \frac{\partial log\left[p_H(d)\right]}{\partial\widehat{log(D)}}<0\frac{\partial log\left[p_H(d)\right]}{\partial log(A)}>0 \quad \frac{\partial log\left[p_H(d)\right]}{\partial log(\theta)}>0$$

$$log(w)=K_w+\frac{\alpha(\delta-1)\,log(A)-\delta\left(1-\beta-\gamma\right)log(\theta)-\alpha(\delta-1)\left(1-\beta-\gamma\right)\left[\widehat{log(h)}+v\widehat{log(D)}\right]}{\delta\left(1-\beta-\gamma\right)+\alpha\beta(\delta-1)} \tag{12}$$

Theoretical predictions:

$$\frac{\partial log(w)}{\partial\widehat{log(h)}}<0 \quad \frac{\partial log(w)}{\partial\widehat{log(D)}}<0\frac{\partial log(w)}{\partial log(A)}>0 \quad \frac{\partial log(w)}{\partial log(\theta)}<0$$

$$log(N)=K_N+\frac{[\delta-\alpha(\delta-1)]\,log(A)+\delta\left(1-\gamma\right)log(\theta)+\alpha(\delta-1)\left(1-\gamma\right)\left[\widehat{log(h)}+v\widehat{log(D)}\right]}{\delta\left(1-\beta-\gamma\right)+\alpha\beta(\delta-1)} \tag{13}$$

Theoretical predictions:

$$\frac{\partial log(N)}{\partial\widehat{log(h)}}>0 \quad \frac{\partial log(N)}{\partial\widehat{log(D)}}>0\frac{\partial log(N)}{\partial log(A)}>0 \quad \frac{\partial log(N)}{\partial log(\theta)}>0$$

As it is possible to see from equations 11 to 13, the direct effect of height of building is to decrease house prices. Since house prices are lower workers can be compensated by lower wages and then firms can hire more people leading to higher city population. A more spread city, higher D, will also produce lower wages, houses prices and higher population. An increase in city-specific productivity is predicted to increase wages, house prices and population. Higher amenities are expected to increase population and then house prices but they are expected to decrease wages, since people are willing to give up part of their wages to live in cities with higher amenities. The different predicted effect of productivity and amenity on wages is key in order to separately identifying the effect of height of buildings and horizontal spread of city on these two city-specific variables, as I will explain in the next section.

2.4 Reduced form equations

The estimation of the effects of vertical and horizontal spread of a city on house prices, wages and density can be used to retrieve their effect on the two unobserved variables: productivity and amenities. The estimation strategy relies on obtaining reduced form estimate of the effect on house prices, wages and density, and then use them to understand which effect on productivity and amenities are consistent with the model under discussion. To obtain reduced form elasticities I will follow the strategy proposed by Glaeser (2008) and I assume that both productivity and amenities depend linearly on both the horizontal and vertical city development. Let's define λ_{AH} and λ_{AD} as the effect of increasing the exogeneous part of building height and city size on productivity, and $\lambda_{\theta H}$ and $\lambda_{\theta D}$ as their effects on amenities. Further, I define κ_i as constants, and μ_i as the residual terms that explains productivity and amenities which are uncorrelated with $\hat{h}$ and $\hat{D}$.

$$log\,(A) = K_A + \lambda_{AH}\widehat{log\,(h)} + \lambda_{AD}\nu\widehat{log\,(D)} + \mu_A \tag{14}$$

$$log\,(\theta) = K_\theta + \lambda_{\theta H}\widehat{log\,(h)} + \lambda_{\theta D}\nu\widehat{log\,(D)} + \mu_\theta \tag{15}$$

These equations model the fact that developing a city vertically or horizontally might have different externalities on productivity and amenities, which are not taken into account by agents in their maximization problem. The goals of this paper are to estimate the λ terms, understand if these externalities are present, and assess if they differ according to the direction of city development, vertical or horizontal. The first advantage of my model is to allow me to microfound the externality from horizontal city size, without assuming that city size is given. Secondly, I can separately obtain estimate of the effect of vertical and horizontal city development which are internally consistent with my model.

Combining equations 11, 12 and 13 with the equations describing the externality of city height on amenities, productivity and city size (14 and 15) I can derive reduced form equations for house prices, wages and city population (equations 16, 18 and 21). The reduced form elasticities depend on model parameters as it shown in equations 17, 19 and 21.

$$log\left[p_H(d)\right] = K_P + B_{PH}\widehat{log\,(h)} + B_{PD}\nu\widehat{log\,(D)} + g_1(d) + \mu_P \tag{16}$$

$$B_{Pi} = \frac{(\delta-1)\left[\lambda_{Ai} + \beta\lambda_{\theta i} - \left(1-\beta-\gamma\right)\right]}{\alpha\beta(\delta-1) + \delta\left(1-\beta-\gamma\right)} \text{ for i=}\{H,D\} \tag{17}$$

$$log\,(w) = K_w + B_{wH}\widehat{log\,(h)} + B_{wD}\nu\widehat{log\,(D)} + \mu_w \tag{18}$$

$$B_{wi} = \frac{\alpha(\delta-1)\lambda_{Ai} - \delta\left(1-\beta-\gamma\right)\lambda_{\theta i} - \alpha(\delta-1)\left(1-\beta-\gamma\right)}{\alpha\beta(\delta-1) + \delta\left(1-\beta-\gamma\right)} \text{ for i=}\{H,D\} \tag{19}$$

$$log\,[N] = K_N + B_{Nh}\widehat{log\,(h)} + B_{ND}\nu\widehat{log\,(D)} + \mu_N \tag{20}$$

$$B_{Ni} = \frac{[\delta-\alpha(\delta-1)]\lambda_{Ai} + \delta\left(1-\gamma\right)\lambda_{\theta i} + \alpha(\delta-1)\left(1-\gamma\right)}{\alpha\beta(\delta-1) + \delta\left(1-\beta-\gamma\right)} \text{ for i=}\{H,D\} \tag{21}$$

The reduced form estimates show the total effect of height and city size on house prices, wages and density. In fact, height and city size will have a negative direct effect on house prices and wages, and a positive direct effect on population. However, the total effect of height and city size might have a different sign because of their interaction with productivity and amenities. In fact, if one of the city development influences positively (negatively) productivity there will be an indirect positive (negative) effect on house prices, wages, and population. On the other hand, if city development influences positively (negatively) amenities there will be an indirect positive (negative) effect on house prices and population, and a indirect negative (positive) effect on wages. The different effect of θ and A on wages will allow me to identify the externalities parameters, λ, which quantify the effect of city development on amenities and productivity. Vertical and horizontal city development might have different effect on productivity and amenities and this fact will be captured by different magnitudes and sign of the estimated elasticities.

2.5 Retrieving the externality parameters

Regressing house prices and wages on height and city size it is possible to estimate the reduced form coefficients $\hat{B}_{PH}$, $\hat{B}_{PD}$, $\hat{B}_{wH}$, and $\hat{B}_{wD}$, which reflect the total effect of height and city size on these variables. The estimation will be done using instrumental variables techniques. The reduced form coefficients depends on the direct effect of height and city size and their indirect effect via productivity and amenities, as it is shown in equations 17, 19 and 21. Once the reduced form coefficients are estimated it is possible to retrieve the externality coefficients. In fact, Equations 17 and 19 provide a system of 4 equations ($\hat{B}_{PH}$, $\hat{B}_{PD}$, $\hat{B}_{wH}$, and $\hat{B}_{wD}$) and 4 unknowns (λ_{AH}, $\lambda_{\theta H}$, λ_{AD}, and $\lambda_{\theta D}$). Solving this system of equations I can calculate the λ parameters as it is shown in equations 22 and 23[2]. Using this procedure I can interpret the reduced form elasticities in terms of what would happen to productivity and amenities if a city develop vertically or horizontally.

$$\lambda_{\theta i} = \alpha \hat{B}_{Pi} - \hat{B}_{wi} \text{ for i=}\{H, D\} \tag{22}$$

$$\lambda_{Ai} = \frac{\beta(\delta-1)\hat{B}_{wi} + \delta\left(1-\beta-\gamma\right)\hat{B}_{Pi} + (\delta-1)\left(1-\beta-\gamma\right)}{\delta-1} \text{ for i=}\{H, D\} \tag{23}$$

The parameter α can be estimated as the share of household expenditure in housing. The parameters β and γ can be estimated as the share of labour and traded-capital in the production of firms. Finally, δ, the convexity of construction costs to height, can be estimated from the first stage equation of height of buildings.

3 Data and identification strategy

In the previous section I have shown how I can link reduced form elasticities of the effect of vertical and horizontal city development on wages and house prices to results about their effect to productivity and amenities. Equations 11 and 12 are consistent with an equilibrium model in which agents have the same utility between cities and

[2] In order to account for the possible non-linear effects of city size the estimated elasticities from IV regressions should be divided by ν before plugging them in Equations 22 and 23.

inside the same city. The model produces an additional reduced form equation, linking population to city expansion. However, as I have shown in the previous section two equations are enough to estimate the two externality parameters of interest (λ_A and λ_θ). Therefore I will estimate the wage and house prices reduced form elasticities using only the following models (Equations 24 and 25) since using aggregating data for population will decrease the number of observations for my estimation.

$$log\left(p_{H_{i(m),t}}\right) = \kappa_m^P + \kappa_t^P + \omega_{PH} log\left(h_{i,t}\right) + \omega_{PD} log\left(D_{m,t}\right) + \mu_{i,t}^P \tag{24}$$

$$log\left(w_{i(m),t}\right) = \kappa_m^w + \kappa_t^w + \omega_{wH} log\left(h_{i,t}\right) + \omega_{wD} log\left(D_{m,t}\right) + \mu_{i,t}^w \tag{25}$$

where i, m, and t refers to individual i living in metropolitan area (MSA) m at time t. In my model different constant variables defined at MSA level can influence house prices and wages, therefore I include in the reduced form models both MSA fixed effects, κ_m, and year fixed effects, κ_t. Equations 24 and 11 differ in that I do not include distance from city center as additional control. This is motivated by the fact that the source of exogeneous variation used does not vary inside a city and then my instruments will not be correlated with distance from the CBD. Then, distance from CBD can safely enter inside the error model $\mu_{i,t}^P$ without worries of endogeneities. This equations are estimated with IV methods.

The main source of data is the Metropolitan sample of the American Housing Survey (AHS) and the supplemental sample of housing units in Chicago, Detroit, New York, Northern New Jersey, and Philadelphia in the National sample of the AHS. The AHS is a panel database of observations at housing level active from 1975 in the United States which includes information about individual housing, neighbourhood, demographic, and labour characteristics. A random sample of residential houses in the U.S. is interviewed and I use as individual information only responses by the householder in the housing unit.

The variable w is proxied by the wage and salary income of the householder, while p_H are represented by the current market value of unit[3]. Every house is asked the number of stories in the same unit, and this proxies h. To compute the total size of city I compute the weighted sum of the residential lot size of every house in one particular city in one particular year. In order to have the same definition of the variables throughout years I use the AHS data from 1998 to 2013, which includes information for more than 460.000 houses in 55 cities and a total of 113 cities-years couples. I use the sample of people with a positive wage and salary for both the estimation of the reduced form of wages and house prices [4].

Table 1 reports the summary statistics for the main variables of interests for each MSA. In particular, it includes the mean number of floors, the residential lot size area (in squared km) and what would have been the radius of the city if this area is distributed as a circle inside the city. As it is possible to notice the tallest cities are New

[3]This is a self-reported measure, households should assess what they think is the value of their houses if they were selling it.

[4]This is given by the fact that house prices is a self-reported measure and some missing observations are present which change the estimating sample in the house price equation. However, non-responses and the possible measurement error are not likely to be correlated with any geological, geographical and external shocks and my estimations do not suffer of any related bias.

York, NY (5.5 stories), Washington, DC MD VA (3.6), Boston, MA (3.4), Chicago, IL (3.3), and Philadelphia, PA NJ (3). On the other hand, the shortest cities, in terms of number of floors, are Tucson, AZ (1.3), Oklahoma City, OK (1.3), Riverside-San Bernardino, CA (1.4), Forth Worth-Arlington, TX (1.4), and Phoenix, AZ (1.4).

The most spread cities are Washington, DC MD VA (with an implied radius of 24.4 km), Philadelphia, PA NJ (23.55 km), Detroit, MI (23.3 km), Minneapolis St. Paul, MN (22.5 km), and Nashville, TN (21.98 km). To have a first understanding about whether a city is more vertical or horizontal developed I have computed a height density proxy by dividing the mean number of floors by the imputed city radius. Using this number it appears that the city which are more vertically developed are San Francisco, CA, Buffalo, NY, Providence, RI, Anaheim Santa Ana, CA, and Milwaukee, WI. Conversely, the most horizontal spread cities are Oklahoma City, OK, Nashville, TN, Houston, TX, San Antonio, TX, and Los Angeles, CA.

3.1 Instruments

Estimation of equations 24 and 25 by means of OLS is likely to produce biased estimates of the effect of height and city size because of omitted variable and reverse causality biases. Therefore, in order to obtain causal estimates I employ an instrumental variable strategy. In fact, the theoretical reduced form estimates are consistent with a model that considers height and city size as endogeneous variables. My model produces first stage equations in which height and city size depend on exogeneous factors, as it is reported in Equations 3 and 10. In order to include in my econometric model MSA and year fixed effects I need instruments that are both time and geographical variants. In particular, I instrument height of buildings using the interaction between earthquake risk and proportion of houses with elevators in the MSA. In addition, I instrument city size using the interaction between the proportion of land in a MSA which is undevelopable and a population Bartik shifter.

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”. Earthquakes are then one of the main threats to the safety of tall buildings. Therefore, building taller structures in places with higher earthquake risk considerably increases construction costs because of the additional reinforcements needed to guarantee safety. I exploit geographical variation of seismic risk inside U.S. to obtain exogeneous variation to instrument height of buildings. Exogeneity of this instrument is additionally guaranteed by the MSA fixed effects. In fact, if seismic risk is correlated with other natural advantages (such as the presence of water) which might explain population density these fixed effects will control for this possible threats to identification.

Earthquake risk is measured as the mean value of the seismic hazard curve in all the ZIP codes of one MSA. Hazard curve is measured as the 2 percent probability of exceedance in 50 years of mean peak ground acceleration. Data come from the 2014 U.S. Geological Survey (USGS) National Seismic Hazard Maps. The states with the highest earthquake risk are in California, the Pacific Northwest, the Intermountain West and the South (in particular Memphis, TN and Charleston, SC).

Table 1: Summary statistics

MSA	Floors	Lot	Area	Radius	Density	Wage	House price	Year
Anaheim Santa Ana, CA	1.7	1,075	291.1	5.4	0.321	40,661	596,965	2011
Atlanta, GA	2.1	5,048	3239	18.1	0.114	35,135	196,409	2011
Austin, TX	1.8	5,432	2752	16.7	0.105	38,506	255,760	2013
Baltimore, MD	2.9	3,193	2822	16.9	0.174	41,876	298,979	2013
Birmingham, AL	1.7	6,106	1038	10.3	0.166	26,928	164,965	2011
Boston, MA	3.4	3,217	1933	14.0	0.244	43,327	505,857	2013
Buffalo, NY	2.7	4,319	606.4	7.8	0.346	25,710	136,531	2011
Charlotte, NC	1.7	4,735	1348	11.7	0.146	32,536	199,188	2011
Chicago Area, IL	3.4	2,084	3654	19.2	0.175	37,349	252,521	2013
Cincinnati, OH KY IN	2.6	6,262	1672	13.0	0.198	32,543	179,800	2011
Cleveland, OH	3.0	3,549	1009	10.1	0.297	29,251	149,423	2011
Columbus, OH	2.4	4,725	1163	10.9	0.222	34,869	162,686	2011
Dallas, TX	1.6	3,514	1947	14.0	0.115	41,013	184,158	2011
Denver, CO	2.6	2,762	989.2	10.0	0.260	35,318	259,075	2011
Detroit, MI	2.5	3,573	5392	23.4	0.109	30,337	139,155	2013
Fort Worth Arlington, TX	1.4	3,584	1126	10.7	0.132	33,801	153,871	2011
Hartford, CT	2.8	5,746	1699	13.1	0.210	39,049	248,813	2013
Houston, TX	1.7	2,398	4076	20.3	0.084	37,681	180,945	2013
Indianapolis, IN	1.9	4,252	1138	10.7	0.179	32,532	155,332	2011
Jacksonville, FL	1.5	3,689	1729	13.2	0.117	28,773	199,233	2013
Kansas City, MO KS	2.4	6,297	2142	14.7	0.161	31,049	157,465	2011
Las Vegas, NV	1.8	2,078	1181	10.9	0.164	27,805	189,443	2013
Los Angeles Long Beach, CA	1.8	1,800	3217	18.1	0.102	29,297	487,953	2011
Louisville, KY-IN	2.1	6,916	3011	17.5	0.118	27,458	173,278	2013
Memphis, TN AR MS	1.6	4,845	913.1	9.6	0.169	30,233	155,597	2011
Miami Hialeah, FL	2.9	1,610	1999	14.2	0.203	26,799	285,416	2013
Milwaukee, WI	2.7	4,156	720.2	8.5	0.315	30,398	215,246	2011
Minneapolis St. Paul, MN	2.8	5,086	5010	22.5	0.126	40,795	236,392	2013
Nashville, TN	1.8	9,298	4771	22.0	0.083	32,857	223,507	2013
New Orleans, LA	1.5	1,928	420.7	6.5	0.231	27,199	198,234	2011
New York Area, NY	5.5	2,440	4090	20.4	0.271	38,625	529,811	2013
Norflok Virginia Beach, VA NC	1.8	3,147	701.1	8.4	0.212	32,082	257,831	2011
Northern New Jersey Area, NJ	3.0	2,599	3873	19.8	0.149	41,533	388,824	2013
Oakland, CA	1.9	1,810	484.3	7.0	0.269	42,258	504,124	2011
Oklahoma City, OK	1.3	6,150	2751	16.7	0.081	30,222	157,968	2013
Orlando, FL	1.6	2,381	1673	13.0	0.124	29,271	171,376	2013
Philadelphia, PA NJ	3.0	3,432	5475	23.6	0.128	33,915	259,832	2013
Phoenix, AZ	1.5	1,645	777.1	8.9	0.164	28,529	203,354	2011
Pittsburgh, PA	2.8	5,345	1957	14.1	0.196	29,856	151,619	2011
Portland, OR WA	2.1	5,111	1662	13.0	0.161	33,391	276,048	2011
Providence, RI	2.7	4,135	697.1	8.4	0.325	30,696	270,911	2011
Richmond, VA	2.0	8,252	3550	19.0	0.106	33,685	226,602	2013
Riverside San Bernardino, CA	1.4	2,268	1273	11.4	0.123	29,866	259,410	2011
Rochester, NY	2.7	6,064	2069	14.5	0.188	27,839	150,277	2013
Sacramento, CA	1.5	4,511	1414	12.0	0.127	32,567	273,124	2011
Salt Lake City Ogden, UT	2.2	1,440	489.6	7.0	0.309	26,021	154,270	1998
San Antonio, TX	1.5	4,931	3202	18.0	0.085	28,938	168,052	2013
San Diego, CA	1.8	2,890	922.7	9.7	0.190	35,360	468,160	2011
San Francisco, CA	2.8	1,612	236.2	4.9	0.564	48,971	864,621	2011
San Jose, CA	1.7	2,413	411.4	6.5	0.268	54,479	722,973	2011
Seattle, WA	2.3	3,204	3166	17.9	0.129	45,349	362,837	2013
St. Louis, MO IL	2.3	5,502	2275	15.2	0.150	31,020	171,597	2011
Tampa St. Petersburg Clearwater, FL	1.7	2,573	2497	15.9	0.107	27,697	164,642	2013
Tucson, AZ	1.3	2,826	970.1	9.9	0.133	24,895	180,997	2013
Washington, DC MD VA	3.6	4,011	5907	24.5	0.147	56,646	442,813	2013
Total sample	2.2	3,890	2,100	14	0.180	34,087	267,714	

MSA is the Standard Metropolitan Statistical Area, $Year$ refers to the last year in which observations from that MSA were in the sample, $Floor$ is mean number of stories in buildings, Lot is the mean lot size in squared meters, $Area$ is the area in the MSA occupied by residential buildings in squared km, $Radius$ is $\frac{\sqrt{Area}}{\pi}$, $Density$ is the ratio between floors and radius, $Wage$ is mean level of wage and salaries of householder, $House\ price$ is the mean market value of the units.
Source: own calculation from American Housing Survey. Averages and sums weighted by AHS survey weights.

The history of tall buildings is deeply related with the evolution of elevators. According to Bernard (2014) elevators made tall buildings possible and they are one of the most important engineering discoveries to boost vertical development of cities. I use elevators as an additional instrument for height of buildings. In particular I use the proportion of houses with an elevator in each MSA in order to obtain both cross-section and time variation of the exogeneous part of height. This data comes from the American Housing Survey. This variable represent the accessibility of the elevator technology in every city at any point of time.

The elevator sector is currently experiencing important technology improvements, in particular because of aging population. The introduction of electric panels with inverter for speed variation and the pneumatic vacuum elevators, which do not require a separate machine room for the engine, are making elevators cheaper and more available, as it is shown in Figure 5. New technology electric elevators require less space for installation mainly because many of the components are installed on the top of the elevator shaft. In addition to the fact that elevators can now being installed almost anywhere, elevators became also considerably faster because of the new electric panels. In fact, elevator speed improved considerably since the first Otis elavator (which speed was around 12 m/min). Elevators built in the 70s and 80s average speed is 0.7 m/s, those built in the 2000s arrives to 2 m/s, with even higher speeds in skyscrapers elevator (Burj Khalifa elevator speed is around 10 m/s)[5].

Moreover, despite the elevator costs in U.S. are similar, the installation costs can vary significantly by location because of material availability, specialized labour force, and local and state building and elevator codes. Controlling for both MSA and year fixed effect I avoid to capture the part of elevator improvement which is coming from a general productivity development or any other local characteristics that can influence housing supply. Therefore, the instrument I use for height of buildings is the log ratio between seismic risk and proportion of houses in a MSA with elevators. This combined instrument is both time and geographical variant, and then is not collinear with any fixed effect.

Saiz (2010) demonstrated that geography plays an important role for the city development. Similarly, Harari (2015) shows that the presence of water bodies can influence the shape of cities. Therefore, presence water bodies or steepness are important obstacles to the horizontal development of cities. I use the percentage of undevelopable area produced by Saiz (2010) to instrument city size. Saiz (2010) defines undevelopable area as area corresponding to steep slopes (above 15 %), oceans, lakes, wetlands and other water features.

However, according to Saiz (2010) physical constraints might be binding only in larger metropolitan areas. Indeed, he interacts physical constraints with population level to estimate housing supply elasticities. I use a Bartik shifter to obtain exogeneous population pressure to city size. Bartik instruments have been developed from the work of Bartik (1991) and they exploit the fact that national economic changes across industries affect in a differ way local economies based on their industrial composition. That is, if a particular sector is experiencing a national boom given by ex-

[5]The information about elevator improvements have been obtained from discussions with workers at Schindler elevator company, Milan

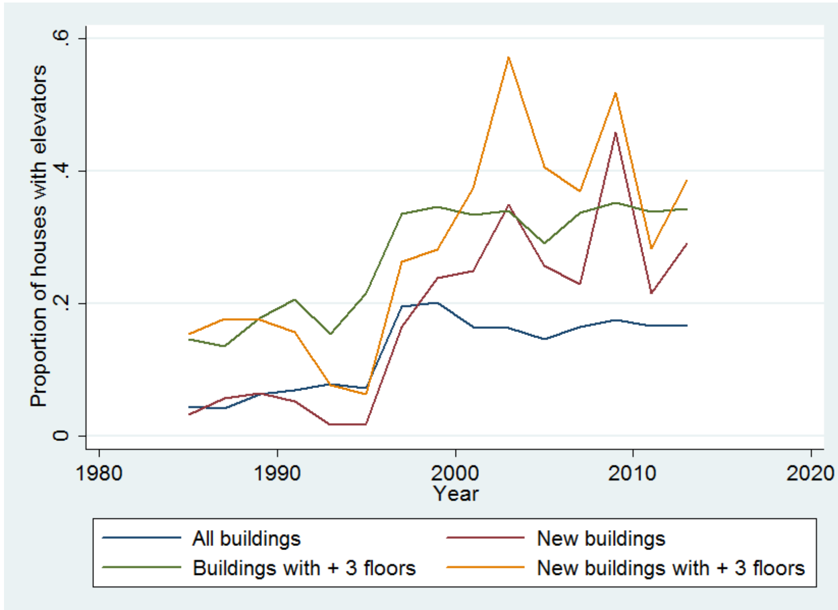


Figure 5: Time series of proportion of houses with elevators in the U.S.
Proportion of houses with elevators: weighted average of houses with elevators in U.S. in one year.
Source: own calculations based on National Sample of American Housing Survey.

ogeneous reasons to local labor demand, cities which are more specialized in that industry will be predicted to increase labour demand and then to increase population. I define the Bartik shock in Equation 26, where j correspond to industry, m is MSA, $-m$ are all other cities that are that MSA, emp is employment level and $t = 0$ is 1998, the year in which my sample begins. Data for employment level at industry level comes from the U.S. Census County Business Pattern (CBP)[6]. Hence, my geographical and time variant instrument is the log ratio of the Bartik shifter and the proportion of undevelopable area[7].

$$Bartik_{m,t} = \sum_j \left(emp_{j,-m,t} - emp_{j,-m,0}\right) \frac{emp_{j,m,0}}{emp_{US,m,0}} \tag{26}$$

3.2 First stage regressions

My model is internally consistent with my estimation strategy since it produces first stage equations for both height and city size. That is, I can predict the exogeneous part of height and city size by regressing those variables on exogeneous instruments, as it is shown in the model Equations 3 and 10, respectively. Relevance condition of the instruments is shown separately for height and city size in Tables 2 and 3.

[6]Industries refer to the 2 digits NAICS industries in the CBP. There are 20 industries present. 1998 has been chosen also starting year also because this is the year in which NAICS system substituted the previous SIC industry classification system.

[7]Since the Bartik shifter can be negative I take the log of this interaction summed with the absolute number of the minimal value plus 1 to avoid having the log of a negative number

From the first two columns of Table 2 it is possible to see that alone seismic risk is decreasing height of buildings, while cities with more elevators tends to be taller. I combine these variables by dividing seismic risk by the proportion of elevators in a MSA. Combining these two variables together I obtain my instrument which is likely to negatively affect height of buildings. That is, the negative effect of seismic risk is higher when elevators are less developed. Similarly, the positive effect of elevators is smaller in places with more seismic risk.

I can impute the δ coefficient, the convexity of housing cost with respect to height, by using these estimates and Equation 3, as it is shown in Equation 27. This number is necessary in order to obtain the externality parameters of the effect of height and city size on productivity and amenities. The estimated δ is around 24 and it confirms the convexity assumption made in the model, since it is a number bigger than 1.

$$\begin{aligned} \text{Model first stage: } \log\left[h\left(d\right)\right] &= \frac{1}{1-\delta}\log\left(\delta\right)+\frac{1}{1-\delta}\log\left(c_0\right)+u \\ \text{Estimated first stage: } \log\left(h_{i,t}\right) &= \kappa_m+\kappa_t+\phi^{FS}\log\left(c_{0\,m,t}\right)+u_{i,t} \\ \rightarrow\hat{\delta} &= 1-\frac{1}{\hat{\phi}^{FS}} \end{aligned} \tag{27}$$

Table 2: First stage height buildings

VARIABLES	(1) Log floors	(2) Log floors	(3) Log floors	(4) Log floors
Average hazard curve MSA	-0.00422***			
	(5.71e-05)			
Log proportion elevators MSA		0.347***		
		(0.00271)		
Log hazard / Proportion elevators			-0.0480***	-0.0438***
			(0.00109)	(0.0122)
Observations	380,692	380,692	380,692	380,692
R-squared	0.020	0.129	0.013	0.287
MSA FE	NO	NO	NO	YES
Year FE	NO	NO	NO	YES
Estimation	OLS	OLS	OLS	OLS
Implied delta	237.82	3.88	21.84	23.85
s.e. delta	3.2	.02	.47	6.4

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Log floors: Logarithm of number of stories in the same housing unit of the interviewed person. *Average hazard curve MSA / hazard*: Weighted average 2 percent probability of exceedance in 50 years of mean peak ground acceleration for each MSA. *Proportion elevators MSA / proportion elevators*: weighted average of houses with elevators in U.S. in one year. *MSA FE*: Metropolitan Statistical Area fixed effects. *Year FE*: year fixed effects. *Implied delta*: estimated δ coefficient, coefficient that regulates the convexity of height construction costs. *s.e. delta*: standard errors of the estimated δ coefficient. Regression weighted by AHS survey weights.

Similarly, Table 3 shows the first stage for city size. As imagined, cities are less spread if a bigger proportion of available land is undevelopable and they are more spread if they suffer an exogeneous population shock. I combine these variables by dividing the Bartik shock by the proportion of undevelopable area. The combined instrument has a positive and significant effect of making exogeneously cities more spread horizontally. That is, the positive effect of a Bartik shock in population is bigger when a city has lower limits to develop. Similarly, the negative effect of having undevelopable areas is stronger when there is an adverse labour demand shock.

Table 3: First stage total city size

	(1)	(2)	(3)	(4)
VARIABLES	Log lot	Log lot	Log lot	Log lot
Log undevelopable area MSA	-0.120***			
	(0.00111)			
Log Bartik shifter		0.170***		
		(0.000806)		
Log Bartik / undevelopable area			0.0996***	0.0496***
			(0.000765)	(0.000182)
Observations	371,104	381,888	371,104	371,104
R-squared	0.028	0.043	0.018	0.941
MSA FE	NO	NO	NO	YES
Year FE	NO	NO	NO	YES
Estimation	OLS	OLS	OLS	OLS

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Log lot: Logarithm of weighted sum of residential lot size of houses in one MSA-year. *Undevelopable area MSA* / *undevelopable area*: Percent of MSA area which is undevelopable because of water bodies or steepness. *Bartik shifter* / *Bartik*: Bartik labour demand shock, constructed as in Equation 26. *MSA FE*: Metropolitan Statistical Area fixed effects. *Year FE*: year fixed effects. Regression weighted by AHS survey weights.

The final first stages that used in the estimation of the causal effect of height and city size are reported in Table 4. In those first stages both floors and lot size of the MSA will be regressed on their own instrument but also on the instrument associated to the other variable. Relevance condition is satisfied for both equations. In fact, the interaction between seismic risk and proportion of elevators is significantly affecting floors in a negative way, while the interaction of the Bartik shifter and the proportion of land which is undevelopable positively shift the city size. The δ I will use to retrieve the externality parameters has a value of 24.03 and it is significantly different from 0.

Proper test of relevance and weak instrument problems are reported in Table 5. I separately report F-statistics and under-identification test after regressing the two reduced forms of interests: using wages and house prices as dependent variables. Overall, my instrument appears to be relevant and they do not suffer of weak problems. Indeed, the Kleinbergen-Paap statistic is suggesting that the estimating model is not underidentified and the excluded instruments are relevant. F-statistics and critical values for weak instruments should be adapted for the particular case of multi-endogeneous variables, as it has been showed by Stock and Yogo (2005). Angrist et al.

Table 4: First stage regressions

VARIABLES	(1) Log floors	(2) Log lot
Log hazard / Proportion elevators	-0.0434***	0.0642***
	(0.0124)	(0.00316)
Log Bartik / undevelopable area	-0.00168	0.0485***
	(0.00103)	(0.000192)
Observations	370,019	371,104
R-squared	0.284	0.942
MSA FE	YES	YES
Year FE	YES	YES
Estimation	OLS	OLS
Implied delta	24.03	.
s.e. delta	6.57	.

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Log floors: Logarithm of number of stories in the same housing unit of the interviewed person. *Log lot*: Logarithm of weighted sum of residential lot size of houses in one MSA-year. *hazard*: Weighted average 2 percent probability of exceedance in 50 years of mean peak ground acceleration for each MSA. *proportion elevators*: weighted average of houses with elevators in U.S. in one year. *undevelopable area*: Percent of MSA area which is undevelopable because of water bodies or steepness. *Bartik*: Bartik labour demand shock, constructed as in Equation 26. *MSA FE*: Metropolitan Statistical Area fixed effects. *Year FE*: year fixed effects. *Implied delta*: estimated δ coefficient, coefficient that regulates the convexity of height construction costs. *s.e. delta*: standard errors of the estimated δ coefficient. Regression weighted by AHS survey weights.

(2009) suggests how to create F-statistics for each endogenous regressor separately, partialling out for the effect passing through the other regressor, while Cragg and Donald (1993) propose a single F-statistics for testing overall weak problem of the different instruments. From Table 5 I can conclude that my instruments do not suffer of weak instrument problem either separately or jointly.

Table 5: Relevance and weak instruments test

Statistics	Dependent variable Wage	 House prices
Angrist-Pischke F-statistics: floors	9.88***	7.56**
Angrist-Pischke F-statistics: lot	14399.85****	7823.73****
Cragg-Donald F-statistics	8.26****	6.91***
Kleibergen-Paap LM statistic	9.883****	3.78*

**** Statistics higher than Stock-Yogo critical values for 10% maximal IV size,

*** for 15% maximal IV size, ** for 20% maximal IV size, * for 25% maximal IV size

Appendix A includes different set of specifications to test the robustness of my first stage. In particular, I show that the first stage results are robust to the use of different standard errors clustering and the inclusion of different fixed effects. Moreover, I also perform a placebo test in which I randomly assign values of the instruments and report the distribution of the first stage coefficients.

4 Empirical results

In this section I present the results of the reduced form estimates obtained from regressing wage and house prices on height and total city size. Table 6 contains the results when the dependent variable is wage. Column 1 reports the OLS estimate, column 2 adds MSA and year fixed effects, column 3 presents the IV estimate and finally column 4 regresses wages directly on the instruments. As it is possible to see both height and lot size have not a significant effect on log wages as soon as I instrument those two variables. In order to interpet the no-significant result it is important to keep in mind the possible opposite counterbalancing effect of productivity and amenities on wages. In fact, if height or city size increases productivity this will lead to increasing wages. If amenities also increase then the indirect effect on wages is to decrease them and at the end the final effect on wages might not be significant. Therefore, in order to understand if these indirect effects are explaining the reduced form elasticities I first need to estimate the house price elasticities.

Table 6: Reduced form estimates for wage

VARIABLES	(1) Log wages	(2) Log wages	(3) Log wages	(4) Log wages
Log floors	0.0853***	0.0391***	0.0606	
	(0.00806)	(0.00917)	(0.734)	
Residential lot size MSA	0.0485***	0.127***	0.106	
	(0.00501)	(0.0286)	(0.0825)	
Log hazard / Proportion elevators				0.00365
				(0.0351)
Log Bartik / undevelopable area				0.00492*
				(0.00287)
Observations	226,582	226,582	220,399	220,399
R-squared	0.003	0.025	.	0.024
MSA FE	NO	YES	YES	YES
Year FE	NO	YES	YES	YES
Estimation	OLS	OLS	IV	OLS

Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Log wages: logarithm of wage and salary income of the householder living in the house. *Log floors*: Logarithm of number of stories in the same housing unit of the interviewed person. *Residential lot size MSA*: Logarithm of weighted sum of residential lot size of houses in one MSA-year. *hazard*: Weighted average 2 percent probability of exceedance in 50 years of mean peak ground acceleration for each MSA. *proportion elevators*: weighted average of houses with elevators in U.S. in one year. *undevelopable area*: Percent of MSA area which is undevelopable because of water bodies or steepness. *Bartik*: Bartik labour demand shock, constructed as in Equation 26. *MSA FE*: Metropolitan Statistical Area fixed effects. *Year FE*: year fixed effects. Regression weighted by AHS survey weights.

Reduced form estimates of the effect of vertical and horizontal city development show that both height and city size influences positively house prices. The OLS fixed effects estimation in column 2 of Table 7 suggests that both vertical and horizontal city expansions influences house prices by 0.3 and 0.06 percent, respectively. The IV estimate presented in column 3 additionally suggests that the effect of height on house prices is stronger than the one of city size. In fact, increasing floors of one city by 1 % is associated to increasing house prices by 11.3 %, while increasing the area of the city by 1 % would raise house prices only by 0.9 %. Moreover, considering the presence of non-linearities in city size this last elasticity reduces to 0.23. The difference in the OLS and the IV estimate from column 2 and column 3 can be explained by the presence of reverse causality bias in the OLS estimation. Appendix B illustrates how reverse causality works in this particular context.

Table 7: Reduced form estimates for house prices

	(1)	(2)	(3)	(4)
VARIABLES	Log h price	Log h price	Log h price	Log h price
Log floors	0.435***	0.366***	11.33***	
	(0.0115)	(0.0135)	(4.062)	
Residential lot size MSA	-0.0554***	0.0608*	0.909**	
	(0.00600)	(0.0348)	(0.458)	
Log hazard / Proportion elevators				-0.521***
				(0.0391)
Log Bartik / undevelopable area				0.0111***
				(0.00352)
Observations	143,059	143,059	139,465	139,465
R-squared	0.035	0.150	.	0.134
MSA FE	NO	YES	YES	YES
Year FE	NO	YES	YES	YES
Estimation	OLS	OLS	IV	OLS

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Log h price: logarithm of current market value of house. *Log floors*: Logarithm of number of stories in the same housing unit of the interviewed person. *Residential lot size MSA*: Logarithm of weighted sum of residential lot size of houses in one MSA-year. *hazard*: Weighted average 2 percent probability of exceedance in 50 years of mean peak ground acceleration for each MSA. *proportion elevators*: weighted average of houses with elevators in U.S. in one year. *undevelopable area*: Percent of MSA area which is undevelopable because of water bodies or steepness. *Bartik*: Bartik labour demand shock, constructed as in Equation 26. *MSA FE*: Metropolitan Statistical Area fixed effects. *Year FE*: year fixed effects. Sample: people with positive wage and salary. Regression weighted by AHS survey weights.

The results show are robust to different kind of specifications, which are reported in Appendix C. Results are consistent to the use of different form of clustered standard errors and the inclusion of several additional controls and geographical fixed effects. In particular, results do not change with inclusion of variables representing the house quality (number of baths and bedrooms, year in which the structure was built, and the square footage of the unit) or the profile of the householder (education, age, race and sex). Moreover, since improvements in the elevator sector were importantly fostered by aging population I report additional robustness checks controlling for the overall the age of houses and population. Finally, Appendix D shows the effects of vertical and horizontal city development on housing density.

4.1 Retrieving the externality coefficients

In Section 2.5 I have shown how I can use my model to map reduced form elasticities to consideration about the externality effects of developing a city vertically or horizontally. In fact, once the reduced form elasticities of wages and house prices have been estimated it is possible to use Equations 22 and 23 to estimate λ_{AH}, λ_{AD}, $\lambda_{\theta H}$, and $\lambda_{\theta D}$, that is the effect of height and city size to productivity (A) and amenities (θ) respectively.

In order to obtain the λ coefficients I need to make assumptions on the following parameters of my model: α (the share of housing into the utility), β (the share of labour into the production function), γ (the share of traded capital into the production function), and δ (the convexity of vertical construction costs). I will use the values proposed by Glaeser (2008) such that $\alpha = 0.3$, $\beta = 0.3$ and $\gamma = 0.6$. Moreover, I have shown in Section 3.2 that δ can be obtained from my first stage equation and I will use a value of 24.3.

Using the IV reduced form elasticities estimated in Tables 6 and 7 and correcting the coefficients of city size for the possible non-linearities it is possible to confirm my previous hypothesis that both productivity and amenities are increasing with height of buildings and city size. Indeed, height externality coefficients are $\lambda_{AH} = 1.322$ and $\lambda_{\theta H} = 3.338$. Similarly, city size externality coefficients are $\lambda_{AD} = 0.140$ and $\lambda_{\theta D} = 0.042$.

To have a first interpretation of these coefficients I can normalize them with respect to the increase in population that city development might generate. Using the reduced form coefficients of total population with respect to height and city size, Equation 21, it is possible to compute the effect of vertical and horizontal city development on total population. In particular, the increase of 1% in the number of building floors is predicted to increase total city population by 12.8%. The elasticity with respect to city size is 1.2%.

Normalizing the reduced form elasticities with respect to population increase I can separately obtain by how much wages and house prices are expected to increase if population increase by either increasing cities vertically or horizontally. In fact, if population increases because of an horizontal city expansion wages are predicted to increase by 0.02, an elasticity which is very similar to the estimation of agglomeration economies done by Combes et al. (2010). The elasticity of house prices with respect to population increasing horizontally is 0.19. On the other hand, if population increases via a vertical increase of the city, wages and house prices are expected to increase by 0.004 and 0.89, respectively.

From the previous normalization of reduced form coefficients it is evident that vertical city expansion have a higher positive effect on house prices and a smaller one on wages, with respect to a horizontal city expansion. This is in line with a model in which vertical city expansion and horizontal city expansion have very similar positive externalities on the overall city level productivity but increasing a city vertically lead to higher gains in amenities. In fact, the normalized λ coefficients (with tilde symbol) are $\tilde{\lambda}_{AH} = 0.103$, $\tilde{\lambda}_{\theta H} = 0.261$, $\tilde{\lambda}_{AD} = 0.115$ and $\tilde{\lambda}_{\theta D} = 0.034$.

If a city population increase by 1% following a vertical expansion of the city, productivity is expected to increase by 0.103%, while if city would have expand horizontally productivity raises by 0.115%. That is, the increase in productivity caused by adding an additional floor to every building in one city is equivalent to the increase caused by increasing city radius by 5.86 km[8]. Similarly, a 1% population increase would increase city amenities by 0.261% in case of a vertical city expansion and by 0.034% in case of a horizontal city expansion. Namely, adding one floor to all the building of the city is equivalent to increase city radius by 50 km in terms of amenity gains, since the $\tilde{\lambda}_{\theta D}$ corresponds to 13% of $\tilde{\lambda}_{\theta H}$.

The higher impact of vertical city expansion on amenities derived from the use of the structural approach proposed in this paper is consistent with what can be obtained using measurable proxies of neighbourhood quality from the American Housing Survey. In fact, Table 8 reports the IV estimate of the effect of height of buildings and total city size on a subjective rating of the neighbourhood quality as a place to live. Both vertical and horizontal expansions are predicted to increase quality of the place but height increase is stronger. Moreover, the ratio of the coefficients of horizontal and vertical expansions is similar to what I have obtained previously, that is around 15 percent.

Table 8: Reduced form estimate for neighbourhood quality

	(1)
VARIABLES	Rating neighbourhood as place to live
Log floors	3.925***
	(1.443)
Residential lot size MSA	0.478***
	(0.168)
Observations	317,308
MSA FE	YES
Year FE	YES
Estimation	IV

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Rating neighbourhood as place to live: interviewed rating of neighbourhood as place to live from 1 (worst) to 10 (best). *Log floors*: Logarithm of number of stories in the same housing unit of the interviewed person. *Residential lot size MSA*: Logarithm of weighted sum of residential lot size of houses in one MSA-year. *MSA FE*: Metropolitan Statistical Area fixed effects. *Year FE*: year fixed effects. Regression weighted by AHS survey weights.

I can decompose the estimated elasticity of house prices and wages with respect to height and city size using my model. From Equations 28 to 33 I present the decomposition of the reduced form elasticities considering the separate effect coming from productivity and amenities. Two numbers are reported, black numbers refer to the estimated numbers, while blue numbers normalize those estimation in terms of a 1% increase in population driven by the either a vertical or horizontal city expansion. That is, for the coefficients attached to height I divided those numbers by $\frac{\partial log(N)}{\partial \widehat{log(h)}}$ and by $\frac{\partial log(N)}{\partial \widehat{log(D)}}$ for the city size coefficients.

[8] Over the sample under consideration the mean of floors is 2.32 and the mean city radius is 15.21 km

$$log\left(p_H\right) = K_P + B_{PH}\widehat{log\left(h\right)} + B_{PD}\widehat{log\left(D\right)} + \mu_P \tag{28}$$

$$\hat{B}_{PH} = \underbrace{\frac{(\delta-1)\,\lambda_{AH}}{\alpha\beta(\delta-1)+\delta\left(1-\beta-\gamma\right)}}_{\text{Effect via A=[4.645;0.363]}} + \underbrace{\frac{(\delta-1)\,\beta\lambda_{\theta H}}{\alpha\beta(\delta-1)+\delta\left(1-\beta-\gamma\right)}}_{\text{Effect via }\theta\text{=[7.037;0.55]}} - \underbrace{\frac{\left(1-\beta-\gamma\right)}{\alpha\beta(\delta-1)+\delta\left(1-\beta-\gamma\right)}}_{\text{Direct effect=[-0.351;-0.027]}} = [11.33; 0.885] \tag{29}$$

$$\frac{\hat{B}_{PD}}{\nu} = \underbrace{\frac{(\delta-1)\,\lambda_{AD}}{\alpha\beta(\delta-1)+\delta\left(1-\beta-\gamma\right)}}_{\text{Effect via A=[0.491;0.405]}} + \underbrace{\frac{(\delta-1)\,\beta\lambda_{\theta D}}{\alpha\beta(\delta-1)+\delta\left(1-\beta-\gamma\right)}}_{\text{Effect via }\theta\text{=[0.088;0.073]}} - \underbrace{\frac{\left(1-\beta-\gamma\right)}{\alpha\beta(\delta-1)+\delta\left(1-\beta-\gamma\right)}}_{\text{Direct effect=[-0.351;-0.19]}} = [0.227; 0.188] \tag{30}$$

$$log\left(w\right) = K_w + B_{wH}\widehat{log\left(h\right)} + B_{wD}\widehat{log\left(D\right)} + \mu_w \tag{31}$$

$$\hat{B}_{wH} = \underbrace{\frac{\alpha\,(\delta-1)\,\lambda_{AH}}{\alpha\beta(\delta-1)+\delta\left(1-\beta-\gamma\right)}}_{\text{Effect via A=[1.393;0.109]}} - \underbrace{\frac{\delta\left(1-\beta-\gamma\right)\lambda_{\theta H}}{\alpha\beta(\delta-1)+\delta\left(1-\beta-\gamma\right)}}_{\text{Effect via }\theta\text{=[-1.227;-0.096]}} - \underbrace{\frac{\alpha\,(\delta-1)\left(1-\beta-\gamma\right)}{\alpha\beta(\delta-1)+\delta\left(1-\beta-\gamma\right)}}_{\text{Direct effect=[-0.105;-0.008]}} = [0.0606; 0.0047] \tag{32}$$

$$\frac{\hat{B}_{wD}}{\nu} = \underbrace{\frac{\alpha\,(\delta-1)\,\lambda_{AD}}{\alpha\beta(\delta-1)+\delta\left(1-\beta-\gamma\right)}}_{\text{Effect via A=[0.147;0.122]}} - \underbrace{\frac{\delta\left(1-\beta-\gamma\right)\lambda_{\theta D}}{\alpha\beta(\delta-1)+\delta\left(1-\beta-\gamma\right)}}_{\text{Effect via }\theta\text{=[-0.015;-0.013]}} - \underbrace{\frac{\alpha\,(\delta-1)\left(1-\beta-\gamma\right)}{\alpha\beta(\delta-1)+\delta\left(1-\beta-\gamma\right)}}_{\text{Direct effect=[-0.105;-0.057]}} = [0.0265; 0.0219] \tag{33}$$

A 1 %increase in city population driven by increasing height of buildings in the city is predicted to increase house prices by 0.885%. However, 40% of the positive increase in house prices can be explained by the fact that height is leading to an increase in productivity that subsequently increase house prices. Other 60% is explained by height increase city amenities and then increasing house prices. A similar increase in city population that is however driven by increasing the city horizontally is eventually increasing house prices by 0.188%. Of this 0.188 elasticity, 85% is explained by increasing productivity and 15% by increasing amenities.

Similarly, the elasticity of wages with respect to height of buildings normalized by population increase is 0.0047. This small number is explained by the counterbalancing effect between the increase in productivity and amenities that height generates. This is given by the fact productivity is expected to increase wages and amenities to decrease it. Nevertheless, since a horizontal city expansion is predicted to increase amenities in a lower amount than productivity this leads to a higher wage elasticity, i.e. 0.0219.

4.2 Heterogeneous effects

Following the previous results, an increase in building height and spread of cities is likely to increase overall city productivity, while the positive reponses in city amenities is stronger in case of an increase in the vertical expansion of a city. However, these results can mask important heterogeneities with respect to several important dimensions. The externality effects can be stronger or weaker in different cities around the U.S., such that studying the presence of heterogeneity effects can shed light on some of the mechanisms which play a role for explaining why vertical and horizontal city expansion can impact productivity and amenities.

In order to assess the presence of heterogeneous effects with respect a variable X I will augment the econometric models expressed in Equations 24 and 25 by adding terms capturing the interaction between vertical and horizontal expansions with the variable X. These new models are reported in Equations 34 and 35. When the X variable change both over MSA and over time I will add it as an additional control too. The average treatment effect (ATE) of height (and similarly for city spread) on wages (and similarly for house prices) can be found as $E\left[\frac{\partial log(w)}{\partial log(\hat{h})}\right] = \omega_{wH1} + \omega_{wH2}E[X]$. These ATEs are the one to be used to compute the externality parameters in the case heterogeneity.

$$\begin{aligned} log\left(p_{H_{i(m),t}}\right) =& \kappa_m^P + \kappa_t^P + \omega_{PH1}log\left(h_{i,t}\right) + \omega_{PH2}log\left(h_{i,t}\right) \times X_m + \\ &+ \omega_{PD1}log\left(D_{m,t}\right) + \omega_{PD2}log\left(D_{m,t}\right) \times X_m + \mu_{i,t}^P \end{aligned} \tag{34}$$

$$\begin{aligned} log\left(w_{i(m),t}\right) =& \kappa_m^w + \kappa_t^w + \omega_{wH1}log\left(h_{i,t}\right) + \omega_{wH2}log\left(h_{i,t}\right) \times X_m + \\ &+ \omega_{wD1}log\left(D_{m,t}\right) + \omega_{wD2}log\left(D_{m,t}\right) \times X_m + \mu_{i,t}^w \end{aligned} \tag{35}$$

In order to account for the endogeneity in height and total city lot size I estimate the models in Equations 34 and 35 using control function approach (see Wooldridge, 2015). This methods consists in two steps: firstly I will estimate the first stage equations and predict the residuals from these equations, secondly I will estimate Equations 34 and 35 introducing also the first stage residuals in order to control for the endogeneity of h and D. I apply bootstrap techniques to obtain consistent estimates of the standard errors.

I study whether the externality coefficients estimated in the previous Section are heterogeneous with respect to the following dimensions: the presence of density restrictions laws, the fact of being a coastal location, the level of education of the city population, and the specialization towards manufacturing of the city. The Density Restriction Index (DRI) I use is the average for each MSA index computed by Gyourko et al. (2008) in 2008. I consider city as being coastal locations if part of the city touches any sea or the great lake. Data comes from the USGS. The overall level of education of the city is measured by the proportion of people with a BA degree or more using the representative sample provided by the AHS. Finally, I compute the proportion of jobs in manufacturing for each MSA in 1998 using data from the U.S. Census County Business Pattern[9].

Tables 9 and 10 present the estimation of the reduced form for wages and house prices with heterogeneity. As it is possible to evince the effect of height on wages is mitigated if a city has density restrictions or lies on a coast. Similarly, the effect of

[9] I consider the following sectors as manufacturing: Mining, Quarrying, and Oil and Gas Extraction (NAICS 21), Construction (NAICS 23) and Manufacturing (NAICS 31-33)

height of buildings on house prices is mitigated by density restrictions and the coastal location but also by having more educated people in the city. The effect of height on both wages and house prices is stronger for cities more specialized towards manufacturing. On the other hand, the presence of density restrictions makes the effect of lot size stronger with respect to wages and weaker with respect to house prices. Having more educated people or more manufacturing reduce the overall positive effect of city spread on wages. Cities on the coast tends also to have a lower impact of city spread on house prices.

Tables 9 and 10 give evidence of heterogeneity in the reduced form effects of vertical and horizontal city expansion. In order to understand if these heterogeneities play a role via productivity or amenities I compute the externality coefficients in each of the possible heterogeneity cases considered. Results are reported in Table 11. I report the Average Treatment Effect (ATE) of the externality parameters (measured taking the expectation of the variable X creating heterogeneity) but also the Average Treatment on the Treated (ATT) and the Average Treatment on the Control (ATC) computed taking the maximum and the minimum possible value of the variable X, respectively.

Cities with more density restrictions limits the effect of height on both wages and house prices, and as a result the externality parameters of vertical expansion seems not to be affected by this heterogeneity. Nevertheless, controlling for density restrictions increase the average effect of city spread on wages and decrease its effect on house prices. As a result cities with more density restrictions, the effect of city size on amenities, $\tilde{\lambda}_{\theta D}$, is smaller than the case of no heterogeneity. The other externality coefficients considering density restrictions are on average similar to the case of no heterogeneity and the case with density restrictions. However, the elasticity of height to productivity is bigger when a city has been constrained in his density previously (the ATC of $\tilde{\lambda}_{AH}$ is bigger than the corresponding ATT), while the opposite is true for amenities externalities.

Controlling for coastal location the externality coefficients of productivity with respect to both height of buildings and city size are stronger, while the one of amenities are weaker. That is, coastal locations can produce higher amenities alone but they do not create necessarily higher productivity. If a city is on the coast (ATT) the externality on productivity is stronger by expanding horizontally rather than vertically, probably because it is already more vertical developed. However, this comes at the cost that the externality of amenities is weaker by expanding horizontally.

An important strand of the literature, began with the work of Combes et al. (2008), has considered how sorting is important to explain part of the agglomeration economies. In order to control for sorting these works usually estimate the effect of a population or density proxy in a wage or productivity equation controlling for worker fixed effects and then exploit geological and historical instruments (see Combes et al., 2010). Likewise, in my estimations the effect of lot size on wages is reduced by controlling for the interactive effect it has on education. Moreover, the elasticity of wages with respect to floors normalized by the average effect it creates on population equals 0.008, a value closer to the elasticity found by Combes et al. (2010) controlling for sorting.

Using my approach, the estimated externality coefficients of height on amenities and productivity do not seem to change because of this sorting possibility. However,

Table 9: Wage reduced form estimation in presence of heterogeneity

VARIABLES	(1) Log wage	(2) Log wage	(3) Log wage	(4) Log wage
Log floors	0.0407	0.281	0.101	0.174
	(0.908)	(0.922)	(0.879)	(0.961)
Residential lot size MSA	0.0837	0.115	0.286***	0.407***
	(0.0830)	(0.0795)	(0.0776)	(0.141)
Floor X Density restrictions	-0.148***			
	(0.0454)			
Lot X Density restrictions	0.180*			
	(0.0953)			
Floor X Coast		-0.102***		
		(0.0166)		
Lot X Coast		0.0605		
		(0.0371)		
Floor X High education			0.0607	
			(0.109)	
Lot X High education			-0.503***	
			(0.103)	
High education MSA			12.64***	
			(2.402)	
Floor X Manufacturing				0.986***
				(0.196)
Lot X Manufacturing				-0.971*
				(0.545)
Observations	460,702	460,702	460,702	460,702
MSA FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Estimation	CF	CF	CF	CF
s.e.	Bootstrap	Bootstrap	Bootstrap	Bootstrap
ATE Floors	0.0024	0.2303	0.1215	0.3549
ATE Lot	0.1302	0.1453	0.1145	0.2288

Bootstrapped standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Log wages: logarithm of wage and salary income of the householder living in the house. *Log floors* / *Floor*: Logarithm of number of stories in the same housing unit of the interviewed person. *Residential lot size MSA* / *Lot*: Logarithm of weighted sum of residential lot size of houses in one MSA-year. *Density restrictions*: Average MSA density restriction index computed by Gyourko et al. (2008) in 2008. *Coast*: dummy for MSA being a costal location. *High education*: weighted proportion of people with a BA degree or more in MSA. *Manufacturing*: proportion of jobs in manufacturing in MSA in 1998. *MSA FE*: Metropolitan Statistical Area fixed effects. *Year FE*: year fixed effects. *CF*: Control Function approach. *s.e.*: standard errors used. *ATE Floors*: Average treatment effect of *Log floors*, computed as $E\left[\frac{\partial log(w)}{\partial log(\hat{h})}\right] = \omega_{wH1} + \omega_{wH2} E[X]$, where X is the variable used for heterogeneity. *ATE Lot*: Average treatment effect of *Log Lot*, computed as $E\left[\frac{\partial log(w)}{\partial log(\hat{D})}\right] = \omega_{wD1} + \omega_{wD2} E[X]$. Regression weighted by AHS survey weights.

Table 10: House price reduced form estimation in presence of heterogeneity

	(1)	(2)	(3)	(4)
VARIABLES	Log h price	Log h price	Log h price	Log h price
Log floors	13.00***	12.69***	12.71***	12.14***
	(1.045)	(1.086)	(1.029)	(1.016)
Residential lot size MSA	0.759***	0.723***	0.658***	0.588***
	(0.0877)	(0.0971)	(0.0947)	(0.186)
Floor X Density restrictions	-0.388***			
	(0.0625)			
Lot X Density restrictions	-0.557***			
	(0.107)			
Floor X Coast		-0.383***		
		(0.0247)		
Lot X Coast		-0.179***		
		(0.0435)		
Floor X High education			-0.352**	
			(0.143)	
Lot X High education			0.207	
			(0.163)	
High education MSA			-3.542	
			(3.867)	
Floor X Manufacturing				4.200***
				(0.322)
Lot X Manufacturing				0.553
				(0.561)
Observations	226,582	226,582	226,582	226,582
MSA FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Estimation	CF	CF	CF	CF
s.e.	Bootstrap	Bootstrap	Bootstrap	Bootstrap
ATE Floors	12.90	12.50	12.59	12.91
ATE Lot	0.61	0.63	0.73	0.69

Bootstrapped standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Log h price: logarithm of current market value of house. *Log floors* / *Floor*: Logarithm of number of stories in the same housing unit of the interviewed person. *Residential lot size MSA* / *Lot*: Logarithm of weighted sum of residential lot size of houses in one MSA-year. *Density restrictions*: Average MSA density restriction index computed by Gyourko et al. (2008) in 2008. *Coast*: dummy for MSA being a costal location. *High education*: weighted proportion of people with a BA degree or more in MSA. *Manufacturing*: proportion of jobs in manufacturing in MSA in 1998. *MSA FE*: Metropolitan Statistical Area fixed effects. *Year FE*: year fixed effects. *CF*: Control Function approach. *s.e.*: standard errors used. *ATE Floors*: Average treatment effect of *Log floors*, computed as $E\left[\frac{\partial log(w)}{\partial log(\hat{h})}\right] = \omega_{wH1} + \omega_{wH2}E[X]$, where X is the variable used for heterogeneity. *ATE Lot*: Average treatment effect of *Log Lot*, computed as $E\left[\frac{\partial log(w)}{\partial log(\hat{D})}\right] = \omega_{wD1} + \omega_{wD2}E[X]$. Regression weighted by AHS survey weights.

it is possible that more educated people receive more amenities in less dense cities and then the average externality coefficient of horizontal expansion on amenities is smaller than in the case of no heterogeneity. Conversely, the ATT of the productivity effect of horizontal expansion is weaker than the ATC, that is more educated people might be less productive in less dense cities.

Finally, considering the possible heterogeneous effect of the city sectoral composition the average productivity externality of both vertical and horizontal expansions is stronger, while the ones on amenities are weaker than in the case without heterogeneity. This is consistent with the possibility that more manufacturing oriented cities are less productive but have higher amenities alone. To conclude, in next Section I estimate city-specific externality coefficients, in order to understand which cities would gain or loose the most after a vertical expansion. I also compute counterfactuals about which cities would gain the most by removing its density restrictions.

Table 11: Externality parameters in presence of heterogeneity

Heterogeneity	Coefficient	ATT	ATC	ATE
Density restrictions	$\tilde{\lambda}_{AH}$	0.095	0.102	0.100
	$\tilde{\lambda}_{AD}$	0.147	0.112	0.120
	$\tilde{\lambda}_{\theta H}$	0.272	0.265	0.267
	$\tilde{\lambda}_{\theta D}$	-0.051	0.031	0.012
Coast	$\tilde{\lambda}_{AH}$	0.109	0.159	0.112
	$\tilde{\lambda}_{AD}$	0.128	0.117	0.123
	$\tilde{\lambda}_{\theta H}$	0.256	0.252	0.254
	$\tilde{\lambda}_{\theta D}$	-0.003	0.022	0.010
Education	$\tilde{\lambda}_{AH}$	0.107	0.105	0.106
	$\tilde{\lambda}_{AD}$	0.131	0.146	0.141
	$\tilde{\lambda}_{\theta H}$	0.259	0.261	0.261
	$\tilde{\lambda}_{\theta D}$	0.003	-0.020	-0.012
Manufacturing	$\tilde{\lambda}_{AH}$	0.148	0.109	0.118
	$\tilde{\lambda}_{AD}$	0.031	0.168	0.136
	$\tilde{\lambda}_{\theta H}$	0.221	0.256	0.248
	$\tilde{\lambda}_{\theta D}$	0.157	-0.055	-0.005
No heterogeneity	$\tilde{\lambda}_{AH}$			0.103
	$\tilde{\lambda}_{AD}$			0.115
	$\tilde{\lambda}_{\theta H}$			0.261
	$\tilde{\lambda}_{\theta D}$			0.034

$\tilde{\lambda}_{AH}$: elasticity of productivity to height. $\tilde{\lambda}_{\theta H}$: elasticity of amenities to height. $\tilde{\lambda}_{AD}$: elasticity of productivity to city size. $\tilde{\lambda}_{\theta D}$: elasticity of amenities to city size. ATT: Average Treatment on the Treated (estimated using maximum possible of variable giving heterogeneity), ATC: Average Treatment on the Control (estimated using minimum possible of variable giving heterogeneity), ATE: Average Treatment Effect (estimated using average of variable giving heterogeneity).

4.3 City-specific externality coefficients

Combining all the heterogeneities together it is possibile to obtain city-specific externality coefficients that can give information about which cities are going to gain or loose the most by expanding vertically. In order to estimate city-specific externality coefficients I will estimate Models 34 and 35 introducing all the heterogeneities previously considered together. The results of this estimation are reported in Table 12. Subsequently, knowing the specific levels of density restrictions, coast location, education and manufacturing of each city I will compute separetely each λ coefficient. In Table 13 it is possible to observe that the cities which would increase productivity the most by expanding vertically are San Jose, Charlotte, Austin, Portland and Louisville. On the other side, Miami, New York, Tampa, Jacksonville and Baltimore are the MSAs which would gain the most in terms of amenities.

Table 12: House price and wage reduced form estimations in presence of heterogeneity

	(1)	(2)
VARIABLES	Log wage	Log h price
Log floors	0.0360	11.78***
	(0.875)	(1.110)
Residential lot size MSA	0.293**	0.906***
	(0.137)	(0.179)
Floor X Coast	-0.0849***	-0.310***
	(0.0177)	(0.0239)
Lot X Coast	0.0510	-0.207***
	(0.0363)	(0.0389)
Floor X Density restrictions	-0.0695	-0.102*
	(0.0432)	(0.0604)
Lot X Density restrictions	0.152*	-0.670***
	(0.0882)	(0.0945)
Floor X High education	0.151	0.235
	(0.118)	(0.176)
Lot X High education	0.0159	0.0373**
	(0.0174)	(0.0183)
Floor X Manufacturing	0.766***	3.520***
	(0.211)	(0.313)
Lot X Manufacturing	-0.611	-0.127
	(0.531)	(0.551)
Observations	460,702	226,582
MSA FE	YES	YES
Year FE	YES	YES
Estimation	CF	CF
s.e.	Bootstrap	Bootstrap
ATE Floors	0.1683	12.33
ATE Lot	0.2509	0.62

Bootstrapped standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Log wages: logarithm of wage and salary income of the householder living in the house. *Log h price*: logarithm of current market value of house. *Log floors* / *Floor*: Logarithm of number of stories in the same housing unit of the interviewed person. *Residential lot size MSA* / *Lot*: Logarithm of weighted sum of residential lot size of houses in one MSA-year. *Density restrictions*: Average MSA density restriction index computed by Gyourko et al. (2008) in 2008. *Coast*: dummy for MSA being a costal location. *High education*: weighted proportion of people with a BA degree or more in MSA. *Manufacturing*: proportion of jobs in manufacturing in MSA in 1998. *MSA FE*: Metropolitan Statistical Area fixed effects. *Year FE*: year fixed effects. *CF*: Control Function approach. *s.e.*: standard errors used. *ATE Floors*: Average treatment effect of *Log floors*, computed as $E\left[\frac{\partial log(w)}{\partial log(\widehat{h})}\right] = \omega_{wH1} + \omega_{wH2} E[X]$, where X is vector considering all the variables used for heterogeneity. *ATE Lot*: Average treatment effect of *Log Lot*, computed as $E\left[\frac{\partial log(w)}{\partial log(\widehat{D})}\right] = \omega_{wD1} + \omega_{wD2} E[X]$. Regression weighted by AHS survey weights.

The last exercise I conduct is to produce counterfactuals of what would happen if cities would decrease density restrictions. Cities with lower density restrictions tend to be more vertically oriented and less spread, as it is showed in the estimations performed in Table 14. Therefore, the first effect of decrease density restrictions is to in-

Table 13: Top 5 and bottom 5 cities for increase in productivity given increase in height

MSA	$\tilde{\lambda}_{AH}$	MSA	$\tilde{\lambda}_{\theta H}$
San Jose, CA	0.117	Miami Hialeah, FL	0.262
Charlotte, NC	0.113	New York Areas	0.262
Austin, TX	0.113	Tampa St. Petersburg Clearwater, FL	0.261
Portland, OR - WA	0.113	Jacksonville, FL	0.261
Louisville, KY-IN	0.112	Baltimore, MD	0.260
Baltimore, MD	0.105	Louisville, KY-IN	0.253
Jacksonville, FL	0.105	Portland, OR - WA	0.253
Tampa St. Petersburg Clearwater, FL	0.104	Charlotte, NC	0.253
New York Areas	0.103	Austin, TX	0.253
Miami Hialeah, FL	0.103	San Jose, CA	0.249

Elasticities computed using the estimation from Table 12. $\tilde{\lambda}_{AH}$: elasticity of productivity to height. $\tilde{\lambda}_{\theta H}$: elasticity of amenities to height. $\tilde{\lambda}_{AD}$: elasticity of productivity to city size.

crease height of the city, which consequently influences wages, house prices, amenities and productivity.

Using Model 35 for every city it is possible to compute the effect of reducing density restrictions (DRI) on wages (and similarly on house prices) via its effect on height (and similarly on city spread) as $\frac{\partial log(w)}{\partial DRI} = \omega_{wH1}\frac{\partial log(h)}{\partial DRI} + \omega_{wH2}\frac{\partial log(h)}{\partial DRI} X_m + \omega_{wH2}\frac{\partial X_m}{\partial DRI} log\,(h)$[10]. From Table 15 it evinces that New York, Chicago, Washington, Minneapolis and Philadelphia are the cities which would increase productivity the most by reducing density restrictions.

Table 14: OLS estimation of the relationship between density restrictions and vertical and horizontal expansion

VARIABLES	(1) Log floors	(2) Log lot MSA
Density Restrictions Index	-0.150***	0.360***
	(0.0113)	(0.00826)
Observations	372,953	374,149
R-squared	0.250	0.777
State FE	YES	YES
Year FE	YES	YES
Estimation	OLS	OLS

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Log floors : Logarithm of number of stories in the same housing unit of the interviewed person. *Log lot MSA*: Logarithm of weighted sum of residential lot size of houses in one MSA-year. *Density restrictions Index*: Average MSA density restriction index computed by Gyourko et al. (2008) in 2008. *State FE*: State fixed effects. *Year FE*: year fixed effects. Regression weighted by AHS survey weights.

[10] Defining X as the vector of all covariates, which include also the density restrictions, I assume that $\frac{\partial X_m}{\partial DRI} = 0$ for any X different from DRI, while $\frac{\partial DRI}{\partial DRI} = 1$

Table 15: Top 10 cities for increase in productivity given reduction of density restrictions

MSA	$\tilde{\lambda}_{AH}$	Average DRI
New York Areas	0.196	0.35
Chicago Areas	0.178	0.13
Washington, DC MD VA	0.177	0.07
Minneapolis St. Paul, MN	0.175	0.12
Philadelphia, PA NJ	0.174	0.58
Boston, MA	0.174	0.61
Cleveland, OH	0.174	0.16
Pittsburgh, PA	0.172	0.27
Denver, CO	0.171	0.22
Milwaukee, WI	0.170	0.18

$\tilde{\lambda}_{AH}$: elasticity of productivity to height. Computed using the estimation in Table 14. *DRI*: Average MSA density restriction index computed by Gyourko et al. (2008) in 2008.

5 Concluding remarks

In this paper I study how tall cities and more spread cities are different with respect to different city outcomes. I have obtained elasticity with respect to height of buildings and the horizontal spread of cities of observable outcomes, such as wages, house prices and housing density, but also of theoretical concepts, such as productivity and amenities. Credible reduced form estimates of the effect on wages, house prices and housing density have been obtained using instrumental variable techniques, exploiting exogeneous variation coming from geological, geographical and technological factors. Subsequently, I build a spatial equilibrium model which encompasses and microfounds the role of height and spread of cities to obtain theoretical predictions of their effects but also give a strategy to map reduced form estimates into conclusions about productivity and amenities.

It has been concluded that both both taller cities and more spread cities are associated with higher levels of house prices and housing density, but these elasticities are much stronger for cities with higher height. No significant difference has been found with respect to wages, but I show that this result derives from the different indirect effects via productivity and amenities. In fact, I estimate that both height and spread of cities are expected to generate positive externalities to productivity and amenities. Since the theoretical effect of productivity on wages is positive and the theoretical effect of amenities on wages is negative, this rationalize the estimated elasticity of wages. I have found that height and spread of cities are predicted to increase productivity in similar way, but height is expected to increase amenities more substantially.

From my results it is possible to argue that is reasonable to use size and density interchangeably while speaking of agglomeration economies in term of productivity. However, from revealed preferences it seems that agents obtain higher amenities by living in more dense cities than in more spread, given the population of the city. Because of important heterogeneity concerns, the gains of expanding vertically or horizontally are different for U.S. cities. In this paper I have presented which cities would have gain the most by expanding vertically and reducing density restrictions.

References

Ahlfeldt, G. M. and D. McMillen (2017). Tall buildings and land values: height and construction cost elasticities in chicago, 1870–2010. *Review of Economics and Statistics* (0).

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015). The economics of density: Evidence from the berlin wall. *Econometrica 83*(6), 2127–2189.

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.

Alonso, W. (1964). Location and land use. toward a general theory of land rent. *Location and land use. Toward a general theory of land rent.*.

Angrist, J. D., J.-S. Pischke, and J.-S. Pischke (2009). *Mostly harmless econometrics: an empiricist's companion.* Princeton university press Princeton.

Bartik, T. J. (1991). Who benefits from state and local economic development policies? *Books from Upjohn Press.*

Bernard, A. (2014). *Lifted: A Cultural History of the Elevator.* New York University Press.

Bertaud, A. and J. K. Brueckner (2005). Analyzing building-height restrictions: predicted impacts and welfare costs. *Regional Science and Urban Economics 35*(2), 109–125.

Borck, R. (2014). Will Skyscrapers Save the Planet? Building Height Limits and Urban Greenhouse Gas Emissions. CESifo Working Paper Series 4773, CESifo Group Munich.

Brueckner, J. K., S. Fu, Y. Gu, and J. Zhang (2015). Measuring the stringency of land-use regulation and its determinants: The case of china's building-height limits. Technical report, Working paper, May 19 2015.

Brueckner, J. K. and K. S. Sridhar (2012). Measuring welfare gains from relaxation of land-use restrictions: The case of india's building-height limits. *Regional Science and Urban Economics 42*(6), 1061–1067.

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). 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.*

Cragg, J. G. and S. G. Donald (1993). Testing identifiability and specification in instrumental variable models. *Econometric Theory 9*(02), 222–240.

Curci, F. (2018). The taller the better? agglomeration determinants and urban structure. mimeo.

Diamond, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980–2000. *The American Economic Review 106*(3), 479–524.

Duranton, G. and M. A. Turner (2018). Urban form and driving: Evidence from us cities. *Journal of Urban Economics*.

Glaeser, E. L. (2008). *Cities, Agglomeration and Spatial Equilibrium*. New York: Oxford University Press.

Glaeser, E. L. and J. D. Gottlieb (2009). The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States. *Journal of Economic Literature 47*(4), 983–1028.

Gyourko, J., A. Saiz, and A. Summers (2008). A new measure of the local regulatory environment for housing markets: The wharton residential land use regulatory index. *Urban Studies 45*(3), 693–729.

Harari, M. (2015). Cities in bad shape: Urban geometry in india. Job market paper, Massachusetts Institute of Technology.

INSDAG (2013). Multi-storey buildings. Technical report, Institute for Steel Development and Growth.

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.

Liu, C. H., S. S. Rosenthal, and W. C. Strange (2018). The vertical city: Rent gradients, spatial structure, and agglomeration economies. *Journal of Urban Economics*.

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.

Mills, E. S. (1972). Studies in the structure of the urban economy.

Muth, R. (1969). Cities and housing: The spatial patterns of urban residential land use. *University of Chicago, Chicago*.

Roback, J. (1982). Wages, rents, and the quality of life. *The Journal of Political Economy*, 1257–1278.

Rosen, S. (1979). Wage-based indexes of urban quality of life. *Current issues in urban economics 3*.

Saiz, A. (2010). The geographic determinants of housing supply. *Quarterly Journal of Economics 125*(3).

Stock, J. H. and M. Yogo (2005). Testing for weak instruments in linear iv regression. *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*.

U.N. (2014). World urbanization prospects: the 2014 revision. Technical report, United Nations.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources 50*(2), 420–445.

A Robustness of the first stage

In this Section I report robustness checks for the first stage estimation performed in Section 3.2. Results for the first stage of the vertical expansion of the city are reported in Table 16, while results for the horizontal expansion are reported in Table 17. Two main concerns can arise about the robustness of my instrument: inference and unobserved heterogeneity at higher level than MSA. Firstly, residuals in the first stage regression can be correlated at metropolitan area because houses characteristics can be similar. This correlation would create inference problem because standard errors would be biased. In order to obtain more consistent standard errors I have clustered standard errors at MSA (see column (2) of Tables 16 and 17) and at MSA-Year level (see column (3) of Tables 16 and 17). Results about significance of my instruments do not change with respect to my baseline first stage (column (1) of Tables 16 and 17).

A second concern arises about the variation I am exploiting to obtain the causal effects of interest. In particular, the identification strategy used in this paper relies upon variation arising from geological and geographical variables. MSAs close to sea or major water sources might be fundamentally different from MSAs in the other parts of the U.S. and this can potentially invalidate the results obtained. In order to control for this potential cofounding factors column (4), (5) and (6) of Tables 16 and 17 control for different kind of fixed effects (FE): state FE, U.S. Census Region FE and Region times Year FE [11]. My first stage results are robust across all these specifications.

I conduct a final robustness check about my first stage estimation by conducting a placebo experiment on the instruments. This experiment consists in assigning randomly values of the instruments of one MSA to another MSA. For each instrument (earthquake risk, undevelopable area, Bartik shifter, and elevator technology) and for every MSA I therefore draw one other MSA and I assign to the original MSA the instrument value of the drawn MSA. I then create the interactions between earthquake risk and elevator technology and between Bartik shifter and undevelopable area, and I subsequently estimate the two first stage equations. I repeat this procedure 200 times. Since the instruments are randomly draw I need to reject the null hypothesis of non-significance of the instrument only 5 % of the times. Moreover, the distribution of the placebo coefficients should be skewed away from my real first stage coefficients.

In Figures 6(a) and 6(b) I report the distribution of the placebo coefficients of interaction between earthquake risk and elevator technology in the vertical expansion first stage and the one of the interaction between Bartik shifter and undevelopable area in the horizontal expansion first stage, respectively. As it is possible to see from the Figures my real first stage coefficients happen to be outliers in the distribution of

[11] 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.

Table 16: Robustness checks for vertical expansion first stage

VARIABLES	(1) Log floors	(2) Log floors	(3) Log floors	(4) Log floors	(5) Log floors	(6) Log floors
Hazard / elevators	-0.0434***	-0.0434**	-0.0434***	-0.0430***	-0.0430***	-0.0452***
	(0.0124)	(0.0173)	(0.0126)	(0.0123)	(0.0123)	(0.0151)
Bartik / undevelopable	-0.00168	-0.00168*	-0.00168**	0.0115	0.0115	0.0103
	(0.00103)	(0.000849)	(0.000719)	(0.0127)	(0.0127)	(0.0173)
Observations	370,019	370,019	370,019	362,280	362,280	362,280
R-squared	0.284	0.284	0.284	0.283	0.283	0.283
MSA FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
State FE	NO	NO	NO	YES	NO	NO
Region FE	NO	NO	NO	NO	YES	NO
RegionYear FE	NO	NO	NO	NO	NO	YES
Estimation	OLS	OLS	OLS	OLS	OLS	OLS
s.e.	Robust	Cl MSA	Cl MSAYear	Robust	Robust	Robust

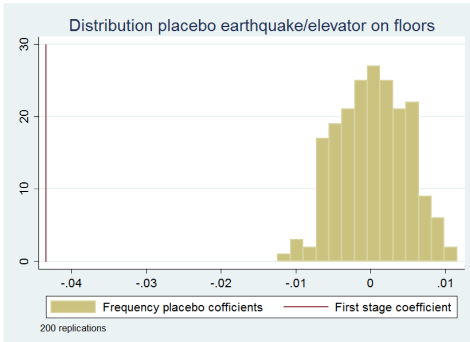
Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Log floors: Logarithm of number of stories in the same housing unit of the interviewed person. *hazard*: Weighted average 2 percent probability of exceedance in 50 years of mean peak ground acceleration for each MSA. *elevators*: weighted average of houses with elevators in U.S. in one year. *undevelopable*: Percent of MSA area which is undevelopable because of water bodies or steepness. *Bartik*: Bartik labour demand shock, constructed as in Equation 26. All explanatory variables are in logarithms. *MSA FE*: Metropolitan Statistical Area fixed effects. Year FE: year fixed effects. *State FE*: State fixed effects. *Region FE*: Census Regions fixed effects. *RegionYear FE*: Census region × year fixed effects. *s.e.*: standard errors used. *Cl*: clustered standard errors. Regression weighted by AHS survey weights.

Table 17: Robustness checks for horizontal expansion first stage

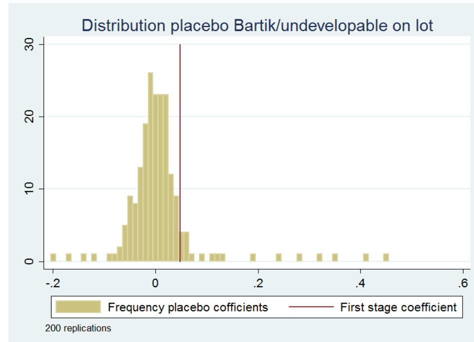
VARIABLES	(1) Log lot	(2) Log lot	(3) Log lot	(4) Log lot	(5) Log lot	(6) Log lot
Hazard / elevators	0.0642***	0.0642	0.0642	0.0679***	0.0679***	0.345***
	(0.00316)	(0.190)	(0.148)	(0.00319)	(0.00319)	(0.00293)
Bartik / undevelopable	0.0485***	0.0485***	0.0485***	0.161***	0.161***	0.188***
	(0.000192)	(0.0143)	(0.0110)	(0.00307)	(0.00307)	(0.00358)
Observations	371,104	371,104	371,104	363,365	363,365	363,365
R-squared	0.942	0.942	0.942	0.941	0.941	0.973
MSA FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
State FE	NO	NO	NO	YES	NO	NO
Region FE	NO	NO	NO	NO	YES	NO
RegionYear FE	NO	NO	NO	NO	NO	YES
Estimation	OLS	OLS	OLS	OLS	OLS	OLS
s.e.	Robust	Cl MSA	Cl MSAYear	Robust	Robust	Robust

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Log lot: Logarithm of weighted sum of residential lot size of houses in one MSA-year. *hazard*: Weighted average 2 percent probability of exceedance in 50 years of mean peak ground acceleration for each MSA. *elevators*: weighted average of houses with elevators in U.S. in one year. *undevelopable*: Percent of MSA area which is undevelopable because of water bodies or steepness. *Bartik*: Bartik labour demand shock, constructed as in Equation 26. All explanatory variables are in logarithms. *MSA FE*: Metropolitan Statistical Area fixed effects. Year FE: year fixed effects. *State FE*: State fixed effects. *Region FE*: Census Regions fixed effects. *RegionYear FE*: Census region × year fixed effects. *s.e.*: standard errors used. *Cl*: clustered standard errors. Regression weighted by AHS survey weights.

the placebo coefficients.



(a) Vertical expansion first stage



(b) Horizontal expansion first stage

Figure 6: Distribution of placebo coefficients

Panel a): Distribution of placebo coefficients of interaction between earthquake risk and elevator technology in the vertical expansion first stage. *Log floors*: Logarithm of number of stories in the same housing unit of the interviewed person. *earthquake*: Weighted average 2 percent probability of exceedance in 50 years of mean peak ground acceleration for each MSA. *elevators*: weighted average of houses with elevators in U.S. in one year. In every replication the value for the earthquake risk and proportion of elevators for one MSA have been randomly reassigned to another MSA. The Figure reports the distribution of the coefficients of regressing log floors on the placebo earthquake / elevator controlling for MSA and year fixed effects for every replication.

Panel b): Distribution of placebo coefficients of interaction between Bartik shifter and undevelopable area in the horizontal expansion first stage. *Log lot*: Logarithm of weighted sum of residential lot size of houses in one MSA-year. *undevelopable*: Percent of MSA area which is undevelopable because of water bodies or steepness. *Bartik*: Bartik labour demand shock, constructed as in Equation 26. In every replication the value for the Bartik shock and proportion of undevelopable area for one MSA have been randomly reassigned to another MSA. The Figure reports the distribution of the coefficients of regressing log lot on the placebo Bartik / undevelopable controlling for MSA and year fixed effects for every replication.

All explanatory variables are in logarithms. *First stage coefficient* reports the real, non-placebo, first stage coefficient. 200 replications. Regression weighted by AHS survey weights.

B Reverse causality bias in the OLS estimation of house prices

In this section I discuss why reverse causality in the house price equation estimation is in line with the results obtained in Section 4. The model used in this paper predicts that increase in height can influence house prices via the effect they have on housing supply (see Equation 4) but also that higher house prices can stimulate higher height (see Equation 3), a result in line with Ahlfeldt et al. (2015). The relationship running from height to house prices can be written as in Equation 36, while the relationship from house prices to height can be written as in Equation 37.

$$log\left(p_H\right) = \xi_1 log\left(h\right) + \epsilon_1 \tag{36}$$

$$log\left(h\right) = \xi_2 log\left(p_H\right) + \epsilon_2 \tag{37}$$

I am interested in the relationship running from height to house prices. Therefore, the model that is estimated by OLS can be written as in Equation 38, by combining Equations 36 and 37. OLS estimates will be biased because of the simultaneity expressed in Equations 36. The OLS bias is equal to $\frac{\xi_1-1}{1-\xi_2} - \xi_1$. OLS estimate will be downward bias with respect to IV if the condition in Equation 39 is met.

$$\text{OLS: } log\left(p_H\right) = \frac{\xi_1 - 1}{1-\xi_2} log\left(h\right) + \frac{1}{1-\xi_2}\left(\epsilon_1 + \epsilon_2\right) \tag{38}$$

$$\text{Downward bias OLS: } \frac{\xi_1 - 1}{1-\xi_2} < \xi_1 \rightarrow \xi_1 < \frac{1}{\xi_2} \tag{39}$$

Equation 3 of my model gives the prediction that the condition for downward bias is satisfied when the IV estimate of the effect of height on house price is lower than the estimated value of the convexity of height construction, δ-1. Then the downward bias condition can be written as $\xi_1 < \delta - 1$ and this condition is satisfied since the estimated elasticity of house prices with respect to height was around 11 and δ was estimated to be around 24. Moreover, using Equation 3 we can conclude that $\xi_2 = \frac{1}{\delta-1}$ and this is coherent with the OLS estimation of Equation 37 reported in Table 18. In fact, the elasticity of floors to house prices, ξ_2, is 0.041, which should be equal to $\frac{1}{\delta-1} = \frac{1}{24-1} = 0.043$.

Table 18: OLS estimation effect house prices on number of floors

	(1)
VARIABLES	Log floors
Log current market value unit	0.0410***
	(0.00156)
Observations	238,096
R-squared	0.304
MSA FE	YES
Year FE	YES
Estimation	OLS

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Log of current market value of house is the proxy for house price. *Log floors*: Logarithm of number of stories in the same housing unit of the interviewed person. *MSA FE*: Metropolitan Statistical Area fixed effects. *Year FE*: year fixed effects. Regression weighted by AHS survey weights.

C Robustness of reduced form estimation

In this section I report robustness checks for the reduced form estimations performed in Section 4. Similar to the robustness checks conducted about the first stage I check that the results are robust using clustered standard errors and fixed effects at higher level of geographical aggregation. Tables 19 and 20 reports the robustness checks for the reduced form estimations of wages and house prices, respectively. Column (1) of these tables reports the baseline IV estimation. Column (2) and (3) use clustered standard errors at MSA and MSA times Year level. Column (7), (8) and (9) control for different kind of fixed effects: state FE, U.S. Census Region FE and Region times Year FE.

An additional robustness I perform is to include individual controls about the householder and the house quality. In particular, column (4) of Table 19 controls for individual characteristics of the householder that can influence its own wages: dummies for education levels, age, race and sex. Similarly, column (4) of Table 20 controls for individual characteristics of the house quality that can influence its own house price: number of baths and bedrooms, year in which the structure was built, and the square footage of the unit.

The variation in the technology improvement in the elevator sector is partly determined by the higher demand for elevators caused by aging population. Aging population can potentially influence overall wages, therefore column (5) of Table 19 controls for the mean age in the MSA of the householder. Similarly, access to elevators can potentially be higher in places with more recent housing constructions, which can sequentially influences house prices. Column (5) of Table 20 reports the robustness checks of the house prices estimation by controlling for the mean age of houses in the MSA.

Finally, column (6) of Tables 19 and 20 re-estimate the reduced form model in the subsample of observations which are not in New York, Chicago and Northern New Jersey. This is done because the definition of these cities derived from the National sample American Housing Survey is different from the one in the Metropolitan sample.

As it is possibly to see in Tables 19 and 20, the results obtained are robust to a variety of different robustness checks. The magnitude, direction and significance of the coefficients is maintained in the majority of the specification considered. An only exception is the significance of the height effect on house prices controlling for region times year fixed effects. However, conclusions about the externality parameters are not changed even in this specification.

Table 19: Robustness checks for wage reduced form estimation

VARIABLES	(1) Log wages	(2) Log wages	(3) Log wages	(4) Log wages	(5) Log wages	(6) Log wages	(7) Log wages	(8) Log wages	(9) Log wages
Log floors	0.0606	0.0606	0.0606	0.0357	0.312	0.989	-0.0109	-0.0109	-1.560
	(0.734)	(1.850)	(1.458)	(0.691)	(0.770)	(0.684)	(0.682)	(0.682)	(1.314)
Log lot	0.106	0.106	0.106	0.0894	0.236**	0.219***	0.0464	0.0464	-0.0860
	(0.0825)	(0.164)	(0.126)	(0.0663)	(0.118)	(0.0658)	(0.195)	(0.195)	(0.203)
Observations	220,399	220,399	220,399	220,399	220,399	203,024	215,588	215,588	215,588
MSA FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
State FE	NO	NO	NO	NO	NO	NO	YES	NO	NO
Region FE	NO	NO	NO	NO	NO	NO	NO	YES	NO
RegionYear FE	NO	NO	NO	NO	NO	NO	NO	NO	YES
Individual controls	NO	NO	NO	YES	NO	NO	NO	NO	NO
Aggregate age	NO	NO	NO	NO	YES	NO	NO	NO	NO
Sample	All	All	All	All	All	No big 3	All	All	All
Estimation	IV	IV	IV	IV	IV	IV	IV	IV	IV
s.e.	Robust	Cl MSA	Cl MSAYear	Robust	Robust	Robust	Robust	Robust	Robust

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Log wages: logarithm of wage and salary income of the householder living in the house. *Log floors*: Logarithm of number of stories in the same housing unit of the interviewed person. *Log lot*: Logarithm of weighted sum of residential lot size of houses in one MSA-year. *MSA FE*: Metropolitan Statistical Area fixed effects. *Year FE*: year fixed effects. *Region FE*: Census Regions fixed effects. *RegionYear FE*: Census region × year fixed effects. *Individual controls*: dummies for education levels, age, race and sex. *Aggregate age*: weighted mean age in the MSA of the interviewed. *No big 3*: Sample without New York, Chicago and Northern New Jersey. *s.e.*: standard errors used. *Cl*: clustered standard errors. Regression weighted by AHS survey weights.

Table 20: Robustness checks for house price reduced form estimation

VARIABLES	(1) Log price	(2) Log price	(3) Log price	(4) Log price	(5) Log price	(6) Log price	(7) Log price	(8) Log price	(9) Log price
Log floors	11.33***	11.33*	11.33**	12.15**	10.01***	14.13***	11.62***	11.62***	17.95
	(4.062)	(5.795)	(4.862)	(5.235)	(3.421)	(5.447)	(4.177)	(4.177)	(15.05)
Log lot	0.909**	0.909	0.909*	0.912*	0.880**	1.004**	2.861***	2.861***	2.119
	(0.458)	(0.602)	(0.465)	(0.507)	(0.399)	(0.433)	(1.079)	(1.079)	(2.408)
Observations	139,465	139,465	139,465	124,572	139,465	129,643	136,202	136,202	136,202
MSA FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
State FE	NO	NO	NO	NO	NO	NO	YES	NO	NO
Region FE	NO	NO	NO	NO	NO	NO	NO	YES	NO
RegionYear FE	NO	NO	NO	NO	NO	NO	NO	NO	YES
Individual controls	NO	NO	NO	YES	NO	NO	NO	NO	NO
Aggregate age	NO	NO	NO	NO	YES	NO	NO	NO	NO
Sample	All	All	All	All	All	No big 3	All	All	All
Estimation	IV	IV	IV	IV	IV	IV	IV	IV	IV
s.e.	Robust	Cl MSA	Cl MSAYear	Robust	Robust	Robust	Robust	Robust	Robust

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Log price: logarithm of current market value of house. *Log floors*: Logarithm of number of stories in the same housing unit of the interviewed person. *Log lot*: Logarithm of weighted sum of residential lot size of houses in one MSA-year. *MSA FE*: Metropolitan Statistical Area fixed effects. *Year FE*: year fixed effects. *Region FE*: Census Regions fixed effects. *RegionYear FE*: Census region × year fixed effects. *Individual controls*:number of baths and bedrooms, year in which the structure was built, and the square footage of the unit. *Aggregate age*: weighted mean age in the MSA of the interviewed. *No big 3*: Sample without New York, Chicago and Northern New Jersey. *s.e.*: standard errors used. *Cl*: clustered standard errors. Regression weighted by AHS survey weights.

D Additional results: housing density

An additional reduced form equation that is possible to estimate using my database is the effect of city development on housing density, measured by the number of units in one building. If height and city size increases population, then their are expected to increase housing density. However, if these two variables lead to a decrease in amenities housing density is expected to decrease. A similar result is possible to obtain if city development does not change population but only lead to cities with bigger size and lower units per building. Table 21 suggests that the effect of both height and city size on housing density is positive and then it gives further evidence to the possibility that these variables increased amenities.

Table 21: Reduced form estimates for housing density

	(1)	(2)	(3)	(4)
VARIABLES	Log h density	Log h density	Log h density	Log h density
Log floors	1.470***	1.685***	1.610**	
	(0.00957)	(0.00967)	(0.697)	
Residential lot size MSA	-0.215***	0.0468	0.158**	
	(0.00482)	(0.0287)	(0.0742)	
Log hazard / Proportion elevators				-0.0749*
				(0.0416)
Log Bartik / undevelopable area				0.00338
				(0.00290)
Observations	226,579	226,579	220,396	220,396
R-squared	0.340	0.429	.	0.103
MSA FE	NO	YES	YES	YES
Year FE	NO	YES	YES	YES
Estimation	OLS	OLS	IV	OLS

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Log h density: logarithm number of units in the building of interview person. *Log floors*: Logarithm of number of stories in the same housing unit of the interviewed person. *Residential lot size MSA*: Logarithm of weighted sum of residential lot size of houses in one MSA-year. *hazard*: Weighted average 2 percent probability of exceedance in 50 years of mean peak ground acceleration for each MSA. *proportion elevators*: weighted average of houses with elevators in U.S. in one year. *undevelopable area*: Percent of MSA area which is undevelopable because of water bodies or steepness. *Bartik*: Bartik labour demand shock, constructed as in Equation 26. *MSA FE*: Metropolitan Statistical Area fixed effects. *Year FE*: year fixed effects. Regression weighted by AHS survey weights.

Flight from Urban Blight:
Lead Poisoning, Crime and Suburbanization

Federico Curci* and Federico Masera†‡

Abstract

In the post World War II period, most U.S. cities experienced large movements of population from the city centers to the suburbs. In this paper we provide causal evidence that this process of suburbanization can be explained by the rise of violent crime in city centers. We do so by proposing a new instrument to exogenously predict violent crime. This instrument uses as time variation the U.S. national levels of lead poisoning. Cross-sectional variation comes from a proxy for soil quality, which explains the fate of lead in soil and its subsequent bioavailability. Using data for more than 300 U.S. cities, results show that the increase in violent crime from the level in 1960 to its maximum in 1991 decreased the proportion of people living in city centers by 15 percentage points. This increase in crime moved almost 25 million people to the suburbs. As a result of suburbanization, we find that people remaining in the city center are more likely to be black people, consistent with the "white flight" phenomenon. We then demonstrate that this suburbanization process had aggregate effects on the city. Exploiting a spatial equilibrium model, we determine that violent crime had externalities on productivity and amenities.

*Universidad Carlos III de Madrid. E-mail: fcurci@uc3m.es

†University of New South Wales. E-mail: f.masera@unsw.edu.au

‡Curci thanks Ricardo Mora for the excellent supervision. We thank Luca Armilli, Nathaniel Baum-Snow, Stephen Billings, Julio Cáceres, Irma Clots-Figueras, Klaus Desmet, Jesús Fernández-Huertas, Pauline Grosjean, Christian Hilber, Adelheid Holl, Matthias Kredler, Salvatore Lo Bello, Olivier Marie, Florian Mayneris, Jaime Millán, Luigi Minale, Peter Nillson, William Parienté, Nishith Prakash, Michele Rosenberg, Kevin Schnepel, William Strange, Jan Stuhler, Bruno Van der Linden, Gonzague Vannorenberghe, and Hasin Yousaf for relevant inputs and comments. We also thank participants to the UC3M reading group, UC3M PhD workshop, 17th SAET Conference on Current Trends in Economics, 2017 Barcelona Workshop on Regional and Urban Economics, 42nd Simposio de la Asociación Española de Economía, 2017 Econometric Society European Winter Meeting, 2017 Asian and Australasian Society of Labour Economics Conference, 5th IEB Workshop on Urban Economics, 8th ifo Dresden Workshop on Regional Economics and seminars at Université Catholique de Louvain, University of New South Wales, Stockholm University, Università Commerciale Bocconi, Navarra Center for International Development, Universidad Autónoma de Madrid, CUNEF, CaixaBank Research and Universitat de Barcelona/IEB. Financial support from the Spanish Ministry of Economics and Competitiveness (ECO2012-33053) is acknowledged.

1 Introduction

In the last century, both developed and developing countries experienced at the same time two important urban phenomena: urbanization and suburbanization. Urbanization refers to the movement of people from rural to urban areas. Suburbanization represents the movement of population from city centers to low density suburban areas. The increase in the number of people living in the suburbs is not just caused by city growth, as shown by Angel et al. (2011). U.S. cities provide an emblematic example of this suburbanization process. According to Baum-Snow (2007), U.S. population living in city centers declined by 17 percent between 1950 and 1990 despite population in urban areas increased by 72 percent.

The advantages and drawbacks of city growth have been largely studied. Urbanization reflects agglomeration economies and higher productivities. At the same time, large cities might suffer congestions and urban distress. In particular, crime is higher in bigger cities (Glaeser and Sacerdote, 1999). The movement of people from city centers to suburbs can underline the negative effects of density. In this paper we provide novel causal evidence for a mechanism that links the increase in violent crimes in U.S. city centers between the 1960s and the 1990s to suburbanization. We then show the consequences of this reallocation of people within cities in terms of racial segregation, and overall city productivity and amenities. Suburbanization of people has implications in terms of congestions and transport costs, decreasing amenities in cities. It also affects location of firms and then city productivity. Therefore, in this paper we show that suburbanization is crucial to explain how city structure can influence productivity and amenities externalities offered by cities, something that has received the attention of a limited number of studies.

While in 1960, 43% of the urban population in the U.S. was living in city centers, this proportion dropped to 33% in 1990. In this paper, we argue that the amenity value of city centers in the U.S. decreased because of crime, leading people to suburbanize. In fact, U.S. cities experienced a dramatic increase in violent crimes at the same time that population suburbanized (see Figure 1(a)). The violent crimes rate rose from 23 crimes per 10,000 inhabitants in 1960 to 163 crimes per 10,000 inhabitants in 1991. Similarly, cities in which violent crime increased the most had the strongest decrease in proportion of people living in city centers (see Figure 1(b)). When crime rates decreased, after 1991, the general trend for suburbanization did not revert.[1]

[INSERT FIGURE 1 HERE]

The goal of this research is to provide causal evidence and to quantify the effect of crime on suburbanization. We do so by proposing a new instrument to exogenously predict violent crime rate in the city centers of all U.S. cities. The time variation of our instrument is provided by U.S. national levels of lead consumption. Medical literature recognizes that exposure to lead as a child alters the formation of the brain and increases aggressive behaviour in adulthood. We exploit the specific timing of the effect of lead on crime to be sure that we are not capturing the effect that lead might have on other outcomes. Lead emissions by cars in the U.S. increased dramatically

[1] There are some recent studies that provide evidence of the return of some categories of population to city centers, in particular white people with college degree (see Baum-Snow and Hartley, 2016 for a review). However, these individuals represent a small proportion of the U.S. population.

until 1972, and 19 years later crime rates reached their peak. Given that our identification strategy takes advantage of this increase in lead exposure, we concentrate our analysis on the period between 1960 and 1991. Lead emissions by cars accumulate in the soil and then can be ingested by humans via soil resuspension. We obtain geographical variation for our instrument by exploiting the chemical literature evidence showing that lead bioavailability to humans increases when lead deposits in soils of a particular pH level. We use this information to instrument violent crime rates using the interaction between the lagged national level of tetraethyl lead used in cars and a function of the pH of the soil. We construct this instrument using machine learning techniques as described in Section 4.2.

We estimate the effect of violent crime on suburbanization using a newly assembled database for more than 300 U.S. cities. Our results show that the increase in violent crimes from the level of 1960 to their maximum in 1991 was responsible for a decrease in the proportion of people living in city centers by 15 percentage points. Increases in violent crime led more than 25 million people to leave city centers. However, we encounter that the increase in crime rate did not change the total city population. Higher crime rates in the city centers drove people to relocate within cities but not between cities. We also find that this suburbanization process was associated with the so-called "white flight". As a city center became more violent, white people moved to the suburbs, leading to an increase in the percentage of black people in the city center. Moreover, we provide evidence that the increase in violent crime was not only responsible for residential suburbanization but also it induced firms to leave city centers and locate in the outskirts of the city. Employment decentralization followed residential suburbanization and not the opposite. These results are confirmed after several robustness and specification tests.

After showing that violent crime had an effect on the distribution of people and firms inside a city we explore whether this phenomenon generated aggregate effects at the city level. In addition to finding that violent crime did not decrease the overall population of the city, we prove that violent crime increased both house prices and median incomes. To rationalize these city-aggregate results, we estimate a spatial equilibrium model, based on Glaeser (2008), in which people decide in which city to locate and in equilibrium utility should be equalized between cities. We assume that violent crime can affect city amenities and productivity. We exploit the model to map reduced form elasticities of the effect of violent crime on house prices, wages and total city population, to structural parameters that describe the effect that violent crime has on city amenities and productivity. We find that higher violent crime rates and the consequent relocation of people inside a city decrease city amenities but increase overall city productivity. Our structural estimates imply that the increase in violent crime in the city center between 1960 and 1991 led to a decrease in the average amenities of the city of 23.2%. We provide suggestive evidence that the effect of crime on productivity is entirely caused by the effect that crime has on employment decentralization.

This paper first contributes to the literature about the determinants of suburbanization. Several reasons have been identified as contributors of the suburbanization of U.S. cities. Causal evidence of the effect of highways on suburbanization has been provided by Baum-Snow (2007).[2] Brueckner and Rosenthal (2009) argue that suburbanization is explained by the fact that high-income households have higher demand for newer housing stock, which develops faster in suburban locations. Boustan (2010) shows that the large migration of black population from the rural South of the U.S. led to whites leaving the cities. Reber (2005) provides evidence that white flight has been stronger in districts with court-ordered desegregation plans.[3] Several early studies have more generally related urban blight to the flight from city centers.[4]

We show that the increase in crime rates from the 1960s to the 1990s is an important reason for U.S. suburbanization. According to our estimates, the relative increase in crime between city centers and suburbs from 1960 to 1991 implies a 35% decrease in the population of city centers. We can compare this result to similar numbers in the literature. Baum-Snow (2007) provides evidence that the construction of the interstate highway system reduced the population of city centers by 23%. The effect of the great black migration has been estimated to cause a drop in 17% in city center population (Boustan, 2010). In this paper, we demonstrate that the effect of violent crime has important complementarities with these other mechanisms. In fact, we show that the increase in violent crime rates increased the construction of highways and decreased the white population in the city centers, which consecutively further influenced suburbanization. We also find that the suburbanization caused by crime is stronger in cities with more blacks in the city center and in cities with more highways.

The link between crime and relocation of people has been the object of study of a limited amount of analyses. Cullen and Levitt (1999) were the first to causally consider the relationship between crime rates and city population. Their empirical strategy consists of analyzing the effect on a city center population of crime rates instrumented by the lagged changes in the punitiveness of the state criminal justice system and controlling for city and year fixed effects and several city characteristics. They conclude that the decrease in city population because of increased crime rates is mainly due to people migrating out of the city center. Our work differs from Cullen and Levitt (1999) because we exploit exogenous variation at a much finer geographical level of observations that do not correlate with any potential suburbanization confounding mechanism at city level.

Our findings relate to the growing literature on optimal city structure. These studies rely on the classical urban models developed by Alonso (1964), Mills (1972), and Muth (1969), and subsequently expanded by Fujita and Ogawa (1982), Lucas and Rossi-Hansberg (2002), Ahlfeldt et al. (2015), and Allen et al. (2015). Our work also connects to studies about how urban amenities change in response to urban shape and crime. The cornerstone of these works is the spatial equilibrium concept introduced by the seminal works by Rosen (1979) and Roback (1982), and then reviewed by Glaeser (2008) and Moretti (2011). Similar to our work, Harari (2015) estimates the externality effects of urban compactness. Diamond (2016) shows that crime is an

[2] Similar evidence has been found for the case of Spain (Garcia-Lopez et al., 2015) and other European countries (Garcia-Lopez et al., 2015). Similarly, Glaeser and Kahn (2004) and Kopecky and Suen (2010) relate suburbanization to car adoption.

[3] This last result was not confirmed by Baum-Snow and Lutz (2011) who find that school desegregation affected only out-of-city migration and not within-city suburbanization.

[4] See, for example, Bradford and Kelejian (1973), Frey (1979), Grubb (1982), Mieszkowski and Mills (1993), and Mills and Lubuele (1997). Cullen and Levitt (1999) look specifically at the role of crime.

important component of urban amenities, which then influences location of people. We contribute to these literatures by showing that the effect of crime on employment decentralization creates externalities over city amenity and productivity.

Our paper also contributes to a long-standing stream of literature that studies the determinants of violence and crime.[5] In particular, we relate to the strand of this literature that studies the biological determinants of violence. As reviewed by Rowe (2002) and O'Flaherty and Sethi (2015) there is a growing body of evidence on how genetic, medical and environmental factors may increase the propensity of violent behavior. In particular, we build on a medical literature that has shown how lead poisoning is a potent neurotoxin that is closely related to aggressive and violent behavior. In economics, there is a new and growing stream of literature that studies the relationship between lead poisoning and crime (see Section 4.1 for a review of the literature). We contribute to this literature by exploiting a new source of cross-sectional variation given by the type of soil in a city. We then use this new instrument to provide causal evidence of the effect lead poisoning from resuspended lead has on violent crime.

Moreover, we contribute to the literature that studies the effects of crime. Here, the literature has mainly focused on the detrimental effects that a violent environment has on the young, especially when it comes to their educational decisions (Bowen and Bowen, 1999; Henrich et al., 2004). Another important strand of the literature has instead looked at the effects crime has on economic activity, mainly by deterring investments (Daniele and Marani, 2011; Detotto and Otranto, 2010). In this paper, we provide causal evidence of the effects that crime has on suburbanization and how this shapes the location of people inside a city.

Finally, our work is one application of machine learning for construction of instrumental variables (see Athey and Imbens, 2017 for a review). We select a proxy of soil quality between a large set of possible alternatives using an algorithm that finds the instrument that maximizes the relevance condition. We run the first stage of our regression for any possible interval of soil pH and we select the interval that maximizes the F-statistics. We then show how the pH of the soil selected by this algorithm conforms to what is expected by the chemical literature. We argue that the soil quality index selected is in line with the identification assumption required for exogeneity and we show that the results are robust to the use of other proxies.

This paper is structured as follows. Section 2 discusses the empirical strategy to obtain the causal effect of crime on suburbanization. Section 3 describes the data used. The instrumental variable and the identifying assumptions are explained in Section 4. Empirical results are reported and discussed in Section 5. Section 6 examines the possible threats to identification and provide evidence of the robustness of the results. In Section 7, we present the spatial equilibrium model and estimates of the externality effects of crime rates on amenities and productivity. Lastly, Section 8 concludes.

[5]Most of the efforts in this literature have been concentrated in assessing the effect of police, incarceration and the judicial system on the criminal activity (for the most recent literature review, see Chalfin and McCrary, 2017). Another strand of the literature instead has been focused on how different social and economic circumstances may affect crime. Examples of these determinants are income inequality (Kelly, 2000), immigration (Bianchi et al., 2012), gun laws (Ludwig, 1998) and social cohesion (Goudriaan et al., 2006) among many others.

2 Empirical strategy

The empirical model we want to estimate is reported in Equation 1. Our objective is to understand the effect of the increase in violent crimes per capita (*VC*) in the city center (*cc*) of one metropolitan area (*m*) in a particular year (*t*) on the suburbanization of that city. The proxy for suburbanization ($sub_{m,t}$) we use is the proportion of population living in the city center (pop^{cc}) over the total city population, which is the sum of the population in the city center and in the suburbs (pop^{ncc}). In order to control for unobserved heterogeneity we introduce both metropolitan area (MSA) and year fixed effects, τ_m and τ_t, respectively. We also include geographic (g) specific time trends, i.e. $\tau_g \times \tau_t$. In our preferred specification we impose Census District time trends (discussion of these time trends is presented in Section 4.3).

$$sub_{m,t} = \frac{pop^{cc}_{m,t}}{pop^{cc}_{m,t} + pop^{ncc}_{m,t}} = \tau_m + \tau_t + \beta VC^{cc}_{m,t} + \tau_g \times \tau_t + \varepsilon_{m,t} \tag{1}$$

The OLS estimation of the coefficient of the effect of violent crime on suburbanization, β can suffer different biases. Firstly, reverse causality might be present since more suburbanized cities might have poorer city centers, which in turn can increase crime rates in the city center. Evidence of this reverse causality has been found by Jargowsky and Park (2009). Moreover, omitted variable biases can contribute to the inconsistency of our estimation. One possible omitted variable is the proportion of black people living in the city center. Boustan (2010) shows that part of the white flight has followed the influx of black population in the city. This estimation can also suffer omitted variable bias if cities with more highways tend to have higher level of crimes.

We propose a new instrument that, we argue, can exogenously predict crime at city center level: the interaction between the lagged amount of the tetraethyl lead (TL) used in car gasoline in the U.S. and a proxy for the bioavalability of lead to humans in the ciy centers. Lead has an effect of brain development of children, and increase their aggresive behaviour. The highest potential of delinquency is reached at 19 years old.[6] Hence, for any given year, we use tonnes of tetraethyl lead in cars 19 years before to predict crime rates. Section 4.1 discusses in details the relationship between lead and crime.

Lead adsorption in the soil depends on particular soil characteristics. Bioavailability of lead is proxied by a specific function of the average pH of the soil in the city center. In particular, our soil bioavailability indicator is a dummy variable taking value 1 if the average pH of the soil is between the values of 6.8 and 7.7. We obtain this proxy using a machine learning algorithm described in Section 4.2. The first stage of our instrumental variable estimation is reported in Equation 2. One of the main advantages of our instrument is that it varies both in space and time, therefore we can include year and city fixed effects in our first stage.

$$VC^{cc}_{m,t} = \mu_m + \mu_t + \chi TL_{t-19} \times \mathbb{1}\left(6.8 \le pH^{cc}_m \le 7.7\right) + \mu_g \times \mu_t + \epsilon_{m,t} \tag{2}$$

[6] According to United States Department of Justice (1993) in 1965 people with 19 years old had the highest arrest rates for violent crimes.

3 Data

We have assembled an unique database for 306 city centers from 1960 to 2014 in the U.S. combining different data sources. The first data source we use is the F.B.I. Uniform Crime Reporting (UCR) Program Data (United States Department of Justice). For each local enforcement agency, which is coded by a Originating Agency Identifier (ORI) number, this data source provides information about monthly number of crimes for each year for all different kind of crimes.[7] This database also reports the total jurisdictional population under responsibility of that particular ORI. We use this information to compute our suburbanization measure.[8] We use the F.B.I. definition of violent crime, that is the sum of murder and non-negligent manslaughter, total robberies, forcible rape, and aggravated assaults.[9] We use the Law Enforcement Agency Identifiers Crosswalk database to link ORIs to Census Geographic Definitions (National Archive of Criminal Justice Data, 2006). We keep geography fixed at 2000 definition and we aggregate all the information at U.S. Place level.[10]

We merge our database with the data provided by Baum-Snow (2007). This database contains information of several social and economic characteristics of Metropolitan Statistical Area (MSA) and city centers.[11] Moreover, we use the definition of city centers provided by Baum-Snow (2007), that is for each MSA he defines the city center as the U.S. Place with the largest population in 1950.[12] Therefore, for every city center we can compute its jurisdictional population, the population of the rest of places inside the same MSA (that we call suburbs) and the population of the MSA. The suburbanization measure we use is the population in the city center divided by the population in the MSA. Similarly, we construct violent crime rate per capita at city center, suburb, and MSA level.

For the construction of our instrument we use two different databases. First, we use the United States Geological Survey (U.S.G.S.) General Soil Map in order to obtain information about the soil pH (United States Department of Agriculture). We use as

[7]We aggregate monthly data to years data. If crimes were not reported for more than 9 months we reweight the number of crimes by 12 divided the number of months in which data are missing.

[8]F.B.I. population in Census years is very similar to the population obtained by the U.S. Census. For non-Census years F.B.I. produces its own population estimation. We do not believe that the possible measurement error in population relates in any way with our instrument. Moreover, we also present robustness using only Census population data. There are some missing values in the population data, we substitute this value by the mean value of population in the ORI. This procedure does not alter in any way our results and interpretation.

[9]FBI defines these crime as follows. *Murder and nonnegligent manslaughter*: "willful killing of one human being by another". *Robbery*: "taking or attempting to take anything of value from the care, custody, or control of a person or persons by force or threat of force or violence and/or by putting the victim in fear". *Rape*: "penetration, of the vagina or anus with any body part or object, or oral penetration by a sex organ of another person, without the consent of the victim. Attempts or assaults to commit rape included". *Aggravated assaults*: "unlawful attack by one person upon another for the purpose of inflicting severe or aggravated bodily injury. Simple assaults excluded".

[10]According to the U.S. Census, "U.S. places are settled concentrations of population that are identifiable by name. They can be legally incorporated under the laws of the state in which they are located (*Incorporated Places*) or not (*Census Designated Places, CDC*). CDC boundaries are defined by the U.S. Census in cooperation with local or tribal officials, and they usually coincide with visible features".

[11]Baum-Snow (2007) also keeps MSA geography constant over time using definitions from 2000.

[12]We keep only information of ORIs inside one MSA. We drop all the ORIs belonging to multiple counties at the same time. We only keep observations of municipal jurisdiction crime.

pH measure the negative logarithm to the base 10 of the hydrogen ion activity in the soil using the 1:1 soil-water ratio method representative value. For every U.S. Census Place we compute the average pH level and information about earth slope, elevation and precipitation.[13] Second, national consumption of tetraethyl lead as gasoline additive comes from the Bureau of Mines Mineral Yearbooks (United States Bureau of Mines).

Data on firms decentralization comes from United States Census County Business Pattern (CBP) 1974-2013. This database reports the number of employed workforce and payroll for every industry and county. Data until 1998 reports information up to 4 digits SIC industries, and data from 1999 onwards up to 6 digits NAICS industries. We only keep 2 digits industries.[14] We identify in which county the city center of one MSA belongs and we compute the proportion of employment in every industry in the county over the total employment in the same MSA. If one MSA is composed only by one county we assign a missing value to the proportion of employment in every industry in the county over the total employment in the same MSA.

Our final database encompasses more than 9,750 observations from 1960 to 1991, and 7,038 observations from 1992 to 2014. All our main discussion will focus on the period from 1960 to 1991. This is given by the fact that lead poisoning increased until 1972, and the biggest effect on crime of lead poisoning in one year appears 19 years later, when affected children have the maximum probability to commit a crime. We devote Online Appendix E to discus what happens after 1991. In this period lead poisoning decreased in U.S. and as a result also crime rates decreased, but at the same time U.S. cities continue to maintain suburbanized (see Figure 1(a)). Summary statistics of our database from 1960 to 1991 are reported in Online Appendix A.

4 Instrumental variable

4.1 Background on lead poisoning

Lead is a heavy metal with several properties. It has high density, lasts longer and is more malleable than iron, is resistant to corrosion, and has relative abundance. Because of these characteristics lead was adopted historically for several uses: plumbing, solder, painting, bullets, and as a gasoline additive. According to Dapul and Laraque (2014) there are several ways through which children and adults can get exposed to lead. Ingestion sources are lead-based paint, contaminated water by lead pipes, lead settled in soil because of leaded gasoline, paint or other industrial sources,

[13]U.S.G.S. divides the U.S. in different map areas. Every map area is composed by different soils (components), and every component is composed by multiple layers (horizons). We use information only at soil level, that is when the distance from the top of the soil to the upper boundary of the soil horizon is 0. For every map unit we compute the weighted mean of pH of the components, weighting by the component percentage in the map unit. Finally, for every place we compute the weighted mean of pH of the map area, weighted by their area.

[14]We also aggregate data at bigger industries definitions: agriculture, good producing industries, service producing industries and other industries. We aggregate the industries as follows. *Good producing industries*: Mining, Construction and Manufacturing. *Service producing industries*: Transportation, Communications, Electric, Gas, And Sanitary Services, Wholesale Trade, Retail Trade, Finance, Insurance, And Real Estate, Services

food cultivated in contaminated soils, or leaded objects (such as children's toys). The two main inhalation sources of lead has been leaded gasoline and occupational hazards in the construction, soldering, painting, plumbing, automotive and ammunition sectors.

Lead has been used since antiquity. Its use as pigment was documented in Ancient Greece and Roman pipes were largely built with lead. The use of lead for pipes, paint and as gasoline additive has followed different timing. Lead pipes were installed on a major scale in the U.S. since the late 1800s. The danger of lead pipes was increasingly documented in the late 1800s and early 1900s and by the 1920s many cities and towns were prohibiting or restricting their use (Rabin, 2008). Conversely, the use of lead in paint peaked in 1920s, and then its use declined significantly (Mielke, 1999). The Lead-Based Paint Poisoning Prevention Act, which restricted the lead content in paint, was signed in 1971, and finally lead was banned from paint in 1978.

Tetraethyl lead was mixed with gasoline from the 1920s, because it can improve engine compression by raising the octane level of gasoline. The consumption of leaded gasoline skyrocketed in post World War II because of the increase in the use of lead as antiknock gasoline additive and the increase in the number of cars. In 1965, it was discovered that lead had a pollution effect on the environment (Patterson, 1965), and several works followed in order to prove the link between gasoline and lead pollution. Patterson's work also began an intense debate between environmentalists and the strong industrial lead lobby. The phase-down of leaded gasoline in U.S. began in 1975 when the Environmental Protection Agency (EPA) required major gasoline retailers to sell at least one grade of unleaded gasoline that was required to protect new car models with catalytic converters (Nriagu, 1990). The lead phase-down continued during the 1980s when the EPA set new limits for the amount of lead in gasoline. Leaded gasoline was finally banned in 1996.

A large medical and biological literature has given evidence of the health effect of lead, in particular on neurobehavioural development in children (see Roper et al., 1991 and International Programme on Chemical Safety, 1995). Lead is a potent neurotoxin which alters the formation of the brain and as a result influences the formation of cognitive and non-cognitive skills (see Toscano and Guilarte, 2005 and Cecil et al., 2008). According to Roper et al., 1991, "Children are at higher risk for lead exposure because they have more hand-to-mouth activity and they absorb more lead than adults". It has been shown that even low level exposure to lead during childhood is related to cognitive and behavioral outcomes, such as lower IQ, ADHD, and hyperactivity (see Banks et al., 1996, Canfield et al., 2003, Chandramouli et al., 2009, and Nigg et al., 2010). Moreover, early age lead poisoning has been found to relate to antisocial behaviours, such as aggressivity, violence and impulsivity, increasing the risk of delinquency (see Denno, 1990, Needleman et al., 1996, and Needleman et al., 2002). Likewise, prenatal and childhood blood lead concentrations are associated with more criminal offenses (see Stretesky and Lynch, 2001 and Wright et al., 2008). All these reported effect are given by the fact that "lead damages neurotransmitter function in the brain that regulate impulse control" (Aizer and Currie, 2017).

The effects of lead have been object of study in several recent works in economics. Exploiting differences in road proximity and the de-leading of gasoline, Aizer and Currie (2017) find a causal positive effect of lead on juvenile delinquency. The positive relationship between lead and criminal behaviour has been also found by Reyes (2015a), who also exploits variation coming from the phase-down of leaded gasoline. Feigenbaum and Muller (2016) show that water pipe lead exposure increased homicide rates in the 1920s and 1930s, instrumenting lead exposure by city distance from lead refineries. A second strand of works found a negative causal effect of early childhood lead exposure on academic achievement (Aizer et al., 2016, Reyes, 2015b, Grönqvist et al., 2016, and Ferrie et al., 2012). Similarly, Billings and Schnepel (2017) estimate how lead-remediation policies can reverse the negative outcomes of lead poisoning. A last group of research identifies the positive effect of lead exposure on mortality in the 1920s exploiting variation from water pipe lead poisoning coming from different water acidity, measured by the water pH (Troesken, 2008, and Clay et al., 2014).

In this paper we are interested in obtaining exogenous variation of crime rates in the U.S. in the years between the 1960s and the 1990s, a period in which the U.S. experienced a dramatic increase in violent crime. We exploit the massive increase of national consumption of leaded gasoline and its effect 19 years later on violent crime, when poisoned children had the highest potential for delinquency. The time relationship between national levels of violent crime and lagged tetraethyl lead is evident from Figure 2. In fact, the increase in tetraethyl lead matches with the posterior increase in violent crimes, the two time series reaching their peaks in 1972 and 1991, respectively. We exploit official national levels of lead poisoning by gasoline published by United States Bureau of Mines (United States Bureau of Mines) rather than local levels. The reason for this choice is that local lead exposure can be correlated with potential confounders of suburbanization, such as the proportion of highways and cars in a city.

[INSERT FIGURE 2 HERE]

We obtain geographical variation of the effect of lead on crime by exploiting the fact that lead is absorbed differently by different types of soil. Lead released from combustion of leaded gasoline becomes airborne and accumulates in the top 1 to 2 inches of soil. Evidence suggests that the bioavailability of lead in soil reaches its lowest level with a near-neutral soil pH, i.e. when the pH of the soil is around 6.5 or 7 (Reddy et al., 1995, Stehouwer and Macneal, 1999 and Peryea, 2001).[15] Despite the existing evidence on the fact that the bioavailability of lead decreases between acidic and near-neutral soils, where it reaches its minimum value, we are not aware of any study about bioavailability in very alkaline soils (pH higher than 7.5).

Children can ingest residual lead in the soil by eating the soil or inhalating it because of air dust resuspension (Laidlaw and Filippelli, 2008, Zahran et al., 2013, and Aizer and Currie, 2017). In fact, due to resuspension of roadside soil lead can be transported longer distances inside the city and then house dust can be contaminated by the soil attached to shoes (Filippelli et al., 2005, Hunt et al., 2006, Laidlaw and Filippelli, 2008). As a result it is not needed to live right close to a soil surface to get poisoned.[16] Several studies have assessed that the lead entering homes is a combi-

[15] "The bioavailability of lead in soil depends on its solubility, i.e. how tightly it is held by soil particles" (Stehouwer and Macneal, 1999). Lead is more soluble in acidic soils (pH lower than 6.5), and it is less soluble in neutral soils (pH between 6.5 and 7.5). Lead availability is considered to be minimized when the pH of the soil is higher than 6.5 or 7.

[16] We exploit variation coming from natural soil only. Despite important part of cities are paved, in

nation of lead from cars with smaller amounts of lead from paints (Clark et al., 2006, and Laidlaw and Filippelli, 2008). According to these studies this is given by the fact that lead paint particles tend to be larger than the one formed by leaded gasoline and then they do not penetrate cracks in homes.

4.2 Construction of the soil quality proxy

We multiply national lagged levels of tetraethyl lead consumption with a proxy for average soil quality at the city center level to obtain exogenous variation in crime rates. We use a function of the average soil pH in the city center as proxy for lead availability. This is in the same spirit to the use of water pH as instrument for water lead pipe poisoning done by Ferrie et al. (2012), Troesken (2008), and Clay et al. (2014). From the previously reported evidence we know that the proxy for soil quality we need to exploit has to be closer to near-neutral soil pH. For every city center, we combine data for the average pH level from the United States Geological Survey with crime observations from F.B.I. With our data, we can test that the effect of lead on violent crime is weaker at levels of pH close to 7.

We test this hypothesis estimating the marginal effect of lead on violent crime rates. We regress violent crime on the interaction of lead with polynomial of pH, up to the third order, and we plot the marginal effects computed at the mean. Figure 3 reports these results. As it is possible to observe, the effect of lead on violent crime is always positive and decreasing up to a pH of 7.5. Point estimate of the marginal effect then increases in the area of alkaline soils (pH higher than 7.5), but these variations do not appear to be significant.

[INSERT FIGURE 3 HERE]

A priori we are not sure of which interval of pH is the best in expressing lead toxicity. We apply machine learning tools to choose the most adequate instrument between the set of all potential candidates, that are dummy variables taking 1 for an interval between any possible minimum and maximum level of pH. For every possible pH interval we run the first stage regression of violent crime rate per capita over city and year fixed effects and the interaction between tetraethyl lead and the pH interval considered. We select the pH interval that maximizes the F-statistics for the relevance of our instrument.

This is similar to the regression tree method for prediction (see Breiman et al., 1984 for classic reference and Athey and Imbens, 2017 for a review). Regression trees are methods in which the covariate space is sequentially partitioned into subspaces such that the sum of squared residuals (SSR) is minimized. That is, given a variable X, regression tree methods find the value of the split c which divide the sample by $X < c$ versus $X \geq c$ and minimized the SSR. This process can be expanded to multiple covariates and splits. In our context we look for two splits of the variable pH. Moreover, instead of minimizing the SSR directly, we maximize the F-statistics. That is, we

the city centers there is still enough variation in natural soil. Large surfaces of cities are covered by parks and playgrounds. According to Harnik et al. (2015) in 2014 for high density cities in U.S. almost 12 % of their city area is parkland, and New York and San Francisco has around 20 % of their area as parks. This proportion is likely to be considerably bigger at the time of the lead poisoning happen in the 1960s.

maximize the SSR difference between a model in which we predict violent crime using only city and year fixed effects (the included instrument) and a model in which we predict violent crime using the included and excluded (the interaction between national lagged tetraethyl lead and a proxy for soil quality) instruments.

The instrument we select is the interaction between national lagged values of tetraethyl lead and a dummy taking values 1 if the pH of the soil in the city center is between 6.8 and 7.7. From now on we refer to MSA in which the average pH of the city centers is between 6.8 and 7.7 as places with good soil. All the other cities are referred as places with bad soil. Using the pH interval 6.8 to 7.7 provides a F-statistics of the excluded instrument of 262.16. The estimated coefficient of the interaction between tetraethyl lead and this soil quality proxy is -0.00528 with a standard error of 0.000326. The tetraethyl lead proxy used has been divided by its maximum level. Therefore, the effect of increasing tetraethyl lead from 0 to its maximum historical value in U.S. is increasing violent crime by 52 violent crimes less per 10,000 inhabitants in places with good soils. That is, the differential effect of lead in places with good and bad soils account for one third of the overall maximum value of violent crimes in 1990.

We summarize the results of our instrument selection procedure in Figures 4. Panel 4(a) reports all the estimated coefficients of the interaction between lead and any possible pH interval, while Panel 4(b) represents the corresponding F-statistics. As it is predicted from biological theory, the absorption of lead into soil should be weaker close to soil pH neutrality. Panel 4(a) shows exactly that first stage coefficient for dummy variables including pH levels lower than 6.5 in the good soil definition tends to be positive or non significant. As the minimum value of pH is higher than 6.5, then the interaction between lead and soil quality becomes negative and significant. Moreover, the first stage coefficients are robust around our preferred pH interval.

Changing the lower or the upper bound of the pH interval does not change the effect of the interaction between lead and pH. Similarly, the F-statistics of the first stage dramatically increases when the soil quality proxy includes near to neutrality pH levels (see Panel 4(b)). Despite the pH interval we choose is the one with the highest F-statistics, changing the upper or lower pH interval bounds does not alter the relevance of our instrument. We present robustness of our estimation to the use of other pH intervals in Online Appendix D.3.

[INSERT FIGURE 4 HERE]

Given our soil quality proxy, Figure 5 shows that the F-statistics of relevance of our instrument is maximized using the 19th year lag of national levels of lead poisoning, which is consistent with the evidence reported about the age structure of crimes by the FBI.

[INSERT FIGURE 5 HERE]

We give additional evidence of the effect of lead on crime in Online Appendix B.0.1. In Online Appendix B.0.2 we provide evidence that our generated instrument is not an outlier of the distribution of possible instrument by computing standard errors for our soil quality proxy. We perform a placebo exercise by creating random instrument and report the results in Online Appendix B.0.3.

4.3 Identifying assumption: relevance

The first assumption we need for the validity of our instrument is relevance. We obtain time-variation of crime using the national lagged level of lead used as gasoline additive 19 years before. Cross-sectional variation of crime comes from variation in city-averaged soil lead adsorption. Figure 6 reports the different time series of crime between places with good and bad soil. In the 1960s, when tetraethyl lead was beginning its increase, the level of crime rates between good and bad soils was very similar. As treatment took place, crime increased more in places with bad soil in terms of lead adsorption. In fact, peak in crime rates in places with good soil is 66 % of the peak in places with bad soil.

[INSERT FIGURE 6 HERE]

The U.S. map in Figure 7 gives a schematic representation of the exogenous variation we exploit. All of the East Coast has bad soil in terms of lead adsorption. In the rest of the US the pH seems to be more uniformly located across cities. We will exploit only variation inside census regions and districts by controlling for geographical specific time trends, so that results cannot be driven by a East versus rest-of-the-US comparison.[17]

[INSERT FIGURE 7 HERE]

Tables 1 and 2 show the relevance of our instrument reporting the coefficient of the first stage regression and the corresponding F-statistics controlling for time trends at different geographical levels. Since some of the variables in our database are measured yearly and some only in Census years, we report the first stage estimates using all the years in our sample in Table 1 and estimates using only Census years in Table 2. Our first stage coefficient is always significant. Using all the years in the sample our F-statistics range between 262 and 47, in the cases of using only city and year fixed effects or also imposing state specific time trends. Using only census years we obtain sufficient F-statistics imposing year and MSA fixed effects and Census region specific trends. Therefore, we augment the model in Equation 1 using Census division specific trends in the case of using variables measured annually. When we use variables measured only in Census years we control for Census region specific trends.

[INSERT TABLE 1 HERE]

[INSERT TABLE 2 HERE]

[17] Census Divisions are defined by the U.S. Census as: *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 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.
Moreover, Census regions are defined by the U.S. Census as: *West*: Pacific and Mountain; *Midwest*: West North Central and East North Central; *Northeast*: New England and Middle Atlantic; *South*: West South Central, East South Central, and South Atlantic

4.4 Identifying assumption: exogeneity

Our identification assumption is that in years in which national consumption of tetraethyl lead increased places with good and bad soils would have had similar trends in terms of suburbanization other than through differences in violent crime. This assumption is credible if soil pH is as good as randomly assigned. We present two balancing tests to support our claim. First, we demonstrate that places with good and bad soils have parallel trends in both suburbanization and violent crimes prior to the massive increase in tetraethyl lead. Second, we show that places with good and bad soil have similar pre-trends also in terms of other observable characteristics.

Table 3 reports the balancing test for the suburbanization and crime variables. We report both the difference in levels and trends between places with good and bad soils. Moreover, we do this exercise both without controlling for any geographical aggregation fixed effect ("All U.S." columns) and also controlling for Census Division fixed effects ("Inside Division" columns). The pre-trend assumption seems to be guaranteed. As soon as we control for Census Division trends, places with good and bad soil had similar trends between the 1950 and the 1960, that is before the great increase in national use of lead as gasoline additive. To further reassure of the exogeneity of our soil quality proxy we also show level differences. Places with good and bad soil tend to be similar in terms of their pre-treatment level of suburbanization, population and crime as soon as we control for Census Division dummies.

[INSERT TABLE 3 HERE]

Balancing test for other observables are also reported in Tables 3. From Table 3 we can rule out that places with good and bad soils had different trends in other geographic and social characteristics that can influence suburbanization. It seems however that places with good and bad soils have some differences in pre-treatment levels in terms of rent, income, precipitation rates, business and manufacturing employment, public transportation and education. We show in Online Appendix D.2 that our results are robust to the inclusion of these variables as controls. It is interesting to note that places with good and bad soil are similar in terms of agriculture and mining employment. Therefore, we can rule out the possibility that soil pH is affecting suburbanization by changing the relative proportion of land used for urban and agricultural use inside a city.

Table 3 shows that places with good and bad soils are very similar in terms of pre-treatment levels and trends of highway construction. Hence, our results cannot be driven by highway construction, a channel emphasized in previous literature. We show in Online Appendix D.1 that the results we found are robust even controlling for highways, and dealing for its particular endogeneity.

In order to understand what would have happened to violent crime if the lead poisoning shock did not take place, we estimate the time-varying effects of soil pH on crime. We run regressions of the effect of soil pH interacted with year dummies on violent crime, that is Equation 3.

$$VC^{cc}_{m,t} = \mu_m + \mu_t + \chi_t 1(year = t) * good\,soil^{cc}_m + \mu_g \times \mu_t + \epsilon_{m,t} \tag{3}$$

Results of this regression are reported in Figure 8. In line with the results reported in the previous Section, between 1960 and 1991 violent crime increased less in city centers with good soil. As the de-leading phase started this difference shrank. In 1996 lead was completely banned, this means that by 2014 almost all adults have suffered very little lead poisoning and people younger than 18 years old were not poisoned at all. As we observe in Figure 8 there is no statistical difference between good and bad soil city centers today.

[INSERT FIGURE 8 HERE]

This is further evidence of the exogeneity assumption. In particular, cities with good and bad soil started with the same level of violent crime when there was no lead poisoning. They then ended with no differences in violent crime, when lead poisoning was no longer relevant. Therefore, this evidence supports our claim that violent crime would have always be the same in these two kinds of cities if lead poisoning would not have been there.

Figure 8 also provides evidence in favour of the exclusion restriction. If the effect of pH on crime is only passing through its interaction with lead then the results of the estimated regressions of the effect of soil quality by year should be similar to the lagged time series of lead poisoning. The time series of the reduced form coefficients of our soil quality index mimics the lagged time series of lead, strongly supporting that the effect of pH on crime is very likely to pass only through its interaction with lead.

5 Results

5.1 Baseline Results

In this Section we first provide estimates of our main equation of interest, Equation 1, that looks at the effect of violent crime in city centers on suburbanization. As shown in column (1) of Table 4 there is a negative correlation between violent crime and the share of population that lives in the city center. As discussed previously this estimate cannot be interpreted as causal, and because of this we implement our instrumental variable methodology. That is, we predict violent crime using the interaction between lagged national lead levels and a proxy for soil quality. Column (2) shows that places with good soil experienced a slower increase in violent crime and this difference is substantial. In 1991, at the peak of lead exposure for potential criminals, a MSA with bad soil had 0.91 standard deviations more violent crimes with respect to one with good soil.

[INSERT TABLE 4 HERE]

In Column (3) we estimate the causal effect of crime on the share of people that lives in the city center using our instrumental variable strategy. Estimates show that an increase in one standard deviation in violent crime decreases the share of population living in the city center by 7.2 percentage points. The upward bias of the OLS estimate is consistent with the presence of reverse causality bias from suburbanization to violent crimes. Online Appendix C discusses for which values of the estimated coefficients the OLS bias could have been induced by reverse causality.

Column (4) reports the reduced form effect of our instrument on the percentage of people living in the city center. The effect of increasing lead from no poisoning to the maximum level increased suburbanization in places with bad soil by 6.8 percentage points with respect to places with good soil.[18]

In Column (5) we estimate our preferred specification in which we additionally control for census division times year fixed effects. In this regression we are only exploiting differences between good and bad soil city centers that are inside the same census division. Our estimates are now robust to any potential omitted variable common to MSAs in a certain census division. As shown in Column (5), previous results are robust to this specification. According to these estimates an increase in one standard deviation in violent crime decreases the share of population living in the city center by 8.4 percentage points. This implies that if in 1991 the level of crime would have been as low as in 1960 the percentage of people that lived in the city centers would have been 15 percentage points higher[19].

Column (6) shows the same results using as dependent variable the population in city centers. A one standard deviation increase in violent crime decreases the population of the city center by 26%. The overall increase in crime rates from their level in 1960 to that one in 1991 translates into a 46% decline in city center population.

We show that our results are robust to several specifications. In Section 6.1 we show that despite the fact that lead could potentially affect educational outcomes, this channel does not bias our results. Online Appendix D discusses the additional robustness we conduct. We discuss that our results are robusts to the inclusion of possible confounders, such as highways (Online Appendix D.1) and many other possible variables (Online Appendix D.2). In Online Appendix D.3 we demonstrate that our results do not depend on the particular decision of the instrument. We also show that the effect of violent crime on suburbanization does not depend on the particular geographical variation we exploit (Online Appendix D.4). Finally, in Online Appendix D.5 we demonstrate the robustness of the standard errors estimated.

We report estimates for the effect of crime on suburbanization in the de-leading phase, after 1991, in Online Appendix E. We show that when lead poisoning decreased, violent crime rates decreased faster in places with bad soil than in places with good soil. However, in the same period places with bad soil did not decrease suburbanization, providing possible evidence for the persistence of the effect of crime on suburbanization. Online Appendix F discusses how our first and second stage results can potentially vary through time. We show that the effect of resuspended lead on violent crime is constant through decades. Nevertheless, the effect of violent crime rates on suburbanization is declining through time.

The effect of violent crime on suburbanization also presents important heterogeneity with respect to several city characteristics. We present the analysis of the mechanism and channels behind the crime effect on suburbanization in Online Appendix G. We show that this effect is stronger in cities in which the suburbs have lower levels of black population with respect to the city center. Moreover, suburbanization was stronger in cities with higher levels of previous suburbanization, that were richer, with smaller geographical constraints, and where more highways were built.

[18]This is given by the fact that the national level of lead poisoning has been normalized by its maximum value.

[19]Violent crime per capita in city centers increased by 1.79 standard deviations between 1960 and 1991.

5.2 Magnitude of the effect of violent crime on suburbanization

To get a better idea of the size of the effects estimated in Table 4 we construct two counterfactual scenarios: one in which crime remained throughout our sample at the low level of 1960 and one in which crime in city centers increases at the same rate as in the suburbs. The time series of the share of people living in the city center in the U.S. and in these counterfactuals are displayed in Figure 9.

[INSERT FIGURE 9 HERE]

As previously described there was a clear pattern of suburbanization in the period studied. The percentage of people living in the city center moved from 44% in 1960 to 33% in 1990. Our estimates instead predict that if the US had maintained the low levels of violent crime in city centers observed in 1960, we would have seen a process of urbanization of US cities. The percentage of people living in city centers would have increased, reaching 50% in 1991.

Some caveats are necessary to have in mind when interpreting this result. First of all, we do not want to claim here that if the U.S. would have banned the use of lead in gasoline since the beginning we would not have observed any growth in violent crime since the 60s. Many factors would have influenced the violent crime rate in this period and only one of them is lead exposure. Furthermore, it is important to notice that with this counterfactual experiment we are also not exploring what would have been the suburbanization trends if all the MSAs in the U.S. would have been in our control group, namely good soil. In fact, also the MSA we used as control in our estimations have suffered an increase in crime in this period. Moreover, we are not considering that the proportion of people living in city centers could have mechanically decreased because of the limitation in space in the centers to allocate the demographic increase in population in any U.S. city.

It is likely that crime rates would have still increased in U.S. if lead poisoning had not happened. One possibility might have been that crime rates in city centers would have followed the trends that the suburbs experienced.[20] Therefore, we compute a second counterfactual experiment in which crime rates in the city center would have increased at the same rate as in the suburbs. In this case our estimates predict that if the U.S. city centers increased crime as in the suburbs then the proportion of people living in the city center would have increased only marginally. As it shown in Figure 9, the percentage of people in city centers would have increased from 44 % in 1960 reaching 45% in 1991.

Cullen and Levitt (1999) estimate that an increase in 10% in crimes rates in the city center translates into a decline in city center population by 1%. Our estimated effect has a bigger magnitude. We estimate that an increase in 10% in crimes rates in the city center translates into a decline in city center population by 2.6%.[21] The difference in magnitude can be explained by the fact that Cullen and Levitt (1999) use as instrument the punitiveness of the state criminal justice system. For instance, this

[20] We show in the Online Appendix B.1.2 that the lead poisoning shock we are exploiting did not affect in any form the suburbs.

[21] Violent crime per capita in the city center has mean and standard deviation of 0.00577 and 0.00581, respectively. Therefore, increasing violent crimes by one standard deviation corresponds to increasing violent crime by 100.6%. We computed the effect of a 10% increase in crimes rates on population dividing -0.257 by 10.06.

instrument can influence both crime rates in city centers and suburbs. Subsequently, the increase in crime rates in the suburbs can lead to an increase in the population of city centers.[22]

We can compare our findings with the results found in similar studies of other causes of suburbanization. The relative increase in crime between city centers and suburbs from 1960 to 1991 implies a 35% decrease in the population of city centers.[23] This point estimate is higher than similar coefficients found in other studies, but it include them in its confidence interval.[24] According to Boustan (2010), black migration from the South was responsible for a 17% decline in total urban population. Baum-Snow (2007) reports that the construction of the interstate highway system led to a decrease of central city population by 23%.[25] That is, for a city like Philadelphia, with 2 million people living in the city center in 1960, 4,000 more violent crimes in 30 years move away the same number of people from the center to the suburbs as if one highway passing from city center would have been built.[26]

The different suburbanization mechanisms proposed by Baum-Snow (2007) and Boustan (2010) are likely to be complementary with the increase in crime rates. In Section 5.3 we show that suburbanization caused by violent crime has been disproportionately driven by the white population. The abandonment of city centers by whites and the consequent increase in the relative proportion of black population in city centers could have in turn made more white people move to the suburbs, consistent with the story proposed by Boustan (2010). Similarly, we show in Section 7 that the increase in violent crimes stimulate the construction of highways, which can then explain part of U.S. suburbanization (Baum-Snow, 2007).

5.3 Displacement Effects

In this section of the paper we explore whether violent crime does not only change the share of people living in the city center but displaces people from one city to another. Investigating this effect is important for two main reasons: First of all, if this was the

[22] Cullen and Levitt (1999) control in one specification for crime rates in the suburbs and they indeed find a stronger effect of crime rates in city centers on the population in city centers. However, crime rates in the suburbs can be a bad control in that specification.

[23] The 35 % refers to the difference in change population if crime would not have increased from the 1960, 46 %, and if crime would have stayed in city centers as in the suburbs, 11 %. This last number has been found multiplying 0.257 by the standard deviation increase in violent crimes from 1960 to 1991, 1.79, and then dividing it by the relative increase in the number of crime in the city centers from 1960 to 1991 with respect to the same increase in the suburbs, 4. The relative increase in crime between the centers and suburbs has a similar magnitude by computing using the predicted crime in city centers and outside from the first stage regression

[24] The point estimate of the effect of increasing violent crimes in city centers with respect to suburbs from the levels of 1960 to the level of 1991 on the logarithm of population in the city center is 0.35 with a standard error of 0.097

[25] This number has been computed multiplying the effect of building a new highway ray in the city center, -0.09, by the average number of rays built between 1950 and 1990, 2.6

[26] This number have been found dividing the effect of building a new highway ray in the city center, -0.09, by the effect of one violent crime per capita on suburbanization, -0.0843/1.79, normalized by the population of Philadelphia living in the city center in 1960. Philadelphia city center decline from 2 million people in 1960 to 1.5 in 1991. Moreover, Philadelphia city center has 330 violent crimes per 100'000 in 1960, and 1'400 in 1991.

case it would add some difficulties to the interpretation of our estimates. If this happened, it would mean that all cities would be in some way treated by the increase in lead but for different reasons. The places with bad soil would have been treated because of the increase of violent crime in the city center, while the places with good soil would have been treated by an increase of the total population driven by the migrants escaping from the violent cities. Furthermore, it is important to understand which kind of suburbanization process is caused by an increase in violent crimes. We could observe a decrease of the percentage of people living in the city center with respect to the suburbs in a context where both of them are losing population due to an increase in violent crime, and the city center is experiencing this process at a faster pace. The other option instead is that people are moving inside the MSA away from the city center.

[INSERT TABLE 5 HERE]

Estimates in Table 5 show that violent crime does not displace people from one MSA to another but redistributes population from the city center to the suburbs. An increase in violent crime in the city center is not influencing the overall population of the MSA (see Column (1)). Increasing violent crimes by 1 standard deviation decreases the population living in city centers by 26% and increases the population in the suburbs by 14% (see Columns (2) and (3)).

Violent crime in the city centers decreased population in the city centers by a similar magnitude such as the increase in population in the suburbs. In particular, the increase in violent crimes from its level of 1960 to its level in 1991 moved an average of $83,000$ people from the city center to the suburbs.[27] That means that the increase in violent crimes in the city center from their level in 1960 to their maximum level in 1991 is responsible for moving almost 25.5 million people outside of city centers in the all U.S, that is almost 0.8 million people by year.[28] For a city of the size of Philadelphia each violent crime moved on average 4 people away from city centers per year.[29] On the other hand, the increase in violent crimes from its level of 1960 to its level in 1991 increased the population in the suburbs by $62,000$ people.[30]

Columns (4) and (5) show whether the racial demographic composition in the city changed because of the increase in violent crime. First, Column (4) of Table 5 shows how as city centers become more violent the percentage of blacks in the MSA does not change. This is further evidence of the fact that the phenomenon that we are studying is not displacing people from one city to the other but only moving people inside the same MSA. In Column (5) we can indeed observe that there is differential racial movements towards the suburbs. A one standard deviation increase in violent crime

[27] This number has been found multiplying 0.257 by the standard deviation increase in violent crimes from 1960 to 1991, 1.79, and then by the average population of city centers in 1960, $181,030$. The point estimate is $83,282$ and its estimated standard error is $23,294$

[28] This number has been found multiplying $83,282$ by the number of urban cities in our sample, 306

[29] This number have been found dividing the effect of one violent crime per capita on population in city centers, $83,282$, normalized by the population of Philadelphia living in the city center in 1960. Philadelphia city center decline from 2 million people in 1960 to 1.5 in 1991. Moreover, Philadelphia city center has 330 violent crimes per 100'000 in 1960, and 1'400 in 1991.

[30] This number has been found multiplying 0.144 by the standard deviation increase in violent crimes from 1960 to 1991, 1.79, and then by the average population of suburbs in 1960, $240,516$. The point estimate is $61,976$ and its estimated standard error is $23,032$

in the city center increases the share of blacks in the city center by 4.7 percentage points. This constitutes a substantial increase as in the 1960, before the suburbanization process began, 13.4% of the population of the city center was black. This estimate provides evidence of the “white flight", that is the movement of white affluent people to the suburbs. What these estimates show is that at least part of this phenomenon may be explained by the rise of violent crime in the city centers. Moreover, the change in racial composition in the city centers could in turn explain part of the subsequent suburbanization, consistent with the mechanism proved by Boustan (2010).

5.4 Effects on employment decentralization

Glaeser and Kahn (2001) show that cities in the U.S. are characterized by decentralization of employment inside the city. In this Section we want to understand whether violent crime has caused residential suburbanization only or it might also induce decentralization of employment location inside the city. In addition we want to understand if the decentralization of firms can be caused by residential suburbanization. In fact, as a response to the residential suburbanization two processes can happen. First, firms can move to city centers because of residential suburbanization if the increase in vacant housing in the city center decreases land cost and firms are able to reconvert residential areas into business areas. This process would increase further the monocentricity of a city in terms of employment location in the Central Business District (CBD). Second, firms might follow people in the suburbs in order to reduce workers' commuting costs, with the effect of creating new employment centers in the city outside the CBD. Similarly, residential suburbs can create infrastructure in the suburbs, such as highways, that firms can exploit.

As we discuss in Section 7 the decision of decentralization of firms will have important implications for aggregate city variables, such as productivity and amenities of the city. We collect data for every MSA about the distribution of employment between the county in which the city center is located and the rest of the city. From Table 6, column (1), we do not evince that overall firms decentralize as a result of higher violent crimes. However, this result can mask sector heterogeneity in the response to the increase in violent crimes.

Manufacturing is one of the sectors that relocates the most to suburbs after the increase in crime rates in city centers. This is likely because manufacturing relies on the use of large land space which is available in the suburbs. We also find that firms in wholesale trade, retail trade and other services move to the suburbs. Finance, Insurance, and Real Estate is the most important sector which does not decentralize as a result of the crime increase. The reason for which Finance might stay in the center can be related to the fact that knowledge spillovers and spatial proximity to other firms is more important in this sector. In addition, this sector tends to locate more in skyscrapers present in the CBD. Therefore, as a result of the crime and suburbanization shock many firms are relocating in the suburbs but firms in the Finance sector, which continue to stay in the CBD, leading to the possible creation of multiple employment centers in the city with different specializations.

[INSERT TABLE 6 HERE]

We have seen that both people and firms in some sectors move to the suburbs after crime increased in the city centers. We can provide evidence of whether people has followed jobs or the opposite is true. In order to do this we have estimated the effect of violent crime on suburbanization controlling for past levels of employment decentralization and the effect of violent crime on employment decentralization controlling for past levels of suburbanization.[31] Results are reported in Table 7. As shown in Columns (1) and (3) violent crimes caused both residential and employment decentralization in the manufacturing sector. If violent crimes cause people to move to the suburbs and then firms follow people, then when we control for the past level of suburbanization we should not find any effect of violent crime on firm decentralization. This is confirmed in Column (2). In fact, it seems that jobs followed people which have escaped city centers because of violent crimes. However, the effect of crimes on suburbanization is maintained even controlling for past level of employment decentralization in the manufacturing sector (see Column (4)). That is, results suggest that the first effect of violent crime is to make people leaving city centers and, then, firms decide to follow people to the suburbs.

[INSERT TABLE 7 HERE]

6 Threats to identification and further robustness

The exclusion restriction requires that the effect of the instrument on suburbanization is only passing through its effect on crime. In terms of our setting, this means that the interaction between lagged national lead and soil quality is only affecting crime and not any other variable that can influence suburbanization. One strength of our instrument is the use of lagged values of lead poisoning that are a priori only related to crime rates. We use a lag of 19 years because this is the age in which a person has the highest probability of getting arrested for a violent crime in 1965 (see United States Department of Justice, 1993). Unless lead poisoning through soil is affecting an omitted variable with exactly the same lag of 19 years, our estimates will be consistent.

In order to fully exploit timing idiosyncrasies of crime we conduct a robustness test in which we do not only use the maximum propensity of committing crime but all the age structure of crime rate. We discuss in Online Appendix D.3 how we perform this exercise and we show that all our results are robust to this specification.

Despite people can potentially leave the city center at the time they get poisoned by lead, for example because of higher pollution, this mechanism would not invalidate our estimates. This is given by the fact that we exploit the effect that lead has on crime 19 years later and not in the same year of the poisoning. Moreover, we provide evidence that people did not leave city centers immediately. This is confirmed by the fact that places with good and bad soil do not have different pre-trends differences in suburbanization between the 1950 and 1960, as it is shown in Section 4.4. The lead poisoning shock began around the 1940s and its effect via pollution should have been manifested before the 1960s. It was only from the 1970s that public opinion became

[31] Estimations have been conducted in the sample of years from 1974 to 1991 and without using Census region time trends in order to guarantee a sufficiently big F-statistics. In order to control for the possible endogeneity of the 10th lag of suburbanization or firm decentralization we include the interaction between the 29th lag of national lead poisoning and our soil quality proxy.

aware of the possible effect of lead poisoning by gasoline additives and the role of soil quality has not been known until relatively recently (see Reddy et al., 1995).

In Section 6.1 we discard the possibility that lead poisoning can affect suburbanization via its effect on cognitive abilities. We present additional evidence in favour of the exclusion restriction in Online Appendix B. We use the control function approach to give evidence that the effect of lagged lead 19 years before is likely to pass only via crime (Online Appendix B.1). We demonstrate that the particular function of pH we use for our instrument is unlikely to be related to agricultural productivity of one city (Online Appendix B.1.1). We show that there is no spillover from the city centers to the suburbs of crime rates and that the only variation in crimes caused by lead happen in city centers (Online Appendix B.1.2). In fact, it might be possible that crime in the suburbs increased because people poisoned by lead in the city center either relocate to the suburbs or displace to commit crimes to the suburbs. However, we do not find evidence supporting this claim. We find non-statistically significant coefficients of both good soils in the suburbs and in the city center on crime rates in the suburbs. We provide evidence that lead poisoning is only affecting violent crimes and not other crimes, as it is predicted by medical literature (Online Appendix B.1.3).

6.1 Potential confounders: cognitive abilities

Recent works have shown that lead poisoning has an effect on child educational attainment (see Aizer et al., 2016, Reyes, 2015b, Grönqvist et al., 2016, and Ferrie et al., 2012)). Education levels and human capital can bias our results if the effect of education on suburbanization follows the same age-structure as crime. If parents decide to suburbanize because of lower cognitive abilities of student peers of their children, this would not be a problem for our estimates since this effect should manifest before the 19-year lag when children are in school age.

Our estimates would be biased if people decide to suburbanize because of their own lower human capital skills. For example, lead poisoning could make people more anxious or racist and then decide to leave city centers. If this channel takes effectively place we should see that people with lower human capital are the one suburbanizing the most. Table 8 reports evidence against this possibility, by showing the demographic profile of people suburbanizing between 1975 and 1980.[32] People that decide to suburbanize have on average 31 years, they are more likely to be white, they tend to have higher high school performances, even between the category of whites, and they tend to have better occupational outcomes.

[INSERT TABLE 8 HERE]

Therefore, it is unlikely that people suburbanize because they get directly lead poisoned. Whites are more likely to suburbanize, and there is evidence of racial disparities in lead poisoning. In fact, Sampson and Winter (2016) demonstrate that "black neighborhoods exhibited extraordinarily high rates of lead toxicity compared to white neighborhoods". Table 9 shows that the higher propensity of violent crime of the black population is exacerbated by the presence of highways passing through the city

[32] The U.S. Census conducted in 1980 allows us to construct the demographic profile of people who left city centers for suburbs in the last 5 years. We could not construct the same statistics for different Census years.

center in years in which lead poisoning was higher. This is because the black population tends to live closer to highways and then they are more likely to get poisoned by lead. Similarly, we show in Online Appendix G that the effect of resuspended lead on crime is stronger in cities in which the city centers has more highways and a higher proportion of blacks.

[INSERT TABLE 9 HERE]

Additionally, our estimates would be biased if people decide to suburbanize because of lower human capital skills of other people in city center caused by lead poisoning. For this channel to be a problem, lead should affect human capital skills with the same age-structure as the one used to predict crime. In Table 10 we regress a proxy for human capital skills, the percentage of people with high school diploma, on the interaction between our soil quality index and different lead lags.[33] Column (1) shows that lead can influence human capital but in a different way from which it affects crime. We only find the 29th lag of lead significant to predict education outcomes. In contrast, the 19th lag is not significant. One possible explanation for this result is that lead affects violent crime which then influences the return of education, leading to a decrease of education 10 years later. However, the result that on aggregate lead can influence educational outcomes is not consistent with the use of different clustering in the standard errors, which takes into account the possible geographical correlation at Census district level in the errors. In fact, from Column (3) we cannot confirm the effect of the 29th lag of lead poisoning on the share of high school graduates in a MSA.

[INSERT TABLE 10 HERE]

We demonstrate in Table 11 that our estimates of the effect of crime on suburbanization are robust to the inclusion of the interaction of soil quality and lead poisoning 29 years before, which can indirectly influence educational outcomes. Column (1) reports the negative effect of violent crime on suburbanization using Census region times year fixed effects.[34] In column (2) we augment our estimations by also controlling for the 29th year lag of lead poisoning. We obtain similar results as when we do not control for other lags of lead poisoning.

[INSERT TABLE 11 HERE]

If the effect of lead is affecting suburbanization because it decreases human capital in the city and not because it influences crime we should expect a positive IV estimate of the effect of a proxy of human capital on the proportion of people living in city centers. Table 11 column (3) contradicts this hypothesis. While in Column (1) we observe that the IV estimate of violent crime on the share of population in the city center is negative, the effect of the share of high school in one MSA is also negative using resuspended lead as instrument. That is, it is possible that the estimation in Column (3) is biased because crime and education are correlated and resuspended lead is affecting crime alone.

To further discard the possibility that suburbanization might happen because of lower human capital in the city center, we run our estimation of the effect of crime controlling for instrumented values of educational attainment in the city. We in-

[33] We use the percentage of people in the MSA with high school diploma from Baum-Snow (2007) as proxy for educational attainment. Since we have data about percentage of people with high school diploma for decennial years we use the 9th, 19th and 29th lag of lead poisoning.

[34] We use these fixed effects because education is measured every Census year.

strument educational attainment using historic state compulsory education laws collected by Goldin and Katz (2008). In particular, we use two different measures of minimum age of compulsory schooling in 1910 as instruments: the school entrance age and the school leaving age. We obtain a time variant instrument for education by multiplying minimum age of compulsory schooling by the U.S. proportion of people with high school in a particular year. Results are reported in Table 12.

[INSERT TABLE 12 HERE]

Columns (1) and (2) replicate our unconditional results. Columns (3) and (6) jointly estimate the effect of violent crime and education on suburbanization using the two different education instruments (columns (4)-(5) and (7)-(8) show the corresponding first stage regressions). Both column (3) and (6) show that the effect of crime on suburbanization is robust to the inclusion of an instrumented education control.

7 Effect of crime on aggregate city variables

We have shown that violent crime has moved people and firms from city centers to the suburbs. In this Section we explore whether this increase in violent crime and suburbanization has generated any aggregate effects on the city or if suburbanization is just a zero sum reshuffling of resources around the city. This is an important question to tackle in order to gain understanding on how people and firms should be distributed in a city.

We first explore this question by looking at the effect that violent crime had on the total population of the city, house prices and median income. Results in Table 13 show that violent crime had no effect on city population while median income and house prices in the city increased. In fact, from Section 5.3 we already know that violent crime has only created movement of people inside the city and not between cities. From Column (2) and (3) we can see that violent crime has a positive effect on MSA housing affordability both measured as median gross rent per housing unit (Column (2)) or median single family house value (Column (3)). Moreover, we find that the increase in violent crimes is associated with higher income also controlling for the education levels (Columns (4) and (5)).

[INSERT TABLE 13 HERE]

In Table 14 we explore further what are the income effects of violent crime and if violent crime generated any changes in the means of transportation used in a city. We observe, first of all, that while income increased overall in the city these gains have not enjoyed by people living in the city center. In fact, we observe that overall inequalities increase in the city. In Column (1) we can observe that the Gini index within the city increased.

[INSERT TABLE 14 HERE]

In Column (2) of Table 14 we can observe that the ratio between income in the city center and the MSA decreased. In the city center the median income has decreased but not in a statistical significant way (Column (3)). On the other hand, the overall income in the MSA has increased indicating that suburbs have become particularly richer. This results are interesting because they uncover two dynamics that happen as a city becomes more violent and therefore suburbanized. The first is a selection process, where richer individuals move to the suburbs leaving the poorest in the city center. This is also in line with the results shown in Tables 5 and 8 on how the racial distribution of people inside a city changes after an increase in violent crime.

The second can be understood in combination with the fact that the cities that are becoming more violent and more suburbanized are not losing population. A way to make this possible is that incomes in the MSA increase to compensate for the increase in violent crime, the increase in transportation cost, and other negative amenities that the suburbanization process may generate. These two effects combined create the observed changes in income after an increase in violent crime. In the city center, the two effects might go in opposite directions. Because of selection only the poorest people stay but they have to be compensated with an increase in income so that they do not migrate to another city. The overall effect is that their income remains unchanged. Instead in the suburbs the two effects might reinforce each other. The increase in violent crime moves the richest people to the suburbs and added to this they also need to be compensated for all the negative amenities.

It is important to notice that these results do not fit with a model in which violent crime is only decreasing the amenities of a city. In that case, one would expect violent crime to decrease the population and/or to decrease house prices. Another mechanism that can explain these results is that the increase in violent crime has generated some positive externality on productivity. In fact, we know from Table 6 that violent crime increased employment decentralization, and city with multiple employment centers can be more productive.

Finally, we explore if violent crime generated also changes in the way that individuals move around the city. From Column (4) of Table 14 we can see that violent crime boosts the construction of highways possibly to facilitate suburbanization of people. The construction of new infrastructure can potentially have effects on productivity of the city. Moreover, since we know from Baum-Snow (2007) that highways have a positive effect on suburbanization, the effect of violent crime on suburbanization that we have found in Section 5.1 has to be interpreted as the general equilibrium effect of violent crime on suburbanization which do not partial out for the mediating factor of the highway construction. In column (5) we further observe that violent crime decreases the use of public transportation. These two last results confirm the fact that violent crime generated suburbanization and this ultimately decreased the demand for the use of public transportation and increased the demand for highways and the use of car.

From the results of this Section we can speculate that amenities and productivity in the city might have been affected by the increase in violent crimes. In the next Section we build a spatial equilibrium model in which every city is considered as a different economy in order to make sense of these between-cities comparisons and assess whether violent crime has influenced amenities and productivity.

7.1 Spatial equilibrium model

We rationalize the city-wide effects of violent crime using the Rosen-Roback spatial equilibrium model developed by Glaeser (2008) and Glaeser and Gottlieb (2009). The model has its base in the works of Rosen (1979) and Roback (1982), and the cornerstone of the model is the concept of spatial equilibrium: utility should be equalized between people living in different spaces. The same model has been applied to study the aggregate effect of city shape by Harari (2015). Each city (m) differs for its specific level of amenities (θ_m) and productivity (A_m). We assume that violent crime has an effect on these parameters. Our goal is to estimate how large and in which direction the effect of violent crime on these parameters has to be to match the city-wide estimates. The model consists of three agents: workers, firms in production sector and firms in construction sector.

Workers (i) have to decide in which city m to live. We assume perfect mobility between cities. Moreover, they have to decide how much to consume of a consumption good (C) and housing (H). The price of the consumption good is normalized to 1, while the price of housing is specified as p_m^H. They supply inelastically labour and obtain a city-specific wage (w_m). We assume a Cobb-Douglas utility function, where α represents the share of housing into utility. If a city has a higher level of amenities (θ_m) workers receive higher utility from that. The problem of workers is therefore the following one:

$$\begin{aligned} \max_{C_i,H_i} \quad & \theta_m C_i^{1-\alpha} H_i^{\alpha} \\ \text{s.t.} \quad & C_i = w_m - p_m^H H_i \end{aligned}$$

The solution of the worker's problem gives rise to the so-called spatial equilibrium condition. Let's define $\bar{v}_m$ as the indirect utility that should be equalized between cities. Plugging the optimal solution for the amount of housing, $H_i = \frac{\alpha w_m}{p_m^H}$, and consumption good, $C_i = (1-\alpha)\, w_m$, into the utility and taking logs we can write the spatial equilibrium condition as in Equation 4. Lower amenities in one city should be compensated by higher wages or lower house prices to obtain the same utility between cities.

$$log\left(\bar{v}_m\right) = (1-\alpha)\, log\,(1-\alpha) + \alpha\, log\,(\alpha) + log\,(\theta_m) + log\,(w_m) - \alpha\, log\left(p_m^H\right) \tag{4}$$

The representative firm in production sector decides the amount of labour (N) and traded capital (K) to hire to produce the consumption good. Labour is paid the wage level (w_m). Traded capital can be purchased at a price of 1 in any location. Every city is characterized by a specific level of productivity (A_m) and a fixed-supply of non-traded capital ($\bar{Z}_m$). As reported in Glaeser (2008), the assumption of the existence of traded and non-traded capital allows to have firms facing constant returns to scale but to have decreasing returns to scale at city level, and then the presence of a finite number of firms in each city. We assume a Cobb-Douglas production function where β and γ represent the share of labour and traded capital into the production function. The problem of the firms in the production sector follows.

$$\max_{N,K} \quad A_m N^{\beta} K^{\gamma} \bar{Z}_m^{1-\beta-\gamma} - w_m N - K$$

The solution of this problem gives rise to the labour demand condition reported in logarithm terms in Equation 5.

$$log\,(w_m) = log\left(\beta\right) + \frac{\gamma}{1-\gamma} log\left(\gamma\right) + \frac{1}{1-\gamma} log\,(A_m) + \frac{1-\beta-\gamma}{1-\gamma}\left[log\left(\bar{Z}_m\right) - log\left(\bar{N}\right)\right] \tag{5}$$

The last actor of our model are the firms in the construction sector. The representative firm decides how many houses to build (H) in each city and sell them at p^H_m. For each house the construction firm decides the combination of height (h) and lot size (L) of the house to build, such that $H = h \times L$. The quantity of land used should not exceed the potential spread of the city given by geographical or political constraints ($\bar{L}_m$). The cost of the land is p^L_m. In addition to the land cost, the cost of building, *C(H)*, depends on the height of the housing unit to build. The cost of building high is assumed to be convex: adding one more floor to a house lead to a more than proportional increase in construction costs. This assumption is parametrized imposing $\delta \geq 1$.[35] While δ refers to the current technology to build higher, which is common across cities, c_m refers to a city-specific factor that influence the cost of height. The problem of the firms in the construction sector follows.

$$\begin{aligned}
&\max_{H,L} \quad p^H_m H - C(H) \\
&\text{s.t.} \quad H = h \times L; \qquad\qquad L \leq \bar{L}_m \\
&\qquad\;\; C(H) = c_m h^{\delta} L + p^L_m L; \qquad \delta > 1
\end{aligned}$$

The solution of this problem gives rise to the height demand condition reported in logarithm terms in Equation 6.

$$log(h) = \frac{1}{1-\delta}\left[log(c_m) + log(\delta)\right] + \frac{1}{\delta - 1} log\left(p^H_m\right) \tag{6}$$

Markets should clear in equilibrium. The amount of labour hired by the firms in production sector should be equal to the total population of the city (N_m). The demand for consumption good equals its supply. Moreover, the housing market equilibrium requires that the total supply of houses, $h\bar{L}_m$, equalizes its demand, HN. From the housing market equilibrium we can obtain the price equation reported in logs in Equation 7.

$$log\left(p^H_m\right) = \frac{1}{\delta}\left[log(c_m) + log(\delta)\right] + \frac{\delta - 1}{\delta}\left[log(\alpha) + log(w_m) + log(N_m) - log\left(\bar{L}_m\right)\right] \tag{7}$$

Using the spatial indifference (Equation 4), labour demand (Equation 5), and house price equations (Equation 7) it is possible to derive three structural equations, denoted with *, that model the behaviour of house prices, wages and city population in function of city-specific parameters: productivity (A_m) and amenities (θ_m). Let's define K^P_m, K^w_m, and K^N_m as constant terms that influences house prices, wages and population respectively without passing through productivity and amenities. These constant terms also include the effect of non-traded capital, the indirect utility value and the potential land spread of the city given by geographical constraints.[36] The three structural equations are reported in Equations 8 to 10.

[35] Evidence for this assumption has been obtained by Ahlfeldt and McMillen (2015). In fact, they conclude that a reasonable value for δ is 2.7, which is the inverse of elasticity of building height with respect to land prices.

[36] We can derive the theoretical predictions of the effect of $\bar{L}$. More potential land of the cities decreases house prices by increasing housing supply available in the city. Bigger land in the city increases population and then it decreases wages.

$$log\left(p_m^{H*}\right) = K_m^P + \frac{(\delta-1)\left[log\left(A_m\right) + \beta log\left(\theta_m\right)\right]}{\delta\left(1-\beta-\gamma\right) + \alpha\beta\left(\delta-1\right)} \tag{8}$$

$$log\left(w_m^*\right) = K_m^w + \frac{\alpha\left(\delta-1\right) log\left(A_m\right) - \delta\left(1-\beta-\gamma\right) log\left(\theta_m\right)}{\delta\left(1-\beta-\gamma\right) + \alpha\beta\left(\delta-1\right)} \tag{9}$$

$$log\left(N_m^*\right) = K_m^N + \frac{\left[\delta-\alpha\left(\delta-1\right)\right] log\left(A_m\right) + \delta\left(1-\gamma\right) log\left(\theta_m\right)}{\delta\left(1-\beta-\gamma\right) + \alpha\beta\left(\delta-1\right)} \tag{10}$$

The theoretical predictions of our model are that house prices increases if a city becomes more productive or increases its amenities. Wages are positively affected by productivity. A decrease in amenities in the city should be compensated by higher wages in order to equalize utility at each location, ceteris paribus. Finally, city population increases city productivity and amenities.

We assume that violent crime (VC_m^{cc}) can have externality effects on city productivity and amenities that are not taken into account by actors.[37] We assume that the exogenous part of violent crime ($\widehat{VC}_m^{cc}$) has a log-linear influence on these parameters. We define λ^A and λ^θ as the reduced form elasticities of productivity and amenities with respect to violent crime, respectively. We assume that city productivity and amenities are further influenced by constant terms (K_m^A and K_m^θ) and any other non-constant factor not related to violent crime which composes errors (μ_m^A and μ_m^θ).

$$log\left(A_m\right) = K_m^A + \lambda^A log\left(\widehat{VC}_m^{cc}\right) + \mu_m^A \tag{11}$$

$$log\left(\theta_m\right) = K_m^\theta + \lambda^\theta log\left(\widehat{VC}_m^{cc}\right) + \mu_m^\theta \tag{12}$$

We are agnostic about the direction of the effects of violent crime on city-specific parameters. Violent crime can in principle decrease city amenities because people are not willing to live in a city with more crimes. Violent crimes can decrease productivity by influencing human capital accumulation of the population.

Our objective is then to obtain the direction of the effects of violent crime on city-specific parameters that are in line with the estimated regressions of the city-wide effects of violent crime on house prices, wages and city population using the strategy proposed by Glaeser (2008). In order to do this we substitute Equations 11 to 12 into the structural equations (Equations 8 to 10) to obtain the reduced form equations that links violent crime to house prices, wages and city population (Equations 13 to 15). As it is possible to see these equations do not depend anymore on productivity and amenities.

$$log\left(p_m^{*H}\right) = K_m^P + B^P log\left(\widehat{VC}_m^{cc}\right) + \mu_m^P \tag{13}$$

$$log\left(w_m^*\right) = K_m^w + B^w log\left(\widehat{VC}_m^{cc}\right) + \mu_m^w \tag{14}$$

[37] The model can be expanded to include the effect of violent crime on total land spread because for example it can influences zoning and regulation constraints. However, no result that we derive depends on the assumption of no effect of violent crime on total land spread. This is similar to what has been done in Harari (2015)

$$log\left(N_m^*\right) = K_m^N + B^N log\left(\widehat{VC}_m^{cc}\right) + \mu_m^N \tag{15}$$

The reduced form elasticities of house prices, wages and city population with respect to violent crime (the B coefficients) are reported in Equations 16 to 18. These reduced forms coefficients depend on a set of parameters and on the reduced form elasticities of productivity and amenities with respect to violent crime (the λ parameters). In particular, we have a set of three equations (B^P, B^w and B^N) and two unknowns (λ^A and λ^θ). If we know the reduced form elasticities of house prices, wages and city population with respect to violent crime we can potentially recover the reduced form elasticities of productivity and amenities with respect to violent crime.

$$B^P = \frac{(\delta-1)\lambda^A + \beta(\delta-1)\lambda^\theta}{\alpha\beta(\delta-1) + \delta\left(1-\beta-\gamma\right)} \tag{16}$$

$$B^w = \frac{(\delta-1)\alpha\lambda^A - \delta\left(1-\beta-\gamma\right)\lambda^\theta}{\alpha\beta(\delta-1) + \delta\left(1-\beta-\gamma\right)} \tag{17}$$

$$B^N = \frac{[\delta(1-\alpha)+\alpha]\lambda^A + \delta\left(1-\gamma\right)\lambda^\theta}{\alpha\beta(\delta-1) + \delta\left(1-\beta-\gamma\right)} \tag{18}$$

The strategy proposed by Glaeser (2008) requires to first regress house prices, wages and population of MSA on violent crime. In this way it is possible to estimate $\hat{B}^P$, $\hat{B}^w$, and $\hat{B}^N$, the reduced form elasticities of house prices, wages and city population with respect to violent crime. In order to obtain the effect of the exogenous part of violent crime on prices, wages and population we estimate Equations 13 to 15 using our instrumental variable strategy. Moreover, we proxy the constant terms (K_m^P, K_m^w, and K_m^N) by MSA and year fixed effects. The fixed spread of cities is captured by the MSA fixed effects. Moreover, we have demonstrated in Section 4.4 that our instrument is not related to the area of a city. Fixed effects at city level also captures non-traded capital. We also include Census Region time trends in our specification.[38]

Once we estimate $\hat{B}^P$, $\hat{B}^w$, and $\hat{B}^N$ we can recover the effect of violent crime on productivity and amenities (λ_A and λ_θ) that rationalizes its city-wide effects using Equations 19 and 20.

$$\lambda^A = \left(1-\beta-\gamma\right)\hat{B}^N + \left(1-\gamma\right)\hat{B}^w \tag{19}$$

$$\lambda^\theta = \alpha\hat{B}^P - \hat{B}^w \tag{20}$$

In our model we assume that agents are not heterogeneous. There exists an important literature explaining why people sort in different locations based on on their productivity and valuation of amenities (see Combes et al., 2008). We do not suspect that this important mechanism can bias our estimations since we have previously demonstrated that the change in violent crime did not alter the distribution of people between cities. Therefore, the estimated effects of violent crime on productivity and amenities can be interpreted as the effect which is not caused by sorting.

[38] Because we compare cities inside the same Census regions our assumption of perfect mobility of people is more likely to be satisfied.

We assume some parameters in order to obtain the effect of violent crime on productivity and amenities. We approximate the share of housing in utility (α) by the total consumption expenditure to housing to be 0.3, obtained from the U.S. B.L.S. Consumption Expenditure Survey. As for Glaeser (2008), we assume the share of labour and traded capital in production function (β and γ) to be 0.6 and 0.3, respectively.

7.2 Estimating the externality effects

The strategy described in the previous section allows us to map reduced form elasticities of the effect of violent crime on observable variables, such as house prices, income and population in the MSA, on reduced form elasticities of the effect of violent crime on unobservable variables, such as productivity and amenities. That is, we can understand how violent crime might have changed these unobservable variables in a way that is consistent with the model reported in the previous section.

The effect of violent crime on the observable variables is reported in Table 15. These results are slightly different from the results obtained in Section 7 because we need to use a log-log specification in order to link these reduced forms to considerations about the externality effects of violent crime on productivity and amenities. However, we can still conclude that violent crime has a positive effect on city income, and house prices, leaving population unchanged.

[INSERT TABLE 15 HERE]

Using Equations 19 to 20, we can recover how productivity and amenities should have changed in order to rationalize the effect of violent crime on house prices, income and population in the MSA. The calculated externality effects of violent crime on productivity and amenities (λ^A and λ^θ) are reported in Table 16, column (2).[39] We find a negative and significant effect of violent crime on amenities (λ^θ) and a positive effect on productivity (λ^A), with an elasticity of -0.0991 and 0.1243 respectively. Using the theoretical predictions of our model if violent crime might affect amenities as a result income should have increased as we found in Table 15. However, if amenities would have been the only city-specific factor changing this should have been reflected in lower house prices, an effect that we do not observe. Therefore, productivity should have necessarily increased as a result of higher violent crimes.

[INSERT TABLE 16 HERE]

The estimates in Table 20 imply that the increase in violent crime in the city center between 1960 and 1991 led to a decrease in the average amenities of the city of 23.2%.[40] Using equation 4 we can see that this quantity corresponds to the percentage of the wage people would have been willing to sacrifice in 1991 to return to violent crime rates in the city center as in 1960.

---

[39] Standard errors of the coefficients have been obtained using the following bootstrap technique. This strategy consists in first bootstrapping a panel sample from our distribution of observations; subsequently, we have estimated the elasticity of house prices, income and population to violent crime and then compute the corresponding elasticity of amenities and productivity to violent crime; we have replicated this procedure several times and obtained a distribution of these parameters and relevant standard errors.

[40] Log violent crime in 1960 and 1991 was -7.08 and -4.73, respectively. We have computed the change in average amenities between 1960 and 1991 multiplying the elasticity of amenities to violent crime, -0.099, to the change in log violent crime in that period, 2.34.

We now explore if these effects are due directly because of the effect of violent crime or by the effect of other variables that were affected by violent crime and ultimately influenced productivity and amenities. In order to separate the real direct effect of violent crime on city-specific factors we control for several possible mediating factors that we have found to be influenced by violent crime: highways, residential suburbanization and employment centralization. The results are reported in Table 16, columns (3) to (5), respectively. By controlling for these variables, we can partial out the externality effect of violent crime which is not passing through these channels.

From Column (3) of Table 16 we can infer that, despite cities with more violent crime lead to the creation of more highways, this last effect cannot explain why cities have higher levels of productivity. The effect of violent crime on productivity which is not passing through residential suburbanization is still positive, as it has been shown in Column (4) of Table 16. Moreover, controlling for the effect of violent crime on residential suburbanization we find a negative and stronger effect of violent crime on city-amenities, with a calculated elasticity (λ^{θ}) of -0.1068. That is, if city centers increase their violent crimes and people cannot suburbanize then the impact of violent crime on amenities is negative.

Finally, when we control for employment decentralization we do not find any externality effect of violent crime on unobservable variables (see Column (5) of Table 16). The elasticity of violent crime on productivity (λ^{A}) is now not significant and equal to 0.1887. That is, the previous positive effect of violent crime on productivity could be explained by the fact that violent crime led to a creation of different employment centers in the city and this could boost firm productivity. This might mean that some firms were inefficiently located in the city center. Violent crime displaced firms in the suburbs, and this displacement was productivity-enhancing. One possible explanation for this result is that historically it might have been more efficient to have firms in the city center. Modern cities might have a different optimal distribution of firms in the city, for example as predicted by Lucas and Rossi-Hansberg (2002).[41] For an individual firm it might not be optimal to move outside the city center, despite the fact that productivity could increase if all firms coordinated this move outside of the city center. Crime could have acted as a tax to firm location in city center that solved this coordination failure by making people and firms move.

The coefficient of the effect of violent crime on amenities also turns to be insignificant. The calculated elasticity of violent crime on amenities (λ^{θ}) is now -0.164. This might be explained by the fact that part of the decrease in amenities in the city is explained by the fact that violent crime lead to employment to be located further away from the center. As a result amenities could potentially decrease because of higher traffic and congestion externalities.

8 Concluding remarks

In this paper we provide evidence of a debated mechanism that can explain why U.S. cities suburbanized between the 1960s and the 1990s: the increase in violent crime in

[41] Lucas and Rossi-Hansberg (2002) predict that cities in equilibrium should have a business center in the CBD, and further away: residential, business, mixed use, business, and residential areas respectively

city centers. We estimate the causal effect of crime on suburbanization exploiting a new instrument which combines time variation from national level of past lead poisoning and geographical variation from local soil quality. We find that an increase in one standard deviation in violent crime decreased the share of population living in the city center by 8 percentage points. We provide counterfactual evidence that if violent crime in city centers would have increased at the same rate as the suburbs then the proportion of people living in city centers in the U.S. would have been constant between 1960 and 1990.

The advantage of our empirical methodology is to be able to compare all the cities in the U.S. for many decades by exploiting a standardized measure, such as the pH of the soil of a city. The use of lead poisoning as time variation of our instrument has the convenience that it can be employed to predict both the big rise of American crimes between the 1960s and the 1990s and the fall afterwards. More micro-evidence should be provided to show the link between resuspended lead and blood lead levels. This can contribute to the discussion about how much of the crime variation in the U.S. can be explained by lead poisoning.

Further research should be dedicated to the study of the big fall of crimes after the 1990s. In particular, it is crucial to understand why suburbanization did not revert when crimes decreased and what are the mechanisms behind the persistence in suburbanization. One possibility is that the flight from city centers affected amenities in the suburbs by increasing school and housing quality.

Additionally, the interaction between lead poisoning and soil quality can potentially provide quasi-experimental variation that can be used to understand the effect of crime on many other outcomes. Furthermore, our methodology could be also easily applicable to other countries. Expanding the context of analysis to European countries it might possible to understand the importance of several other urban amenities and characteristics to explain suburbanization.

The results we find are important in order to understand how much urban amenities and productivity can explain location of people and firms inside cities. Using a spatial equilibrium model we infer that the increase of violent crimes created important spillover effects at city level by reducing the overall level of city amenities. Inequalities in cities with higher crimes also increased and racial location segregation happened because of white people moving to the suburbs. The model we exploit in this paper can potentially be expanded in the future in order to assess how amenities react differently between suburbs and city centers after violent crimes increased. Another possible extension of the model is to make the land spread of the city endogenous to disentangle the different theoretical effects of crime and suburbanization.

We provide suggestive evidence that the job decentralization caused by higher violent crimes in the city center is the responsible for increasing city productivity. This is consistent with a situation where it is optimal for firms to move to the suburbs but they do not because of coordination failures. Violent crime potentially provide a common shock that solves this sub-optimality. This result points to the fact that cities could achieve gains in productivity by using incentives to move firms to the suburbs. Further research should be devoted in incorporating this coordination failure in our model and show the gains of different employment centers inside cities.

References

Ahlfeldt, G. M. and D. P. McMillen (2015). The vertical city: The price of land and the height of buildings in chicago 1870-2010. Technical report, Spatial Economics Research Centre, LSE.

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015). The economics of density: Evidence from the berlin wall. *Econometrica 83*(6), 2127–2189.

Aizer, A. and J. Currie (2017). Lead and juvenile delinquency: New evidence from linked birth, school and juvenile detention records. Technical report, National Bureau of Economic Research.

Aizer, A., J. Currie, P. Simon, and P. Vivier (2016). Do low levels of blood lead reduce children's future test scores? Technical report, National Bureau of Economic Research.

Allen, T., C. Arkolakis, and X. Li (2015). Optimal city structure. *Yale University, mimeo.*

Alonso, W. (1964). Location and land use. toward a general theory of land rent. *Location and land use. Toward a general theory of land rent.*.

Angel, S., J. Parent, D. L. Civco, A. Blei, and D. Potere (2011). The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050. *Progress in Planning 75*(2), 53–107.

Athey, S. and G. Imbens (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspective 31*(2), 3–32.

Banks, E. C., L. E. Ferretti, and D. Shucard (1996). Effects of low level lead exposure on cognitive function in children: a review of behavioral, neuropsychological and biological evidence. *Neurotoxicology 18*(1), 237–281.

Baum-Snow, N. (2007). Did highways cause suburbanization? *The Quarterly Journal of Economics 122*(2), 775–805.

Baum-Snow, N. and D. Hartley (2016). Accounting for Central Neighborhood Change, 1980-2010. Working Paper Series WP-2016-9, Federal Reserve Bank of Chicago.

Baum-Snow, N. and B. F. Lutz (2011). School desegregation, school choice, and changes in residential location patterns by race. *The American economic review 101*(7), 3019–3046.

Bianchi, M., P. Buonanno, and P. Pinotti (2012). Do immigrants cause crime? *Journal of the European Economic Association 10*(6), 1318–1347.

Billings, S. B. and K. T. Schnepel (2017, July). Life After Lead: Effects of Early Interventions for Children Exposed to Lead. IZA Discussion Papers 10872, Institute for the Study of Labor (IZA).

Boustan, L. P. (2010). Was Postwar Suburbanization "White Flight"? Evidence from the Black Migration. *The Quarterly Journal of Economics 125*(1), 417–443.

Bowen, N. K. and G. L. Bowen (1999). Effects of crime and violence in neighborhoods and schools on the school behavior and performance of adolescents. *Journal of Adolescent Research 14*(3), 319–342.

Bradford, D. F. and H. H. Kelejian (1973). An econometric model of the flight to the suburbs. *Journal of Political Economy 81*(3), 566–589.

Breiman, L., J. Friedman, C. J. Stone, and R. A. Olshen (1984). *Classification and regression trees.* CRC press.

Brueckner, J. K. and S. S. Rosenthal (2009). Gentrification and Neighborhood Housing Cycles: Will America's Future Downtowns Be Rich? *The Review of Economics and Statistics 91*(4), 725–743.

Canfield, R. L., C. R. Henderson Jr, D. A. Cory-Slechta, C. Cox, T. A. Jusko, and B. P. Lanphear (2003). Intellectual impairment in children with blood lead concentrations below 10 μg per deciliter. *New England journal of medicine 348*(16), 1517–1526.

Cecil, K. M., C. J. Brubaker, C. M. Adler, K. N. Dietrich, M. Altaye, J. C. Egelhoff, S. Wessel, I. Elangovan, R. Hornung, K. Jarvis, et al. (2008). Decreased brain volume in adults with childhood lead exposure. *PLoS Med 5*(5), e112.

Chalfin, A. and J. McCrary (2017). Criminal Deterrence : A Review of the Literature. *Journal of Economic Literature 55*(1).

Chandramouli, K., C. D. Steer, M. Ellis, and A. M. Emond (2009). Effects of early childhood lead exposure on academic performance and behaviour of school age children. *Archives of Disease in Childhood 94*(11), 844–848.

Clark, H. F., D. J. Brabander, and R. M. Erdil (2006). Sources, sinks, and exposure pathways of lead in urban garden soil. *Journal of Environmental Quality 35*(6), 2066–2074.

Clay, K., W. Troesken, and M. Haines (2014). Lead and mortality. *Review of Economics and Statistics 96*(3), 458–470.

Combes, P.-P., G. Duranton, and L. Gobillon (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics 63*(2), 723–742.

Cullen, J. B. and S. D. Levitt (1999). Crime, Urban Flight, And The Consequences For Cities. *The Review of Economics and Statistics 81*(2), 159–169.

Daniele, V. and U. Marani (2011). Organized crime, the quality of local institutions and FDI in Italy: A panel data analysis. *European Journal of Political Economy 27*(1), 132–142.

Dapul, H. and D. Laraque (2014). Lead poisoning in children. *Advances in pediatrics 61*(1), 313–333.

Denno, D. W. (1990). *Biology and violence: From birth to adulthood.* Cambridge University Press.

Detotto, C. and E. Otranto (2010). Does crime affect economic growth? *Kyklos 63*(3), 330–345.

Diamond, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980–2000. *The American Economic Review 106*(3), 479–524.

Feigenbaum, J. J. and C. Muller (2016). Lead exposure and violent crime in the early twentieth century. *Explorations in Economic History 62*, 51–86.

Ferrie, J. P., K. Rolf, and W. Troesken (2012). Cognitive disparities, lead plumbing, and water chemistry: Prior exposure to water-borne lead and intelligence test scores among world war two us army enlistees. *Economics & Human Biology 10*(1), 98–111.

Filippelli, G. M., M. A. Laidlaw, J. C. Latimer, and R. Raftis (2005). Urban lead poisoning and medical geology: an unfinished story. *GSA Today 15*(1), 4–11.

Frey, W. H. (1979). Central city white flight: Racial and nonracial causes. *American Sociological Review*, 425–448.

Fujita, M. and H. Ogawa (1982). Multiple equilibria and structural transition of non-monocentric urban configurations. *Regional science and urban economics 12*(2), 161–196.

Garcia-Lopez, M.-A., A. Holl, and E. Viladecans-Marsal (2015). Suburbanization and highways in Spain when the Romans and the Bourbons still shape its cities. *Journal of Urban Economics 85*(C), 52–67.

Garcia-Lopez, M.-A., I. Pasidis, and E. Viladecans-Marsal (2015). Express delivery to the suburbs? transportation effects in heterogeneous european cities.

Glaeser, E. L. (2008). *Cities, Agglomeration and Spatial Equilibrium*. New York: Oxford University Press.

Glaeser, E. L. and J. D. Gottlieb (2009). The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States. *Journal of Economic Literature 47*(4), 983–1028.

Glaeser, E. L. and M. E. Kahn (2001). Decentralized employment and the transformation of the american city. Technical report, National Bureau of Economic Research.

Glaeser, E. L. and M. E. Kahn (2004). Sprawl and urban growth. *Handbook of regional and urban economics 4*, 2481–2527.

Glaeser, E. L. and B. Sacerdote (1999). Why is there more crime in cities? *Journal of political economy 107*(S6), S225–S258.

Goldin, C. and L. F. Katz (2008). Mass secondary schooling and the state: the role of state compulsion in the high school movement. In *Understanding long-run economic growth: Geography, institutions, and the knowledge economy*, pp. 275–310. University of Chicago Press.

Goudriaan, H., K. Wittebrood, and P. Nieuwbeerta (2006). Neighbourhood characteristics and reporting crime: Effects of social cohesion, confidence in police effectiveness and socio-economic disadvantage. *British Journal of Criminology 46*(4), 719–742.

Grönqvist, H., J. P. Nilsson, and P.-O. Robling (2016). Early lead exposure and outcomes in adulthood.

Grubb, W. N. (1982). The flight to the suburbs of population and employment, 1960–1970. *Journal of Urban Economics 11*(3), 348–367.

Harari, M. (2015). Cities in bad shape: Urban geometry in india. Job market paper, Massachusetts Institute of Technology.

Harnik, P., A. Martin, and K. Barnhart (2015). City park facts. *Report. The Trust for Public Land, Washington, DC, USA*.

Henrich, C. C., M. Schwab-Stone, K. Fanti, S. M. Jones, and V. Ruchkin (2004). The association of community violence exposure with middle-school achievement: A prospective study. *Journal of Applied Developmental Psychology 25*(3), 327–348.

Hunt, A., D. L. Johnson, and D. A. Griffith (2006). Mass transfer of soil indoors by track-in on footwear. *Science of the Total Environment 370*(2), 360–371.

International Programme on Chemical Safety (1995). Environmental health criteria 165: Inorganic lead. Technical report, World Health Organization Geneva.

Jargowsky, P. A. and Y. Park (2009). Cause or consequence? suburbanization and crime in us metropolitan areas. *Crime & Delinquency 55*(1), 28–50.

Kelly, M. (2000). Inequality and Crime. *Review of Economics and Statistics 82*(4), 530–539.

Kopecky, K. A. and R. M. Suen (2010). A quantitative analysis of suburbanization and the diffusion of the automobile. *International Economic Review 51*(4), 1003–1037.

Laidlaw, M. A. and G. M. Filippelli (2008). Resuspension of urban soils as a persistent source of lead poisoning in children: a review and new directions. *Applied Geochemistry 23*(8), 2021–2039.

Lucas, R. E. and E. Rossi-Hansberg (2002). On the internal structure of cities. *Econometrica 70*(4), 1445–1476.

Ludwig, J. (1998). Concealed-gun-carrying laws and violent crime: evidence from state panel data. *International Review of Law and Economics 18*(3), 239–254.

Mielke, H. (1999). Lead in the inner cities policies to reduce children's exposure to lead may be overlooking a major source of lead in the environment. *American scientist 87*(1), 62–73.

Mieszkowski, P. and E. S. Mills (1993). The causes of metropolitan suburbanization. *The Journal of Economic Perspectives 7*(3), 135–147.

Mills, E. S. (1972). Studies in the structure of the urban economy.

Mills, E. S. and L. S. Lubuele (1997). Inner cities. *Journal of Economic Literature 35*(2), 727–756.

Moretti, E. (2011). Local labor markets. *Handbook of labor economics 4*, 1237–1313.

Muth, R. (1969). Cities and housing: The spatial patterns of urban residential land use. *University of Chicago, Chicago*.

National Archive of Criminal Justice Data (2006). Law enforcement agency identifiers crosswalk [united states], 2005 [computer file]. icpsr04634-v1.

Needleman, H. L., C. McFarland, R. B. Ness, S. E. Fienberg, and M. J. Tobin (2002). Bone lead levels in adjudicated delinquents: a case control study. *Neurotoxicology and teratology 24*(6), 711–717.

Needleman, H. L., J. A. Riess, M. J. Tobin, G. E. Biesecker, and J. B. Greenhouse (1996). Bone lead levels and delinquent behavior. *JAMA: The Journal of the American Medical Association 275*(5), 363–369.

Nigg, J. T., M. Nikolas, G. Mark Knottnerus, K. Cavanagh, and K. Friderici (2010). Confirmation and extension of association of blood lead with attention-deficit/hyperactivity disorder (adhd) and adhd symptom domains at population-typical exposure levels. *Journal of Child Psychology and Psychiatry 51*(1), 58–65.

Nriagu, J. O. (1990). The rise and fall of leaded gasoline. *Science of the total environment 92*, 13–28.

O'Flaherty, B. and R. Sethi (2015). *Urban Crime*, Volume 5 of *Handbook of Regional and Urban Economics*, Chapter 0, pp. 1519–1621. Elsevier.

Patterson, C. C. (1965). Contaminated and natural lead environments of man. *Archives of Environmental Health: An International Journal 11*(3), 344–360.

Peryea, F. J. (2001). *Gardening on lead-and arsenic-contaminated soils*. Washington State University Cooperative Extension.

Rabin, R. (2008). The lead industry and lead water pipes "a modest campaign". *American journal of public health 98*(9), 1584–1592.

Reber, S. J. (2005). Court Ordered Desegregation: Successes and Failures Integrating American Schools since Brown versus Board of Education. *Journal of Human Resources 40*(3).

Reddy, K. J., L. Wang, and S. P. Gloss (1995). Solubility and mobility of copper, zinc and lead in acidic environments. *Plant and Soil 171*(1), 53–58.

Reyes, J. W. (2015a). Lead exposure and behavior: Effects on antisocial and risky behavior among children and adolescents. *Economic Inquiry 53*(3), 1580–1605.

Reyes, J. W. (2015b). Lead policy and academic performance: Insights from massachusetts. *Harvard Educational Review 85*(1).

Roback, J. (1982). Wages, rents, and the quality of life. *The Journal of Political Economy*, 1257–1278.

Roper, W. L., V. Houk, H. Falk, and S. Binder (1991). Preventing lead poisoning in young children: A statement by the centers for disease control, october 1991. Technical report, Centers for Disease Control, Atlanta, GA (United States).

Rosen, S. (1979). Wage-based indexes of urban quality of life. *Current issues in urban economics 3*.

Rowe, D. C. (2002). *Biology and crime*. Roxbury Los Angeles, CA.

Sampson, R. J. and A. S. Winter (2016). The racial ecology of lead poisoning: Toxic inequality in chicago neighborhoods, 1995-2013. *Du Bois Review: Social Science Research on Race 13*(2), 261–283.

Stehouwer, R. and K. Macneal (1999). Lead in residential soils: Sources, testing, and reducing exposure.

Stretesky, P. B. and M. J. Lynch (2001). The relationship between lead exposure and homicide. *Archives of pediatrics & adolescent medicine 155*(5), 579–582.

Toscano, C. D. and T. R. Guilarte (2005). Lead neurotoxicity: from exposure to molecular effects. *Brain Research Reviews 49*(3), 529–554.

Troesken, W. (2008). Lead water pipes and infant mortality at the turn of the twentieth century. *Journal of Human Resources 43*(3), 553–575.

United States Bureau of Mines, U. B. o. M. Minerals yearbook: Lead, 1941-1986.

United States Department of Agriculture, Soil Survey Staff, N. R. C. S. Web soil survey. http://websoilsurvey.nrcs.usda.gov/.

United States Department of Justice, F. B. I. Uniform crime reporting program data [united states]: Offenses known and clearances by arrest, 1960-2014.

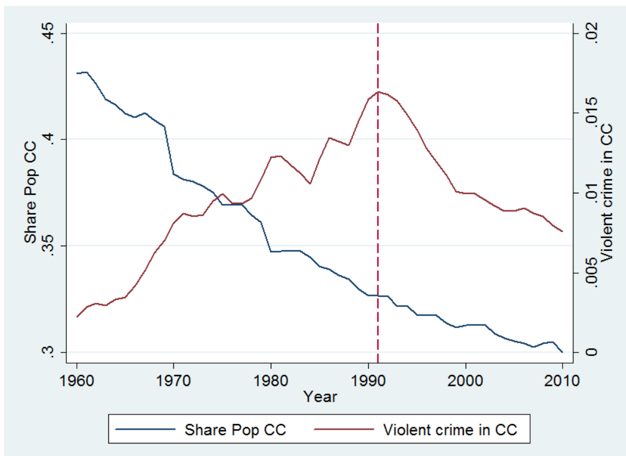
United States Department of Justice, F. B. I. (1993). Age-specific arrest rates and race-specific arrest rates for selected offenses, 1965-1992.

Wright, J. P., K. N. Dietrich, M. D. Ris, R. W. Hornung, S. D. Wessel, B. P. Lanphear, M. Ho, and M. N. Rae (2008). Association of prenatal and childhood blood lead concentrations with criminal arrests in early adulthood. *PLoS Medicine 5*(5), e101.

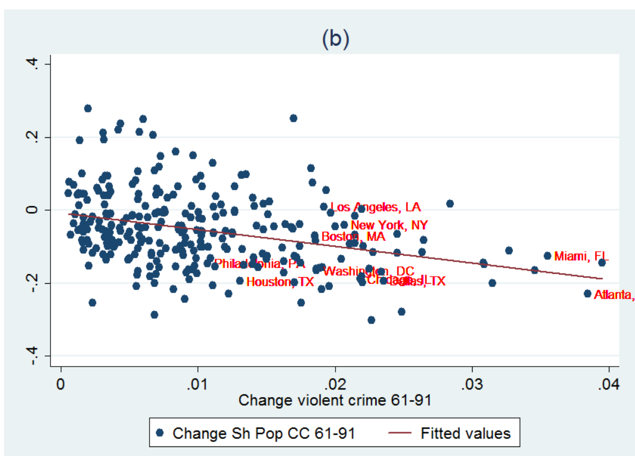
Zahran, S., M. A. Laidlaw, S. P. McElmurry, G. M. Filippelli, and M. Taylor (2013). Linking source and effect: Resuspended soil lead, air lead, and children's blood lead levels in detroit, michigan. *Environmental science & technology 47*(6), 2839–2845.

9 Figures and Tables

9.1 Figures



(a) Time series



(b) Scatter plot

Figure 1: Correlation between violent crime per capita and proportion of people living in the city center

Panel a): Time series share population living in city center (CC), left y axis, and violent crime per capita in city center, right y axis. Panel b): Scatter plot change share population living in city center (CC) between 1961 and 1991 against change violent crime per capita in city center.

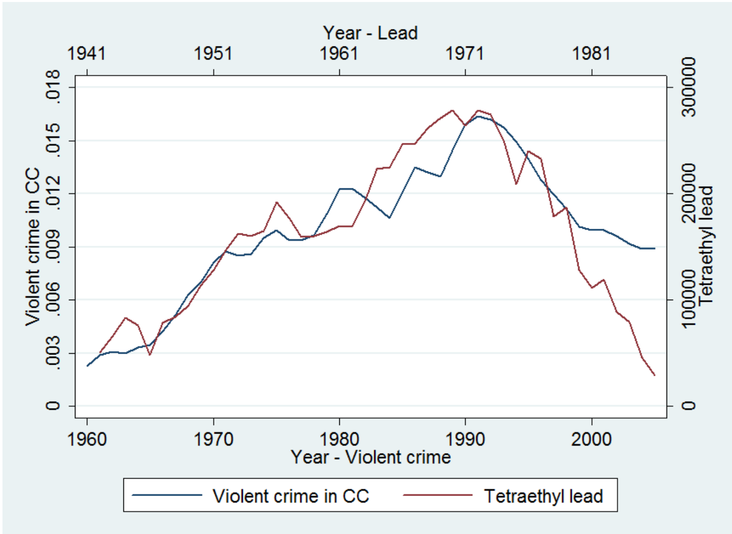


Figure 2: Time series of violent crime rates per capita and the consumption of tetraethyl lead 19 years before

Upward x axis: time series of tonnes of lead consumed in U.S. as gasoline additive. Downward x axis: time series of violent crime per capita in city center

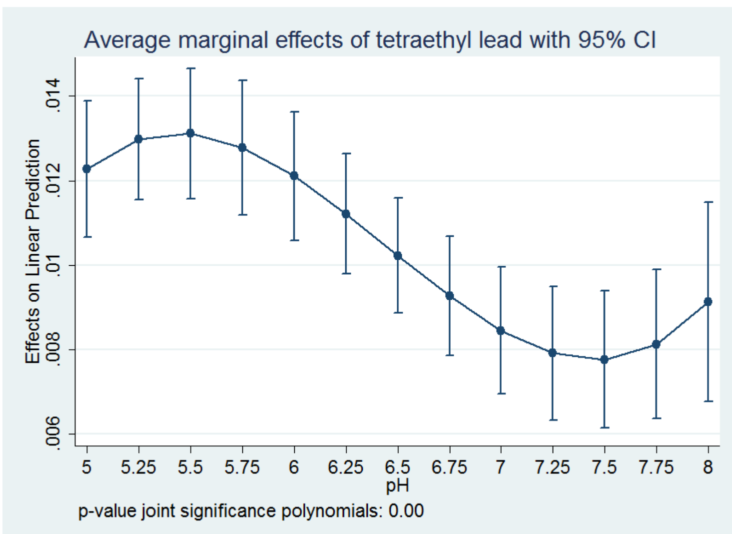
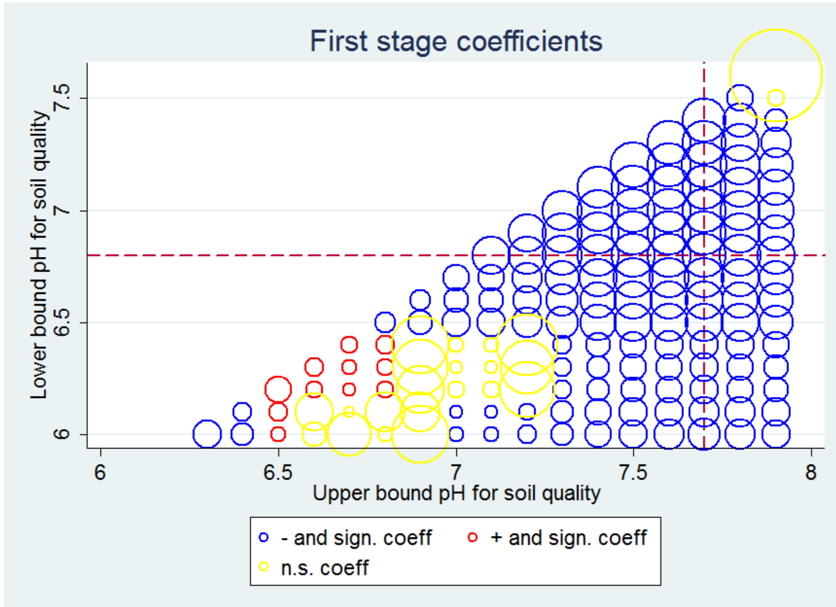
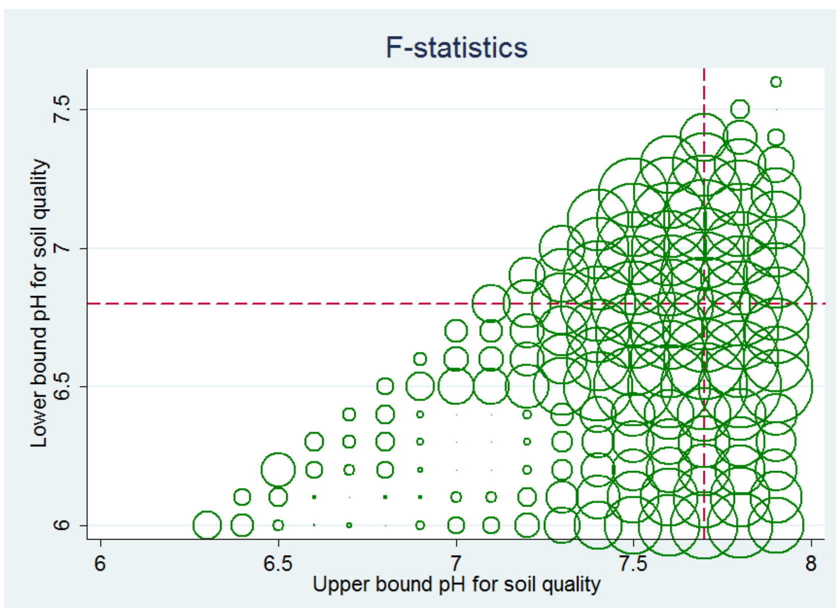


Figure 3: Average marginal effect of tetraethyl lead at different pH values

Marginal effects derived after regressing violent crime per capita in city centers on tetraethyl lead, tetraethyl lead x pH, tetraethyl lead x pH^2, and tetraethyl lead x pH^3. Robust standard errors have been used clustered at city level. Marginal effects reported for value of pH between the 10th and 99th percentile. p-value joint significance polynomials: p-value for the test of joint significance of the coefficients of the following regressors: tetraethyl lead x pH, tetraethyl lead x pH^2, and tetraethyl lead x pH^3. For every city center, we combine data for the average pH level from the United States Geological Survey with crime observations from F.B.I.



(a) Coefficients



(b) F-statistics

Figure 4: Coefficients and F-statistics of the first stage regression for every possible pH interval

Panel a): Coefficients of first stage regression. Panel b): F-statistics of first stage regression. Coefficient and F-statistic of the excluded instrument derived after regressing violent crime per capita on city and year fixed effects and the interaction between tetraethyl lead 19 years before and the soil quality index. Every different circle refers to a different regression for every possible minimum and maximum level of pH. Robust standard errors have been used. The size of the circles refer to the absolute value of the coefficient or F-stat with respect to the coefficient or F-stat in the same category (- and sign.: negative and signficant, + and sign: positive and significant, n.s.: non significant). Dashed lines indicate our chosen soil quality index: pH between 6.8 and 7.7

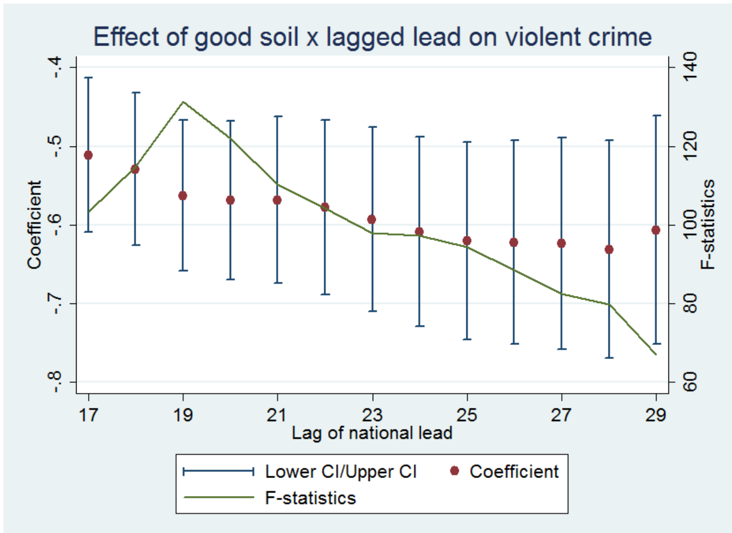


Figure 5: First stage using different lags of national lead poisoning

Left y axis: coefficient of the first stage estimate. Right y axis: F-statistics of the relevance of the instrument. For each possible lag of national lead (X), coefficient obtained after regressing violent crime per capita in the city center on MSA, year and Census division times year fixed effects and the interaction between a dummy taking 1 if pH is between 6.8 and 7.7 and the tonnes of lead consumed in U.S. as gasoline additive X years before, normalized by the maximum level of tetraethyl lead consumption. F-statistics obtained as the F-statistics of the instrument in this regression.

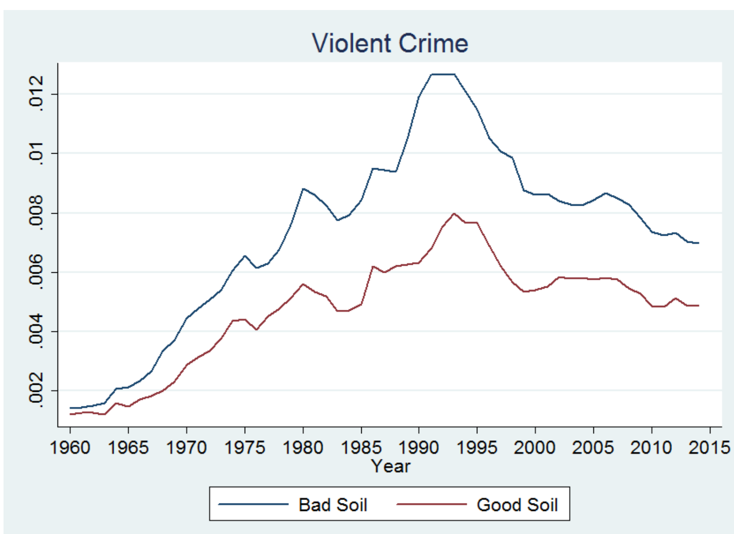


Figure 6: Time series of violent crime per capita in city center by good and bad soil

City centers with good soil: pH between 6.8 and 7.7. City centers with bad soil: pH outside the interval between 6.8 and 7.7.

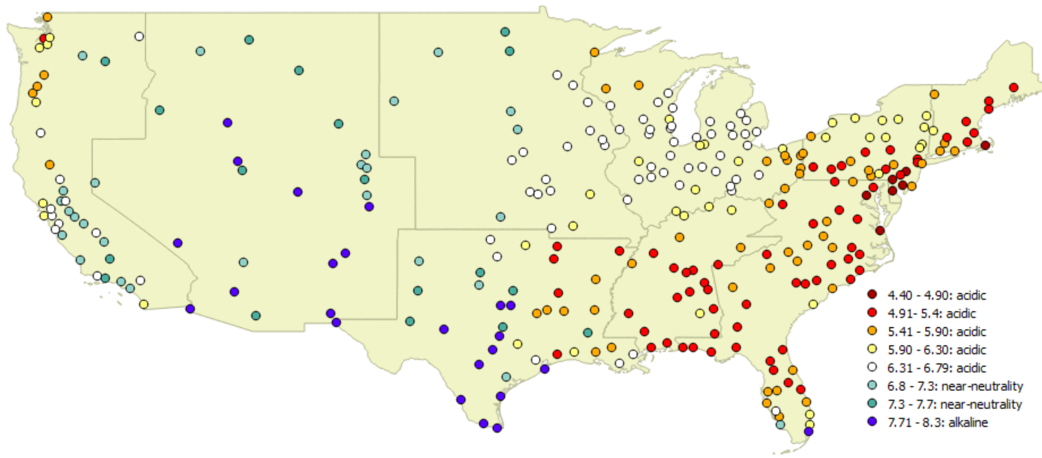


Figure 7: Map of city centers with good and bad soils for lead adsorption

City centers with good soil: pH between 6.8 and 7.7. City centers with bad soil: pH outside the interval between 6.8 and 7.7. The map reports the U.S. Census Divisions

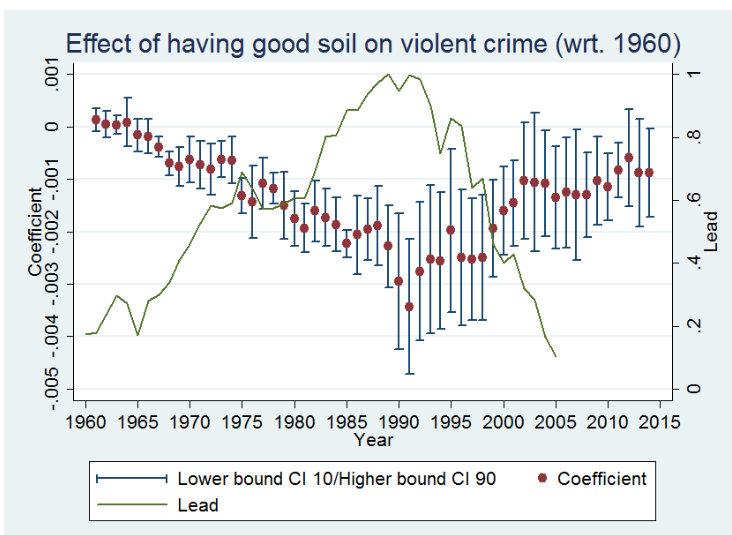


Figure 8: Time series of the effect of the good soil index on crime and time series of lagged tetraethyl lead

Coefficients obtained regressing violent crime per capita on the interaction between the good soil dummy and year dummies, controlling for city, year and Census division times year fixed effects. Standard errors have been clustered at Census Division level. Lower bound CI 10: lower bound confidence interval at 10 % significance level. Higher bound CI 10: higher bound confidence interval at 10 % significance level. City centers with good soil: pH between 6.8 and 7.7. Lead: tonnes of lead consumed in U.S. as gasoline additive 19 years before, normalized by the maximum level of tetraethyl lead consumption. 1960 year dummy has been omitted. Robust standard errors have been used. Left y axis: effect of the good soil index on crime. Right y axis: time series of lagged tetraethyl lead.

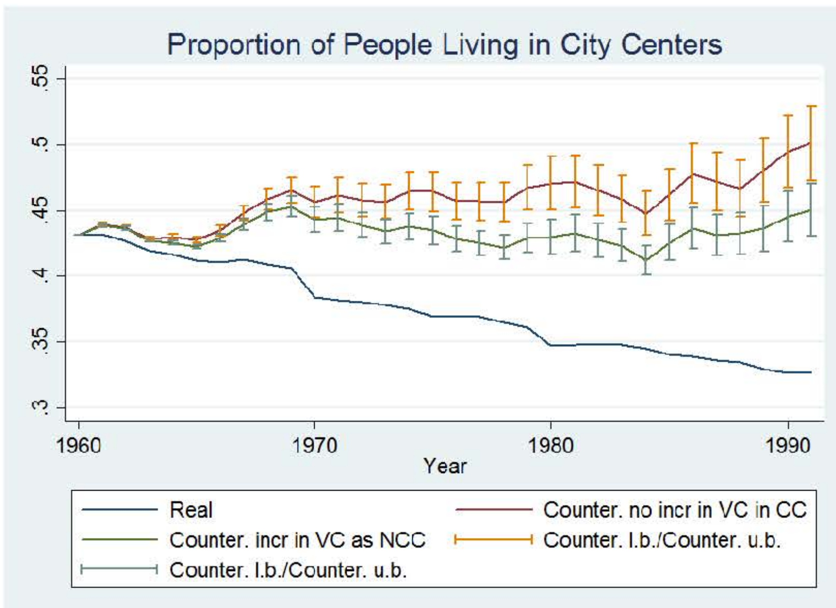


Figure 9: Counterfactual proportion of people in city center in MSA if violent crime in city center would not have increased from 1960 and if violent crime in city center would have increased from 1960 as in the suburbs

Real: actual time series of proportion of people in city center in MSA. Counter. no incr in VC in CC: counterfactual proportion of people in city center in MSA if violent crime (VC) in city center (CC) would not have increased from 1960 with corresponding 95 % confidence interval. Obtained subtracting the actual increase in violent crime from one year in the city center to another multiplied by the causal effect of crime on suburbanization from the actual suburbanization measure. Counter. incr in VC as NCC: counterfactual proportion of people in city center in MSA if violent crime in city center would have increased from 1960 as in the suburbs (NCC) with corresponding 95 % confidence interval. Obtained subtracting the actual difference in change in violent crime from one year to another in the city center with respect to the suburbs multiplied by the causal effect of crime on suburbanization from the actual suburbanization measure. VC: Violent crime per capita. CC (NCC): city center (suburbs). l.b: Lower bound confidence. u.b.: upper bound confidence.

9.2 Tables

Table 1: First stage using different fixed effects and time trends for all the years

VARIABLES	(1) Violent crime	(2) Violent crime	(3) Violent crime	(4) Violent crime	(5) Violent crime
Good soil x Lead	-0.00184***	-0.00528***	-0.00451***	-0.00327***	-0.00276***
	(0.000190)	(0.000326)	(0.000363)	(0.000378)	(0.000400)
Observations	9,515	9,515	9,484	9,484	9,363
R-squared	0.005	0.757	0.763	0.771	0.832
MSA FE	NO	YES	YES	YES	YES
Year FE	NO	YES	YES	YES	YES
C. reg X Year FE	NO	NO	YES	NO	NO
C. div X Year FE	NO	NO	NO	YES	NO
State X Year FE	NO	NO	NO	NO	YES
Year	60-91	60-91	60-91	60-91	60-91
Estimation	OLS	OLS	OLS	OLS	OLS
F		262.16	154.46	74.80	47.78

For the notes see Table 2

Table 2: First stage using different fixed effects and time trends for Census Years

VARIABLES	(1) Violent crime	(2) Violent crime	(3) Violent crime	(4) Violent crime	(5) Violent crime
Good soil x Lead	-0.00206***	-0.00721***	-0.00556***	-0.00396***	-0.00325**
	(0.000611)	(0.00120)	(0.00134)	(0.00140)	(0.00145)
Observations	1,190	1,189	1,185	1,185	1,170
R-squared	0.004	0.744	0.752	0.758	0.831
MSA FE	NO	YES	YES	YES	YES
Year FE	NO	YES	YES	YES	YES
C. reg X Year FE	NO	NO	YES	NO	NO
C. div X Year FE	NO	NO	NO	YES	NO
State X Year FE	NO	NO	NO	NO	YES
Year	CY 60-90	CY 60-90	CY 60-90	CY 60-90	CY 60-90
Estimation	OLS	OLS	OLS	OLS	OLS
F		36.11	17.29	7.99	5.01

Violent crime: Violent crime per capita in the city center. Good soil x Lead: dummy taking 1 if pH in the city center is between 6.8 and 7.7 multiplied by tonnes of lead consumed in U.S. as gasoline additive 19 years before, normalized by the maximum level of tetraethyl lead consumption. C. reg: Census region. C. div: Census division fixed effects. CY: Census year. F: F-statistics on the excluded instrument. Robust standard errors have been used. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Balancing test for economic, social and geographic characteristics

Variable	Average	Levels		Trends	
		All US	Inside Division	All US	Inside Division
Area CC (50)	27.51	3.52	0.98	.	.
Area MSA (50)	2056.36	708.11	-424.82	.	.
Share Pop. CC (50 - 60)	0.47	0.09**	0.02	0.05**	0.01
Population CC (50 - 60)	2.0e+05	-7.7e+04	-1438.35	32129.34*	6486.92
Population MSA (50 - 60)	4.3e+05	-1.7e+05	1842.98	30009.88	40253.98
Pop. Density CC (50 - 60)	6109.20	-1535.46***	30.86	327.39	151.63
Pop. Density MSA (50 - 60)	197.35	-105.09***	-8.35	12.44	27.83
Violent Crime Rate CC (per 10000) (60 - 63)	13.82	-2.17	-0.34	-0.46	0.38
Murder Rate CC (per 10000) (60 - 63)	0.53	-0.11	0.01	-0.05	-0.04
Rape Rate CC (per 10000) (60 - 63)	0.70	0.06	-0.00	0.10	0.15
Robbery Rate CC (per 10000) (60 - 63)	4.71	0.43	-0.15	-0.27	0.19
Agg. Assault Rate CC (per 10000) (60 - 63)	7.89	-2.55**	-0.21	-0.29	0.08
Burglary Rate CC (per 10000) (60 - 63)	59.64	5.28	1.28	-0.54	-2.17
Larceny Rate CC (per 10000)(60 - 63)	169.71	78.02***	29.73	-10.26	-18.26*
Vehicle Theft Rate CC (per 10000) (60 - 63)	22.91	2.75	-1.78	0.24	-0.13
Total Crimes CC (per 10000) (60 - 63)	266.43	83.87***	28.82	-9.56	-17.52
Median Gross Rent (housing unit) MSA (60)	384.91	26.96**	22.12*	.	.
Median Single Family House Value MSA (60)	64628.31	4839.84*	2297.45	.	.
Median Family Income CC (50-60)	22811.72	1995.89***	1246.13*	87.89	-618.41
Median Family Income MSA (50-60)	21400.64	1837.63***	1236.50*	-319.70	-428.57
Annual Precipitation (77)	35.86	-19.38***	-8.86***	.	.
% Possible Sun (77)	59.93	7.91***	2.60	.	.
Average Jan Temp (77)	34.34	-1.49	-3.20	.	.
Average July Temp (77)	75.76	0.25	-0.12	.	.
% Blacks CC (60)	13.41	-7.81***	-1.21	.	.
% Blacks MSA (60)	9.45	-5.47***	-1.49	.	.
% Foreign CC (60)	16.40	-0.41	-0.70	.	.
% Foreign MSA (60)	14.47	0.70	-0.07	.	.
Distance Border or Coast	129.89	139.60***	35.60	.	.
Unemployment Rate MSA (60)	5.17	0.26	0.20	.	.
Labor Force Civilian MSA (50-60)	0.39	-0.01	0.00	0.00	-0.00
Emp. Rate MSA (50-60)	36.99	-1.18	0.26	0.21	-0.40
Emp. Rate Agriculture MSA (50-60)	3.44	0.57	0.54	-0.07	-0.01
Emp. Rate Business Services MSA (50)	2.29	0.50***	0.30**	.	.
Emp. Rate Construction MSA (50-60)	2.51	0.63***	0.19	-0.06	0.09
Emp. Rate Education MSA (60)	2.18	-0.02	-0.16	.	.
Emp. Rate Finance MSA (50-60)	1.17	0.15*	0.01	0.07*	0.03
Emp. Rate Government MSA (60)	1.82	0.29*	-0.03	.	.
Emp. Rate Manufacturing MSA (50-60)	9.18	-5.52***	-1.52**	0.82***	-0.04
Emp. Rate Mining MSA (50)	0.44	0.47	0.52	.	.
Emp. Rate Professional MSA (50)	3.53	0.10	-0.31	.	.
Median Age MSA (50-60)	29.49	-0.93*	-0.29	-1.01***	-0.39
% Over 65y MSA (50-60)	7.81	-0.94**	-0.64	-0.50*	-0.32
% Non-white MSA (50-60)	9.98	-5.17***	-1.77	0.50*	0.40
% Pub. Transportation to Work MSA (60)	6.82	-4.04***	-2.83***	.	.
Median Years of School MSA (50-60)	9.57	1.42***	0.69**	-0.13	-0.01
CC interstate rays (50-60)	0.04	-0.05**	-0.08*	0.17	-0.03
CC total rays (50-60)	0.04	-0.05**	-0.08*	0.17	-0.03
2-digit CC rays (50-60)	0.03	-0.03**	-0.05	0.20	-0.08
All interstate CC rays (50-60)	0.03	-0.03**	-0.05	0.22	-0.02
Federally funded CC rays (50-60)	0.02	-0.02	-0.03	0.26	-0.00
All rays in MSA (50-60)	0.06	-0.07***	-0.06	-0.08	-0.20
2-digit ray in MSA (50-60)	0.06	-0.07***	-0.05	-0.06	-0.23
Federally funded rays in MSA (50-60)	0.03	-0.03*	-0.06	0.11	-0.19
Rays in plan running through MSA	2.10	-0.24	-0.46*	.	.
Rays in plan running through CC	1.90	-0.01	-0.32	.	.

Years in parenthesis refer to first year in which the data are present and if a second number is present it represents the year in which the trend coefficient has been taken. Average: average value of the variable in the first year in which the variable is present. Levels, all U.S.: coefficient obtained regressing the variable in consideration on the good soil dummy. Levels, inside Division: coefficient obtained regressing the variable in consideration on the good soil dummy, controlling for Census Division fixed effects. Trend, all U.S.: coefficient obtained regressing the variable in consideration on the good soil dummy interacted by the year in which the trend coefficient has been taken, controlling for the interaction by the good soil dummy and the all the other years and omitting the interaction between by the good soil dummy and the first year in which the variable is present. Trend, inside Division: coefficient obtained regressing the variable in consideration on the good soil dummy interacted by the year in which the trend coefficient has been taken, controlling for Census Division fixed effects interacted by year fixed effects, the interaction by the good soil dummy and the all the other years and omitting the interaction between by the good soil dummy and the first year in which the variable is present. Robust standard errors are always used. CC: city center. MSA: metropolitan statistical area. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Baseline results: The effect of crime on suburbanization

VARIABLES	(1) Share Pop CC	(2) Violent crime	(3) Share Pop CC	(4) Share Pop CC	(5) Share Pop CC	(6) ln(Pop CC)
Violent crime	-0.0165***		-0.0717***		-0.0843***	-0.257***
	(0.00141)		(0.0121)		(0.0167)	(0.0718)
Good soil x Lead		-0.914***		0.0686***		
		(0.0649)		(0.00885)		
Observations	9,481	9,484	9,481	9,716	9,481	9,484
R-squared	0.960	0.758		0.956	0.934	
MSA FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
C. div x Year	NO	NO	NO	NO	YES	YES
Year	60-91	60-91	60-91	60-91	60-91	60-91
Estimation	OLS	OLS	IV	OLS	IV	IV
F	.	.	264.55	.	77.32	76.96

Share Pop CC: Proportion of MSA population living in city center. Violent crime: Violent crime per capita in the city center standardized. Good soil x Lead: dummy taking 1 if pH in the city center is between 6.8 and 7.7 multiplied by tonnes of lead consumed in U.S. as gasoline additive 19 years before, normalized by the maximum level of tetraethyl lead consumption. C. div: Census Division fixed effects. F: Cragg-Donald Wald F-statistics on the excluded instruments. Standard errors clustered at Census division times year level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: The effect of crime on population displacement

VARIABLES	(1) Log Pop MSA	(2) Log Pop CC	(3) Log Pop NCC	(4) Blacks MSA	(5) Blacks CC
Violent crime	0.0661	-0.257***	0.144***	0.137	4.730***
	(0.0428)	(0.0718)	(0.0534)	(0.809)	(0.795)
Observations	9,481	9,484	9,481	921	916
MSA FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
C. div x Year	YES	YES	YES	NO	NO
C. reg x Year	NO	NO	NO	YES	YES
Year	60-91	60-91	60-91	CY 60-90	CY 60-90
Estimation	IV	IV	IV	IV	IV
F	77.06	76.96	77.06	13.29	12.78

Violent crime: Violent crime per capita in the city center standardized. Pop MSA: Population in the MSA. Pop CC: Population in the city center. Pop NCC: Population in the suburbs. C. div: Census Division fixed effects. C. reg: Census Region fixed effects. F: Cragg-Donald Wald F-statistics on the excluded instruments. Standard errors clustered at Census division times year level (columns 1-3) and Census region times year level (columns 4-5) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6: The effect of crime on proportion of MSA employment in central city county

VARIABLES	(1) All	(2) Manuf	(3) Wholesale	(4) Retail	(5) Finance	(6) Other serv
Violent crime	-0.0108	-0.0462***	-0.0244*	-0.0386***	0.00283	-0.0285***
	(0.0113)	(0.0163)	(0.0143)	(0.0148)	(0.0178)	(0.0103)
Observations	5,334	5,334	5,334	5,334	5,334	5,334
MSA FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
C. div x Year	YES	YES	YES	YES	YES	YES
Year	74-91	74-91	74-91	74-91	74-91	74-91
Estimation	IV	IV	IV	IV	IV	IV
F	19.37	19.37	19.37	19.37	19.37	19.37

Violent crime: Violent crime per capita in the city center standardized. Dependent variable is proportion of MSA employment of SIC industry under consideration in county of the city center. All: all SIC employment. Manuf: Manufacturing; Wholesale: Wholesale Trade; Retail: Retail Trade; Finance: Finance, Insurance, And Real Estate; Other serv: Services). MSA with a unique county have missing values of employment proportion. C. div: Census Division fixed effects. F: Cragg-Donald Wald F-statistics on the excluded instruments. Standard errors clustered at Census division times year level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Timing of residential and employment decentralization

VARIABLES	(1) Share Manuf CC	(2) Share Manuf CC	(3) Share Pop CC	(4) Share Pop CC
Violent crime	-0.0241***	-0.0181	-0.0276***	-0.0478***
	(0.00863)	(0.0187)	(0.00631)	(0.0127)
Share Pop CC (10 years lag)		0.0656***		
		(0.0241)		
Share Manuf CC (10 years lag)				0.0256
				(0.0185)
Observations	5,352	5,332	5,367	2,344
MSA FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Good Soil x Lead 29 years lag	NO	YES	NO	YES
Year	74-91	74-91	74-91	74-91
Estimation	IV	IV	IV	IV
F	74.24	15.32	76.87	19.57

Violent crime: Violent crime per capita in the city center standardized. Share Pop CC: Proportion of MSA population living in city center.Share Manuf CC: proportion of MSA employment of manufacturing industry under consideration in county of the city center. MSA with a unique county have missing values of employment proportion. F: Cragg-Donald Wald F-statistics on the excluded instruments. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Profile of recent suburbanized population in 1980

Variable	Recent suburbanized	Difference wrt people staying in CC
Mean age	31.03	-2.86***
Mean number children	0.67	0.10***
Prop. married	0.62	0.09***
Prop. white	0.85	0.16***
Prop. black	0.11	-0.15***
Prop. high school or higher	0.63	0.14***
Prop. high school or higher of whites	0.64	0.11***
Prop. employed	0.55	0.11***
Prop. unemployed	0.03	-0.00***
Prop. people not working in CC	0.24	0.21***
Mean occupational score	18.69	4.18***

Prop.: Proportion. Recent suburbanized refers to people who in 1980 lives in a not central city in the metropolitan area and in the previous five years they were living in the central city of the same metropolitan area. Difference with respect to people staying in CC has been obtained regressing the variable under interested on a variable indicating whether the person is a recent suburbanized or he continues to live in the central city for the sample of people living in metropolitan areas. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Heterogeneous effects of correlation between crime rates and proportion blacks living in city center

	(1)
VARIABLES	Violent crime
Prop. Blacks CC	0.000379***
	(8.11e-05)
Prop. Blacks CC x Rays CC	-3.10e-05**
	(1.10e-05)
Prop. Blacks CC x Rays CC x Lead	4.90e-05***
	(9.02e-06)
Observations	916
R-squared	0.798
MSA FE	YES
Year FE	YES
C. reg X Year FE	YES
Year	CY 60-90
Estimation	OLS

Violent crime: Violent crime per capita in the city center. Prop. Blacks CC: proportion of black population in city center in the MSA. Rays CC: number of highway rays passing through the city center, source: Baum-Snow (2007). Lead: tonnes of lead consumed in U.S. as gasoline additive 19 years before, normalized by the maximum level of tetraethyl lead consumption. C. reg: Census region fixed effects. CY: census years. Standard errors clustered at Census region times year level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 10: Effect of interaction good soil and different lag of past leads on education outcomes and crime

VARIABLES	(1) Share high school	(2) Share high school
Good soil x Lead (9 years before)	-0.546	-0.546
	(0.708)	(1.860)
Good soil x Lead (19 years before)	0.699	0.699
	(1.895)	(4.270)
Good soil x Lead (29 years before)	-9.485***	-9.485
	(2.132)	(5.846)
Observations	1,174	1,174
R-squared	0.979	0.979
MSA FE	YES	YES
Year FE	YES	YES
C. reg x Year	YES	YES
Year	CY 60-90	CY 60-90
Estimation	OLS	OLS
s.e. cluster	MSA	C. reg x Year

Violent crime: Violent crime per capita in the city center. Share high school: share of population with high school diploma in the MSA. Good soil x Lead (X years before): dummy taking 1 if pH in the city center is between 6.8 and 7.7 multiplied by tonnes of lead consumed in U.S. as gasoline additive X years before, normalized by the maximum level of tetraethyl lead consumption. CY: Census years. C. reg: Census region fixed effects. CY: census years. s.e. cluster: cluster level of the standard errors. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 11: Results controlling for different lagged effect of education level

VARIABLES	(1) Share pop CC	(2) Share pop CC	(3) Share pop CC
Violent crime	-0.0591***	-0.0781***	
	(0.00678)	(0.0289)	
Good soil x Lead (29 years before)		-0.0166	
		(0.0208)	
Share high school MSA			-0.0105**
			(0.00426)
Observations	9,481	9,481	939
MSA FE	YES	YES	YES
Year FE	YES	YES	YES
C. reg x Year FE	YES	YES	YES
Year	60-91	60-91	CY 60-90
Estimation	IV	IV	IV
F	156.25	5.52	50.54

For the notes see Table 12. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 12: Results controlling for instrumented education level

VARIABLES	(1) Sh. pop CC	(2) V. crime	(3) Sh. pop CC	(4) H.s. MSA	(5) V. crime	(6) Sh. pop CC	(7) H.s. MSA	(8) V. crime
Violent crime	-0.0686***		-0.0611***			-0.0827***		
	(0.0214)		(0.0120)			(0.0232)		
Share high school MSA			-0.00183			0.000998		
			(0.00198)			(0.00245)		
Good soil x Lead		-0.957***		-7.135***	-1.024***		-7.008***	-1.018***
		(0.192)		(1.906)	(0.211)		(1.964)	(0.213)
H.s. U.S. x age school entry				0.0131***	-0.00259**			
				(0.00164)	(0.000899)			
H.s. U.S. x age school leave							0.00713***	-0.000962*
							(0.00116)	(0.000544)
Observations	1,184	1,185	917	935	1,181	917	935	1,181
MSA FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
C. reg x Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Year	CY 60-90	CY 60-90	CY 60-90	CY 60-90	CY 60-90	CY 60-90	CY 60-90	CY 60-90
Estimation	IV	OLS	IV	OLS	OLS	IV	OLS	OLS
F	17.51	.	12.24	.	.	7.97	.	.

Sh pop CC: Proportion of MSA population living in city center. V. crime: Violent crime per capita in the city center standardized. Good soil x Lead: dummy taking 1 if pH in the city center is between 6.8 and 7.7 multiplied by tonnes of lead consumed in U.S. as gasoline additive 19 years before, normalized by the maximum level of tetraethyl lead consumption. Share high school MSA: Percentage of people with high school diploma in the MSA. H.s. U.S.: Percentage of people with high school diploma in the U.S. Age school entry: state age school of entry in 1910. Age school leave: State minimum age school leave in 1910. C. reg: Census Region fixed effects. CY: Census Year. F: F-statistics on the excluded instruments. Stock-Yogo weak ID test critical value at 10% maximal IV size: 7.03. Standard errors clustered at Census region times year level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 13: Effect of violent crime on MSA population, house prices, and income

VARIABLES	(1) Pop. MSA	(2) H. rent MSA	(3) H. price MSA	(4) Inc MSA	(5) Inc MSA
Violent crime	95,144	57.94***	21,690**	3,422**	836.2*
	(68,244)	(16.58)	(8,838)	(1,252)	(449.5)
Observations	921	921	920	921	917
MSA FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
C. reg x Year	YES	YES	YES	YES	YES
Year	CY 60-90	CY 60-90	CY 60-90	CY 60-90	CY 60-90
Controls	.	.	.	.	Education
Estimation	IV	IV	IV	IV	IV
F	13.29	13.29	13.41	13.29	12.24

Violent crime: Violent crime per capita in the city center standardized. Pop MSA: Population in the MSA. H. rent MSA: Median gross rent per housing unit in the MSA. H. price MSA: Median single family house value. Inc MSA: Median family income in the MSA. Control for education: control for the percentage of people with high school diploma in the MSA instrumented by the state age school of entry in 1910 multiplied by the overall percentage of people with high school diploma in the U.S. CY: Census years. C. reg: Census Region fixed effects. F: F-statistics on the excluded instruments. Standard errors clustered at Census region times year level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 14: Effect of violent crime on MSA inequalities and transportation

VARIABLES	(1) Gini MSA	(2) Inc CC/MSA	(3) Inc CC	(4) Highway CC	(5) Pub transp MSA
Violent crime	0.00613**	-0.0945***	-358.3	0.255*	-2.212**
	(0.00243)	(0.0210)	(1,394)	(0.135)	(0.946)
Observations	921	921	921	921	921
MSA FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
C. reg x Year	YES	YES	YES	YES	YES
Year	CY 60-90	CY 60-90	CY 60-90	CY 60-90	CY 60-90
Estimation	IV	IV	IV	IV	IV
F	13.29	13.29	13.29	13.29	13.29

Violent crime: Violent crime per capita in the city center standardized. Inc CC/MSA: ratio between the median family income in the city center (CC) and the median family income in the MSA. Inc CC: median family income in the city center. Gini MSA: Simulated Gini coefficient from Baum-Snow (2007). Pub trans MSA: percentage of people using public transport to get to work. Highway CC: highway rays built passing through city center. CY: Census years. C. reg: Census Region fixed effects. F: F-statistics on the excluded instruments. Standard errors clustered at Census region times year level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 15: Elasticities of house rents, income and population with respect to violent crime

VARIABLES	(1) Log h. rent MSA	(2) Log income MSA	(3) Log Pop. MSA
Log violent crime	0.193**	0.157*	0.145
	(0.0836)	(0.0880)	(0.141)
Observations	918	918	918
R-squared	0.830	0.838	0.990
MSA FE	YES	YES	YES
Year FE	YES	YES	YES
C. reg x Year	YES	YES	YES
Year	CY 60-90	CY 60-90	CY 60-90
Estimation	IV	IV	IV
Cluster s.e.	MSA	MSA	MSA
F	11.69	11.69	11.69

H. rent MSA: Median gross rent per housing unit in the MSA. Income MSA: Median family income in the MSA. Pop. MSA: Total population in the MSA. Violent crime: Violent crime per capita in the city center standardized. C. reg: Census Region fixed effects. CY: Census Year. F: F-statistics on the excluded instruments. Standard errors clustered at Census region times year level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 16: Elasticities of amenities and productivity with respect to violent crime

	Control for mediation			
Coefficient	None	Highways	Resid suburbanization	Empl centralization
λ^{θ}	-0.0991**	-0.0972**	-0.1068**	-0.164
	(0.04551)	(0.04371)	(0.04552)	(0.6082)
λ^{A}	0.1243**	0.1219**	0.1260**	0.1887
	(0.05242)	(0.04991)	(0.05111)	(0.62977)

λ_{θ}: elasticity of city amenities with respect to violent crime. λ_{A}: elasticity of city productivity with respect to violent crime. Control for mediation, highways: the estimated regression also includes the number of highways passing through the MSA. Control for mediation, resid suburbanization: the estimated regression also includes the proportion of population in MSA living in city center. Control for mediation, empl centralization: the estimated regression also includes the proportion of employment in MSA located in the central county. Standard errors have been bootstrapped: this strategy consists in first bootstrapping a panel sample from our distribution of observations; subsequently, we have estimated the elasticity of house prices, income and population to violent crime and then compute the corresponding elasticity of amenities and productivity to violent crime; we have replicated this procedure several times and obtained a distribution of these parameters and relevant standard errors. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Flight from Urban Blight:
Lead Poisoning, Crime and Suburbanization
Online Appendix

A Summary statistics

Summary statistics of our database from 1960 to 1991 are reported in Table H1.

[INSERT TABLE H1 HERE]

B Additional evidence on the identification assumption

B.0.1 Relationship between lead poisoning and crime

Table H2 describes how much lead alone can explain violent crime rates in the U.S. Columns (1), (2) and (3) report the R-squared of a regression in which violent crime has been regressed on MSA fixed effects alone, year fixed effects alone and a combination of the two. As it is possible to see fixed effects can explain up to 75 % of the variation in violent crime. Regressing violent crime on lead we obtain a R-squared of 0.23, column (5), while additionally controlling for MSA fixed effects the R-squared becomes 0.47. That is, the time variation of lead alone can explain at most 23 % of the variation in violent crime.

[INSERT TABLE H2 HERE]

We can provide additional evidence that the effect of lead on violent crime should be weaker at levels of pH close to 7. Table H3 regresses violent crimes on the interaction of the national level of tetraethyl lead 19 years before and several dummies at different 0.5 bounds of pH. A negative and significant interaction effect of lead and pH is found only close to soil neutrality, where the pH is between 7 and 7.5.

[INSERT TABLE H3 HERE]

B.0.2 Standard errors of the generated instrument

In this Section we show that our generated instrument is not an extreme realization of the distribution of possible instruments. To do so we have obtained consistent standard errors of the minimum and maximum pH values for the soil quality index, and the F-statistics of the excluded instrument. In order to do this we have implemented a bootstrapping procedure. This strategy consists in first bootstrapping a panel sample from our distribution of observations. Subsequently, we have implemented the procedure for the creation of the instrument: we have run the first stage with any possible interval, and select the one that gives the highest F. We have replicated this procedure several times and obtained a distribution of these parameters and relevant standard errors. The estimated standard errors are 95.10 for F-statistics, 2.623 for the minimum value of the pH interval, and 0.784 for the maximum value of the pH interval. This reassures us that our selected instrument is not an outlier in the distribution of possible instruments.

B.0.3 Placebo distribution of the first stage

As further evidence of the relevance of the instrument we perform a placebo exercise. We first randomly assign the good soil status to cities. Each city has a 14.7% of being assigned the good soil status (the same probability as in our sample of cities). We then perform the first stage regression (Equation 2) with this fake good soil status assignment. We repeat this procedure 10000 times. Figure H1 reports distribution of the parameter χ of equation 2, the effect of the instrument on violent crime, estimated with the fake good soil assignment. The red line is the "real" χ parameter estimate of our first stage.

[INSERT FIGURE H1 HERE]

Results show that the differential effect that lead has on violent crime between city centers with good and bad soil cannot be replicated by any of the 10000 fake assignment. This is further evidence that there is something special about cities with good soil that allow them to have lower violent crime rates. Our claim is that this is due to the effect that the pH has on the bioavailability of lead.

B.1 Evidence on the exclusion restriction using the control function approach

We provide additional evidence that the exclusion restriction is met thanks to the specific timing choice we select for our instrument using the control function approach proposed by Wooldridge (2015). One of the main advantages of the control function is that it provides a test of endogeneity of the OLS estimation of the effect of crime on suburbanization. In the next paragraphs we show that estimating the effect of violent crime on the population of the suburbs might not suffer bias and then we provide evidence that the effect of resuspended lead is only passing through its effect on crime.

The control function approach requires estimating the first stage of our instrumental variable strategy, Equation 2, and then predicting the residuals of the first stage, $\hat{\epsilon}_{m,t}$. Once the residuals of the first stage are obtained it is possible to obtain identical results to the second stage of the instrumental variable strategy estimating the following model, Equation B1, in which the suburbanization proxy is regressed on controls, violent crime and the residuals of the first stage.

$$sub_{m,t} = \tau_m + \tau_t + \beta VC^{cc}_{m,t} + \chi\hat{\epsilon}_{m,t} + \varepsilon_{m,t} \tag{B1}$$

Controlling for the residuals of the first stage it is possible to correct the bias of the OLS estimation of Equation 1, since we control for the correlation between violent crime and the error, $\varepsilon_{m,t}$. Moreover, Wooldridge (2015) proves that the estimation of

Equation B1 produces a heteroskedasticity- robust Hausman test of the null hypothesis $H_0 : \chi = 0$, which means a test of whether violent crime is endogeneous.

Results of this test are reported in Table H4, columns (1) to (4). Test of endogeneity reports that violent crime is endogeneous when we estimate its effect on the population of the city center, Column (1). This is not confirmed when we study its effect on the population of the suburbs, Column (2). Therefore, the OLS estimation of the effect of violent crime on the population of city center or the percentage of people living in city centers is biased, and it is required to use our instrumental variable strategy. Similarly, the effect of education on population of city center is biased, while the estimation of its effect on the population of suburbs it is not (see Columns (3) and (4)).

[INSERT TABLE H4 HERE]

We can provide evidence in favour of the exogeneity of our instrument using the previous regressions. In fact, both violent crime and education are exogenous in the estimation of their effect on population of the suburbs. We can test whether the effect of lead poisoning affect population of the suburbs only through its effect on crime, since violent crime is not correlated with the anything else that explains the population of the suburbs. In Column (5) we show that our lead poisoning instrument has an effect on population of the suburbs. In column (6) we also control for violent crime. The coefficient of our instrument turns insignificant in this new specification providing evidence that its effect passes entirely through its influence on violent crime. Moreover, in column (7) we also control for the education proxy and we show that our results do not change.

B.1.1 Agricultural productivity

Soil quality can potentially affect suburbanization by influencing the relative proportion of land devoted to agriculture in one city. We have already showed in Section 4.4 that places with good and bad soil are similar in terms of pre-trends in agricultural employment. Further, we demonstrate in this section that the particular pH function we use to describe lead adsorption in soil does not affect potential yield of other crops and then the proxy of pH we use is not likely to affect agricultural productivity. In Table H5 we regress a measure of potential yield of several crops on our soil quality proxy[1]. As it is possible to see there is no significant difference in potential yields of crop between places with good and bad soil for lead soil adsorption.

[INSERT TABLE H5 HERE]

This pattern is consistent with the results presented in Figure H2. In fact, the marginal plot of the non-parametric effect of pH on potential yield of crops deos not mimic the non-linear effect on violent crime. The best range of pH for potential yields of the major 5 agricultural products in U.S. does not coincide with the soil quality proxy we use for predicting crime.

[INSERT FIGURE H2 HERE]

If pH of all the soil is anyhow correlated with any geological feature that might matter for suburbanization our estimations would suffer a bias. The inclusion of MSA

[1] Data for potential yield comes from IIASA/FAO (IIASA/FAO). We collect data and compute the potential yield at central city level for the major 5 agricultural products in U.S.: Corn (Maize), Soybeans, Wheat, Alfalfa, Cotton.

fixed effects in all our specifications take care of this particular concern. However, geological properties, in particular bedrock distance, can predict density in cities by influencing skyscrapers construction, as it has been shown by Rosenthal and Strange (2008), Combes et al. (2010), and Curci (2015). We show that this possible bias concern is further minimized since our instrument relies only on variation of soil quality at the top of the surface. In fact, lead expelled by cars accumulate into the top inches of soil (Stehouwer and Macneal, 1999). Therefore, the relevant soil quality index for lead poisoning is the pH of the top of the soil and not the overall pH of soil at any possible layer, which might or not be correlated with any other geological property. Table H6 checks this assumption by conducting the first stage regression using both the soil quality at the top of the soil surface and the mean soil quality index for all the soil in the city center. As it is predicted the only relevant soil quality for crime is the one on the top of the soil surface. This result, in addition with the use of MSA fixed effects in all our specification, further reassures us that the effect of pH on crime is coming from its interaction with lead.

[INSERT TABLE H6 HERE]

B.1.2 Crime spillovers

Understanding the impact of the increase in violent crime in the city center on the flight of population from the center to the suburbs helps us understanding the importance of crime as urban amenity that can influence location of people inside a city. To be sure we are really capturing the effect of a decrease in amenities in the city center we are implicitly imposing that violent crime in the suburbs did not move differently between cities with good and bad soil, for example because of displacement of criminal activity. In the period under analysis violent crime increased in both city centers and suburbs, however this increase was stronger and more significant in city centers[2]. We want to show that the increase in violent crime in the suburbs is not related anyhow with our instrument. That is, places with good and bad soil in the suburbs did not have different trend in criminal activity. Otherwise, our result can be biased because violent crime in the city center might not be the only source of change in population in city center.

Table H7 performs the first stage regression on the violent crime in the suburbs. This table shows that violent crime in the suburbs did not move differently comparing places with good and bad soils in the city center but also comparing places with good and bad soils in the suburbs. The first column of Table H7 shows that the pH level in the city center correlates significantly with the pH in the suburbs. Therefore, places with good (bad) soil in the city center tend to have good (bad) soil in the suburbs. Regressing violent crime in the suburbs on the interaction between national levels of lead and the good soil proxy for the suburbs we obtain a positive coefficient of the instrument. However, as it is shown in Column (3) and (5), this results is caused by the fact that we do not take into consideration the correlation between the pH inside the MSA. In fact, by clustering standard errors at MSA and controlling for the pH in the center we do not find that violent crime in the suburbs diverges between places

[2]Violent crimes in city centers increased from 22.5 crimes per 10,000 in 1960 to 163.6 in 1991. Violent crimes in suburbs increased from 6.4 crimes per 10,000 in 1960 to 47.4 in 1991.

with good and bad soil in the suburbs when the lead shock happen. Similarly, violent crime in the suburbs did not raise because of lead ingestion in the city center (see column (5)). That is, we find no evidence of crime spillover from the center to the suburbs.

[INSERT TABLE H7 HERE]

In Table H8 we show that crime in the city centers depends only on the interaction between lead poisoning and the soil quality of the center and not in the suburbs. In fact, the coefficient of the interaction between interaction between lead poisoning and the soil quality of the center is always negative and significant across specifications, while the coefficient of the interaction between interaction between lead poisoning and the soil quality of the surburbs is not significant as soon as we control for correlation in the errors at MSA level. Column (5) of Table H8 regresses the violent crime in suburbs and city center on a dummy for being in the city center. The positive and statistically significant coefficient shows that city center and suburbs are dramatically different in terms of crimes. City centers tend to have higher crimes. Using this result and the previous robustness checks we can conclude that city centers are the relevant unity of analysis for our exercise and that suburbanization of U.S. cities has been caused by increase in the crime in the centers and not in the rest of the city.

[INSERT TABLE H8 HERE]

We investigate why violent crime did not react differently between places with good and bad soil in the suburb. The first possibility is that places in the suburbs did not receive enough lead poisoning. The second possibility is that population density is a prerequisite for violent crimes. We cannot test directly the first possibility because we do not have data on the distribution of lead emissions inside cities. In Table H9 we provide evidence that the effect of lead poisoning in city center was higher in places with higher density at the beginning of the treatment (1960). Moreover, the negative difference in violent crimes in places with good and bad soil poisoned is smaller when population density is higher. This result is also confirmed by the marginal effect plot presented in Figure H3.

[INSERT TABLE H9 HERE]

[INSERT FIGURE H3 HERE]

B.1.3 Other crimes

Medical literature suggests that lead poisoning can increase the propensity of committing crime since it affects the part of the brain that control impulses and aggression. Therefore, lead poisoning is likely to increase only violent crimes. We conduct a falsification exercise to show that our instrument only affects crimes which are violent and not other crimes that might affect suburbanization. We disaggregate crimes between crimes which are violent according to the FBI, that is murder and non-negligent manslaughter, total robberies, forcible rape, and aggravated assaults, and crimes that are not violent: burglary and larceny. Controlling for the possible spatial correlation in crimes, we find that the only crimes that are affected by our instrument are the violent one and not the others (see Table H10)[3].

[3] Inside the violent crimes all the kind of crimes seems to be affected by resuspended lead but total murder rates

[INSERT TABLE H10 HERE]

C Reverse causality bias in the OLS estimation of suburbanization

In this section we discuss why reverse causality in the effect of crime on suburbanization can explain the difference between the OLS and IV estimates. The relationship running from violent crime (VC) to suburbanization (Sub) can be written as in Equation C1, while the relationship from suburbanization to violent crime can be written as in Equation C2.

$$Sub = \xi_1 VC + \epsilon_1 \tag{C1}$$
$$VC = \xi_2 Sub + \epsilon_2 \tag{C2}$$

I am interested in the relationship running from violent crime to suburbanization. Therefore, the model that is estimated by OLS can be written as in Equation C3, by combining Equations C1 and C2. OLS estimates will be biased because of the simultaneity. The OLS bias is equal to $\frac{\xi_1 - 1}{1-\xi_2} - \xi_1$. OLS estimate will be downward bias with respect to IV if the condition in Equation C4 is met.

$$\text{OLS: } Sub = \frac{\xi_1 - 1}{1-\xi_2} VC + \frac{1}{1-\xi_2}(\epsilon_1 + \epsilon_2) \tag{C3}$$
$$\text{Upward bias OLS: } \frac{\xi_1 - 1}{1-\xi_2} > \xi_1 \rightarrow \xi_1 > \frac{1}{\xi_2} \tag{C4}$$

From the IV estimate we know that $\xi_1 = -0.07$. Let consider the case in which there exists a positive relationship running from suburbanization to violent crime. For example, more suburbanized cities might be more racially segregated and this could potentially increase crime rates. If this is the case this reverse causality cannot explain the upward bias in the OLS estimate, since ξ_2 would be positive and then $\xi_1 > \frac{1}{\xi_2}$ cannot be met. Therefore, the only potential reverse causality that is consistent with our results is one in which more suburbanized cities lead to lower levels of violent crimes.

D Robustness of the results

D.1 Potential confounders: highway

We test the robustness of our suburbanization mechanism with respect to other potential confounders. First, we provide evidence that suburbanization happened more in places with higher crimes irrespective of the highway construction in that city. In order to do this we use data from Baum-Snow (2007). In this work the author regresses the population living in the city center on a proxy of the number of highway rays passing from a city center. He obtains exogeneous variation in highway construction using as instrument the 1947 national interstate highway plan. The identification assumption required to unbiasedly estimate our causal coefficient is that the interaction between national lagged lead levels and good soil for lead adsorption does not correlate with actual highway construction. In Section 4.4 we show that places with good and bad soils have the same pre-treatment levels and trends in highway construction.

In Table H11 we formally test that our instrument is not correlated with either planned and actual highways in U.S. (see columns (1) and (2)). Moreover, places with good soil have not either planned and actual highways with respect to places with bad soil (see columns (3) and (4)).

[INSERT TABLE H11 HERE]

Table H12 jointly estimates the effect of violent crime on suburbanization controlling for the level of highway construction. Columns (1) and (2) report the second and first stage of the effect of highway construction on suburbanization alone. As it is possible to see cities with more highways are associated with higher levels of suburbanization. Using the same sample as Baum-Snow (2007) columns (3) and (4) report the second and first stage of the effect of violent crime alone. Column (5) estimates the effect of violent crime and highway construction jointly, using both our instrument and the one proposed by Baum-Snow (2007) (columns (6) and (7) show the corresponding first stage regressions). It is evident that our effect is robust to control for the highway construction and the estimated coefficient is of similar magnitude.

[INSERT TABLE H12 HERE]

D.2 Potential confounders: other variables

Table H13 provides evidence that our results are robust to the inclusion of several socioeconomic and housing characteristics. Column (2) of Table H13 imposes different trends according the pre-treatment share of suburbanizations. Columns (3) to (6) of Table H13 control for the proportion of blacks in one MSA, the median family income and the proportion of people over 25 and 65 in one city. = Columns (7) of Table H13 controls for median gross rent. If soil pH affects suburbanization through its affect on agricultural productivity this should affect land price and the city housing market. However, controlling for house prices our results are not changed.

[INSERT TABLE H13 HERE]

D.3 Using other instrument definitions

The instrument we use to predict violent crimes is the interaction between a proxy of soil quality for lead adsorption in the soil and the 19th lag of national lead consumed in U.S. In particular, we define as good soil a city center in which the pH of the soil is between 6.7 and 7.7. In this section we show that our results are robust to the specification of our instrument used in our estimations. We both change the definition of our soil quality index and the lags used of lead poisoning and we find consistent results.

We obtain IV estimates of the effect of violent crime on suburbanization using any possible definition of soil quality. We report all the possible results in Figure H4. Changing slightly the minimum and maximum pH used for defining a good soil does not alter the results obtained both in terms of direction and magnitude. Our IV estimate does not lie in any tail of the distribution of the possible effects of crime on suburbanization.

[INSERT FIGURE H4 HERE]

In our baseline specifications we use a linear and binary function of pH to explain violent crime. However, the effect of pH on violent crime can potentially be non-linear. Therefore, in Table H14 we use different functions of pH in the first stage. Column (1) reports our baseline binary specification. Column (2) uses as instrument several pH dummies interacted with past national lead poisoning [4]. Column (3) uses as instrument the interaction between pH and past national lead poisoning , while Columns (4) to (6) also uses the the interaction between pH to the power of 2, 3, 4 and past national lead poisoning, respectively. Results are robust according to the different specification of our instrument. However, the magnitude of our estimated coefficient decreases with more polynomial of pH included in the first stage.

[INSERT TABLE H14 HERE]

We also find robust results of the effect of crime on suburbanization changing the lead poisoning variable used as an instrument. Table H15 reports the results using separately the national levels of lead used as gasoline additive from 15 to 23 years before. As it is possible to see the effect of crime on suburbanization does not change in magnitude or sign moving around our preferred past level of lead poisoning. Moreover, the 19th lag of lead poisoning is the lag with the highest F-statistics between the possible lags giving further credit of its use.

[INSERT TABLE H15 HERE]

In the previous regressions we have used the 19th lag of lead poisoning as part of our instrument since the highest propensity of committing crimes has been found to be at 19 years old. The required identification assumption is that lead poisoning through soil is not affecting some other variable which has the same age-structure effect on suburbanization as crime. To strengthen this assumption we conduct an additional exercise in which we do not only use lead related to the maximum propensity of committing crime but all the age structure of crime rate. In fact, for every year we have constructed the total poisoning of the population relevant for violent crimes by weighting past national level of lead poisoning by the probability of commiting crimes later. That is, our new measure of lead poisoning is $\sum_{j=1}^{70} TL_{t-j} P\left[VC_{a=j}\right]$, where TL is the national level of tetraethyl lead used as gasoline additive t-j years ago and P[VC] is the probability of committing violent crime at a particular age (a) equal to j [5]. For example, in year 1991 we have computed the total lead poisoning as the lead poisoning 19 years before, in 1972, multiplied by the probability of one 19 year old person to commit a crime, plus lead poisoning 20 years before, in 1971, multiplied by the prob-

[4]The pH dummies used are: pH between 5 and 5.5, pH between 5.5 and 6, pH between 6 and 6.5, pH between 6.5 and 7, pH between 7 and 7.5, pH between 7.5 and 8, pH between 8 and 8.5

[5]Probability of committing violent crime has been measured using the age-specific arrest rates provided by United States Department of Justice (2003). We have normalized the probability of committing violent crime at any age by the maximum probability of committing violent crime.

ability of one 20 year old person to commit a crime, and so on for all possible ages from 1 to 70.

The time series of this new lead poisoning measure is reported in Figure H5. This new time series is smoother than using the level of lead poisoning 19 years before. The peak of the two time series coincide. Moreover, weighted lead poisoning time series decays at a slower pace than the lagged lead poisoning variable since there are still people in U.S. poisoned in their youth by gasoline additives, and this explains why the difference in crimes between places with good and bad soil is still not canceled.

[INSERT FIGURE H5 HERE]

We have run our regression of the effect of crime on suburbanization using this new definition of lead poisoning instead of the 19th lag of national tetraethyl lead. In Table H16 we find very similar results suggesting that is unlikely that lead poisoning through soil is affecting some other variable which has the same age-structure effect on suburbanization as crime.

[INSERT TABLE H16 HERE]

D.4 Different fixed effects and weighting

We demonstrate that our results are robust to narrowing the identification strategy to smaller samples. In Table H17 we report results using different geographical specific fixed effects. Our results are consistent also comparing observations uniquely inside a Census Region, Division or States.

[INSERT TABLE H17 HERE]

Our instrument includes geological properties of the soil that affect violent crime via lead poisoning. We have already demonstrated in Section 2 that the pH variable used is as good as randomly assigned between places. We provide further evidence by estimating the effect of crime on suburbanization comparing observations with similar geological characteristics. In fact, we impose common trends for observations with similar levels of slope, precipitation and distance from water or border and we find similar coefficients to our baseline result (see Table H18).

[INSERT TABLE H18 HERE]

We can also use the subsample of bigger cities to re-derive our results. We obtain this by running our estimations using the subsample of cities with more than 100,000 inhabitants but also weighting our regressions by the MSA population. From Table H19 we can conclude that the effect of crime on suburbanization is higher in bigger cities.

[INSERT TABLE H19 HERE]

D.5 Standard errors robustness

In this section we demonstrate that not only the coefficient but also the standard errors of the estimated results are consistent to several different specifications. In particular, we show that our standard errors are robust to considering different kind of geographical correlations and taking into account that our instrument is a predicted variable.

In Section 4.3 we show that the pH variable has some geographical correlation. Moreover, other factors explaining suburbanization can be correlated across space. In Table H20 we present results using different level of clustering in the standard errors in order to take into account theses possible geographical correlations in the error of the second stage. The significance of our result is maintained in all the different specifications.

[INSERT TABLE H20 HERE]

The soil quality index used as part of our instrument has been found with a machine learning procedure in which we have selected the interval of pH which guarantees the maximum F-statistics in the first stage. Our instrument is therefore a predicted variable and this can generate a bias in the standard errors used in the instrumental variable estimations. We have obtained consistent standard errors of the effect of violent crime on suburbanization using a bootstrapping procedure similar to the one described in Appendix B.0.2.

This strategy consists in first bootstrapping a panel sample from our distribution of observations. Subsequently, we have implemented the procedure for the creation of the instrument: we have run the first stage with any possible interval, and select the one that gives the highest F. Finally, we have used the best estimated interval to instrument violent crime and obtained its effect on suburbanization. We have replicated this procedure several times and obtained a distribution of these parameters and relevant standard errors. The estimated coefficient for the effect of violent crime on suburbanization controlling for MSA and year fixed effect is -0.0717 with a standard error of 0.0286. We can see that the significance of the effect of violent crime on suburbanization is consistent to this bootstrapping procedure.

E De-leading phase

Patterson (1965) provided the first scientific evidence that showed that the observed high levels of lead were man-made. This started in the US a series of regulatory changes that slowed down the increase of lead used in gasoline that ultimately reached a peak between 1970-1972. As shown in Figure 2, if we take into account the expected lag that lead has on violent crime, this is consistent with the peak of violent crime in the US that happened in 1991. What this figure shows is that after 1991 with the de-leading process violent crime decreased in the US.

What we then explore in this section is how this decrease in violent crime could be potentially different between city centers with good or bad soil. As discussed previously, cities with good and bad soil started in 1960 at similar of violent crime rate, and as the use of lead increased in the US this lead was ingested by humans at a higher rate in cities with bad soil. This in turn increased the level of violent crime disproportionately in these cities. Subsequently, when the de-leading phase started the lead ingested by humans dropped in the US, especially in the cities with bad soil. Because of this reason we have to expect that violent crime dropped especially in cities with bad soil.

We have previously found evidence of this process estimating the first stage year by year on all the sample, including the de-leading phase, that is Equation 3. Results are reported in Figure 8, in which we have shown that places with bad soil decreased faster violent crime than places with good soil in years in which lead poisoning decreased nationally (after 1991). This is in line with the reasoning we put forward before where cities with bad soil suffered a more substantial change in the lead ingested by humans in the de-leading phase leading to a more substantial drop in the violent crime. Between 1991 and 2014 the violent crime rate dropped by 18.1 violent crimes 10000 inhabitant in cities with good soil. In the same period cities with bad soil experienced a drop of 56.4 violent crimes per 10000 inhabitants.

Now that we know that in the de-leading phase violent crime reacted differently in places with good and bad soil we can explore what happened to suburbanization in the same period. There is one caveat that is important to have in mind when interpreting the results of this analysis. When we study the de-leading phase we are not performing a properly executed difference in difference analysis in which any difference in the suburbanization can be attributed to the different trends in violent crime. The reason is that, as we proved in the rest of the paper, even if cities with good and bad soil started with very similar levels and trends of violent crime and suburbanization in 1960 by 1991 they had very different ones. So when we look at the de-leading phase and we are not only comparing places that experienced different changes in violent crime due to different soils, but that also have different levels of crime and suburbanization. Potentially both levels of these variables are very important in determining the reaction of suburbanization decisions on crime. Still even if the results cannot be considered causal in this sense it is interesting to explore what happened to the suburbanization phenomenon when the de-leading phase started and cities with good and bad soil converged to more similar levels of violent crime.

In order to estimate the evolution of suburbanization in the de-leading process we estimate Equation 1 on the whole sample, including the de-leading period, in which we allow good soil to have a different effect by year on suburbanization. We run a regression of the effect of good soil interacted with year dummies on the suburbanization measure and we study the evolution of the parameters β_t.

$$sub_{m,t} = \tau_m + \tau_t + \beta_t 1(year = t) * good\ soil_{cc(m)} + \varepsilon_{m,t} \tag{1}$$

Results are shown in Figure H6. First, as expected, between 1960 and 1991 cities with good soil experienced a much slower process of suburbanization when comparing them with the cities with bad soil. After 1991, in the de-leading phase, this process did not revert back to the levels of 1960 and places with bad soil continued to have a more accelerated process of suburbanization. In 1991, cities with bad soil had 8.1% less people living in the city center as a percentage of total population. In 2014, this difference increase slightly to 9.6%.

[INSERT FIGURE H6 HERE]

We can confirm the result that cities with bad soil did not decrease suburbanization when violent crime decreased by estimating the effect of violent crime on suburbanization using our entire sample. The exogeneous variation we exploit always comes from using lead poisoning via soil as an instrument. Results are reported in Table H21. Using only the de-leading sample (1992-2014) we obtain a positive and significant effect of violent crime on suburbanization, with an estimated effect such that increasing by one standard deviation violent crime increased suburbanization by 3 percentage points. This is in the line with the results previously found because in that period violent crime decreased the most in places with bad soil but the difference in suburbanization maintained constant. That is, one standard deviation decrease in violent crime after 1991 caused a increase in suburbanization by only 3 percentage points, while the increase in one standard deviation of crimes between the 1960 and 1991 caused an increase in suburbanization by more than 8 percentage points. Estimating the effect of violent crime on suburbanization using the entire sample we obtain a negative and significant effect of a smaller magnitude than when we estimate it on the period 1960-1991 (-0.0564 versus -0.0869).

[INSERT TABLE H21 HERE]

The reasons why the cities with bad soil did not experience a reversal in suburbanization may be different. First of all, an asymmetry is possible, while increases in violent crime can generate suburbanization, a decrease in violent crime may not affect the suburbanization decisions of individuals. Second, in all the de-leading period cities with bad soil had higher levels of violent crime and it is possible that the level of violent crime is the one affecting suburbanization and not the changes. Finally, once a city is highly suburbanized even if violent crime in the city centers goes down people may decide to remain in the suburbs and do not return back to the city centers. Suburbs might generate amenities with time that do not make attractive to return to city centers. Independently on which of these stories drives the results it is possible to see that there is persistence in the suburbanization process. Even if cities with bad and good soil in 2014 have again very similar levels of violent crime, bad soil cities continue to have a much larger proportion of the population living in the suburbs.

F Heterogeneity: Time-varying effects

In this section we study the possible time varying effect of lead on crime and then of crime on suburbanization. First, looking at the effect of lead as the medical literature describes, we should not expect any time varying effect between lead ingestion and individual aggressivity. What may indeed change through time is how this then affects the overall violent crime rate of a city. For studying this we estimate the effect of good soil on violent crime allowing good soil to have a different effect every year, that is Equation 3.

The estimated parameters χ_t indicates the difference in violent crime growth between good and bad soil cities between the reference year (1960) and the year t. If the effects were to be time invariant the response of violent to lead consumption should be the same. We explore this by calculating the change in violent crime per change of lead consumption ($\frac{\Delta VC}{\Delta TL^W}$) for every possible 10 year interval using the following statistic, in which TL refers to the national level of lead poisoning ingested until that year weighted by the propensity of commiting crime at every age described in Appendix D.3[6]:

$$\left(\frac{\Delta VC}{\Delta TL}\right)_t = \frac{\chi_t - \chi_{t-10}}{TL_t^W - TL_{t-10}^W}$$

We plot the results in Figure H7 and we show that the effects of lead on violent crime are relatively time invariant. For almost all 10 year period an increase of lead

[6]We use this proxy instead of the 19th years lag in other to obtain smoother results. Interpretation of the results does not change using the other proxy

from zero to the peak US consumption would have increase violent crime in bad soil cities 0.84 standard deviations. That is, the 10 year effect of good soil on violent crime is similar independently of what time period we use to estimate it.

[INSERT FIGURE H7 HERE]

Finally, we study how the effect of crime on suburbanization may change trough time. For doing so we combine the estimates of the effect of good soil on violent crime through time, Equation 3, with the estimates of the effect of good soil on suburbanization through time, Equation 3. The parameter β_t in Equation 3 indicates the difference in suburbanization changes between good and bad soil cities between the reference year and the year t. Similarly to what we have done in the previous analysis we calculate the change in suburbanization per change in violent crime ($\frac{\Delta sub}{\Delta vc}$) for every possible 10 year interval using the following statistic:

$$\left(\frac{\Delta sub}{\Delta vc}\right)_t = \frac{\beta_t - \beta_{t-10}}{\chi_t - \chi_{t-10}}$$

This statistics can be interpreted as the Wald estimate of the effect of violent crime on suburbanization for a particular sub-period, which is identical to the IV estimate. We then plot these statistic over time in Figure 8(a). We focus on the period in which lead poisoning increased, that is from 1960 to 1991. As depicted by the red line in average for an increase of 1 standard deviation in the violent crime rate the population in the city center decreases by 7.4 percentage points. This effect declined over time. While in the 10 years between 1960 and 1970 an increase of one standard deviation in violent crime generated a decrease of the population living in the city center of 20 percentage points. The suburbanization decisions became less reactive to changes in crime and in the last ten years analyzed between 1981 and 1991 an increase of one standard deviation in violent crime decrease only of 4.8 percentage points the population of the city center. To show robustness of our results we also report the estimate using a 5 years interval in Figure 8(b). For the de-leading phase we confirm the results obtained in Appendix E. In fact, the decrease in violent crime did not make people return to city centers. Moreover, this effect is constant across the period 1992 to 2014.

[INSERT FIGURE H8 HERE]

Because of data limitation we cannot properly study the effect of violent crime on suburbanization at the very beginning of lead poisoning shock. In fact, yearly data about violent crime at within city level and suburbanization are only available from 1960, while lead poisoning started to increase in 1927 and then its first effect on crime might be present already in the early 1950s. We can try to have some insights about the speed of reaction of people to the increase in crime using data for suburbanization in 1950 by Baum-Snow (2007). In order to compute the previous statistics $\left(\frac{\Delta sub}{\Delta vc}\right)$ between 1950 and 1960 we will estimate the effect of good soil on suburbanization between 1950 and 1960. We further assume that the effect of good soil on violent crime is identical to its effect after 1960, which can be estimated in our first stage, and we will weight it by the increase in past lead between 1950 and 1960 that can be obtained interpolating the lead time series backwards.

From Figure H8 we can deduce that suburbanization might have reacted quickly to the increase in violent crime. However, looking at Figure 8(b), which provide the IV on a smaller time window, the effect of violent crime on suburbanization seems to be weaker between 1950 and 1960 than in the period between 1960 and 1965, in which it reaches its maximum. Therefore, it seems that the maximum effect of violent crime on suburbanization happened in the decade between 1960 and 1970, that is at least 10 years after that violent crime had increased consecutively. This might reflect the fact that the decision of suburbanization requires considerable fixed costs to convince people to move away, and that people decide to move only when the increase in violent crime has been manifested for some years.

G Mechanisms and channels

In this section we discuss the possible mechanisms and channels that can explain the effect of crime on suburbanization. In particular, we compute the heterogeneity of our causal effect with respect to many variables. To understand how the effect of crime on suburbanization change according a variable X we estimate the model reported in Equation G1.

$$sub_{m,t} = \tau_m + \tau_t + \beta_0 VC^{cc}_{m,t} + \beta_1 VC^{cc}_{m,t} \times X_m + \beta_2 VC^{cc}_m \times X^2_m + \beta_3 VC^{cc}_m \times X^3_m + \chi \hat{\epsilon}_{m,t} + \varepsilon_{m,t} \tag{G1}$$

To address the endogeneity of the variable violent crime the model we apply the control function approach proposed by Wooldridge (2015) and described in Section 6.1. That is, controlling for the estimated residuals of the first stage, $\hat{\epsilon}$, we control for the endogeneity of the violent crime rate and we do not to produce additional instruments for each interaction variable between violent crime and the variable X. We formally test the presence of heterogeneity of the causal effect of violent crime on suburbanization by testing the joint significance of the coefficient of the interaction variables, that is $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$. Figure H9 reports the marginal effect of violent crime effect on suburbanization with respect to all the city characteristics we consider.

[INSERT FIGURE H9 HERE]

Previous evidence suggests that African Americans commit higher rates of violent crimes (Sampson and Lauritsen, 1997). We have shown in Section 6.1 that the crime propensity of the black subpopulation increases in the years in which lead poisoning was higher (see Table 9). Moreover, we can see from Figure H10 that the effect of resuspended lead on violent crime is stronger in cities with more black population in the city center[7]. We test whether the flight of people from city centers because of increased crime is higher in cities with more blacks.

[INSERT FIGURE H10 HERE]

Panel (a) of Figure H9 shows that there is no heterogeneity in the effect of crime on suburbanization with respect to the percentage of blacks living in the city center at the beginning of the treatment period (1960). However, we find that the effect of crime on suburbanization is stronger in cities in which blacks are overrepresented in the city centers with respect to the suburbs (see panel b). That is, people decide to

[7] Interestingly, the heterogeneous effect of resuspended lead on violent crime is not linear, th effect of resuspended soil on violent crime increases until the point in which the proportion of blacks is approximately 30 % and then it decreases.

leave city centers if suburbs have lower levels of black population. This is in line with the previous evidence about the white flight process. This process can result because of two possible mechanisms that we cannot test. First, whites might perceive that blacks commit more crimes and they escape to suburbs if they believe that in those areas less blacks and then less crime are present. Second, this process can be the result of racism of the white subpopulation leaving the city center.

In Section 5.1 we show that violent crime reduced the population of city centers by 26 %. The movement of a big fraction of people to the suburbs required important fixed costs in the construction of the suburbs and their amenities and infrastructures. Therefore, we expect that the suburbanization process was stronger in cities that had already suburbs developed. Panel (c) of Figure H9 confirms this conjecture. In fact, the effect of crime on suburbanization is stronger in cities that were already more suburbanized in 1950. In Panel (d) we perform the same analysis for the first period of the suburbanization process, 1960-1975. We find that the heterogeneity of previous suburbanization is much stronger in the period in which cities start to lose population from city centers.

We have also considered heterogeneity with respect to several other previous city characteristics: income inequalities, median income, education and median age. We find that cities in which income inequalities were higher tend to have lower effect of crime on suburbanization (panel e of Figure H9). If within city income inequalities represent city segregation then this result can be interpreted as if crime moved more people away from city centers if cities were not previously segregated and there was a higher mix of incomes in the city center.

In panel (f) of Figure H9 we show that the effect of crime on suburbanization is stronger in cities that are richer, while we do not find heterogeneities with respect to the education level in the city (panel g). Moreover, we find that cities in which the median age is closer to the average age of the people suburbanizing (31 years) responded more strongly to the increase in crime.

We then ask if suburbanization was stronger in cities in which geographical constraints to the expansion of the city were present. Saiz (2010) demonstrates that housing supply is severely influenced by the geography of cities. He constructs an index for the undevelopable area of a city due to geographical constraints that takes into account of the presence of water and slopes in the city. For each city in our panel database we compute the same indexes. We draw a radius of 20 km around the centroid of the city center for every city and we compute the proportion of that radius which does not include internal water bodies and wetlands and which does not include terrain with slopes higher than 15 %[8].

Panels (i) and (j) of Figure H9 show the heterogeneity of the effect of crime on suburbanization with respect to the area of a city undevelopable because of water or slopes, respectively. Results suggest that the effect of crime is heterogeneous to the proportion of area with water in a city but not with respect to the area with steep-sloped terrain. In fact, cities which are surrounded by water, because of rivers, lakes

[8] We use several databases. For a map of the coast of the U.S. we use Water polygons OpenStreetMap-Data. For the map of inland waters we use USGS Small-scale Dataset, Global Map: 1:1,000,000-Scale Inland Water Areas of the United States 201406 Shapefile. For slope of the terrain we use the U.S.G.S. General Soil Map

or oceans, suffer a lower flight of people from city centers after the increase in crime.

Baum-Snow (2007) provides evidence that highways play a major role in creating suburbanization of U.S. cities. We have demonstrated in Appendix D.1 that the estimate of the violent crime effect on suburbanization is exogeneous to the level of highways in one city. Figure H10 show that the effect of lead poisoning on violent crime is stronger in cities with more highways. Furthermore, in Section 7 we show that violent crime leaded to the construction of new highways possibly to accommodate suburbanization. Therefore, we expect that the crime effect on suburbanization should be stronger in cities with more highways. Panel (k) of Figure H9 shows that this is exactly the case.

Finally, history of city might matter for the decision of suburbanization. One possibility is that older cities might have more historical amenities in the city center and this can contrast the decrease in amenities given by the increase of violent crimes in the center. Using data from NHGIS (2011) we have proxied the age of a city by historical densities in 1900[9]. We do not find any heterogeneity of the effect of crime with respect to historical densities (see panel l of Figure H9)

[9] Since counties in 1900 does not correspond to current counties. For each county we have associated the county that it would have belonged in 1900. For each MSA population density in 1900 have been computed for all the counties that have a link to counties in that MSA today. 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.

References

Baum-Snow, N. (2007). Did highways cause suburbanization? *The Quarterly Journal of Economics 122*(2), 775–805.

Combes, P.-P., G. Duranton, L. Gobillon, and S. Roux (2010). Estimating agglomeration economies with history, geology, and worker effects. In *Agglomeration economics*, pp. 15–66. University of Chicago Press.

Curci, F. (2015). The taller the better? agglomeration determinants and urban structure.

IIASA/FAO. Global agro-ecological zones (gaez v3.0), 2010.

NHGIS (2011). Minnesota population center's national historical geographic information system: Version 2.0. Database, University of Minnesota.

Patterson, C. C. (1965). Contaminated and natural lead environments of man. *Archives of Environmental Health: An International Journal 11*(3), 344–360.

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

Saiz, A. (2010). The geographic determinants of housing supply. *Quarterly Journal of Economics 125*(3).

Sampson, R. J. and J. L. Lauritsen (1997). Racial and ethnic disparities in crime and criminal justice in the united states. *Crime and justice 21*, 311–374.

Stehouwer, R. and K. Macneal (1999). Lead in residential soils: Sources, testing, and reducing exposure.

United States Department of Justice, F. B. I. (2003). Age-specific arrest rates and race-specific arrest rates for selected offenses, 1993-2001.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources 50*(2), 420–445.

H Figures and Tables

H.1 Figures

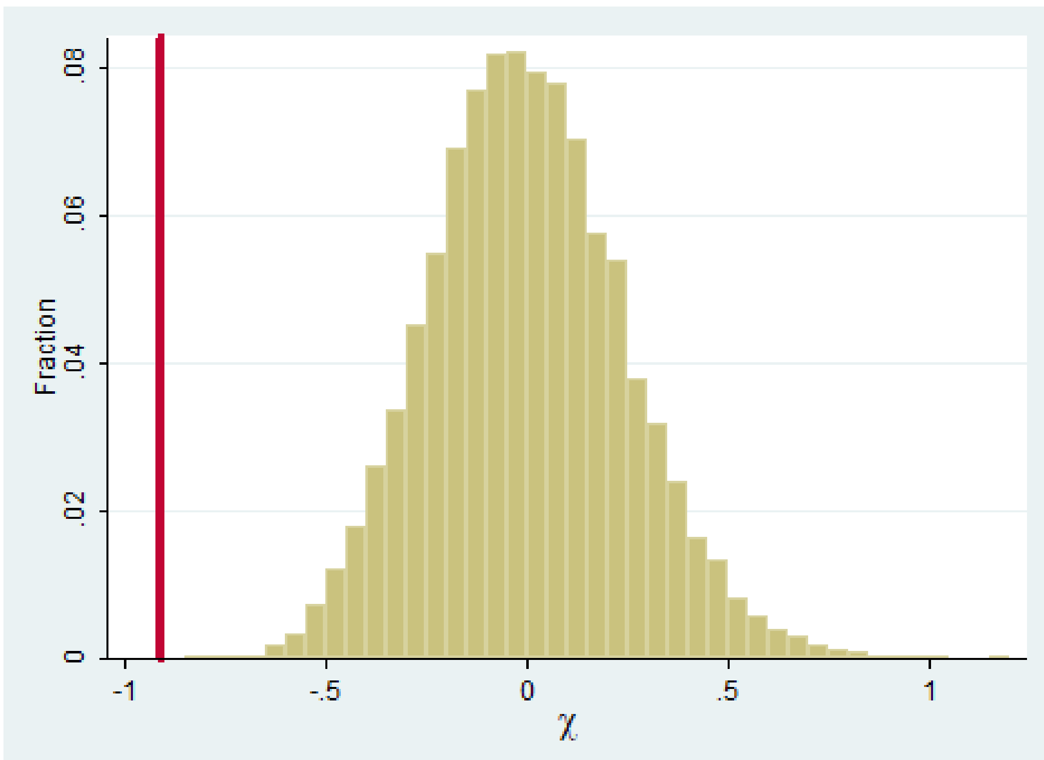


Figure H1: Distribution of the effect of the placebo and real instrument on violent crime

Note: This figure reports the distribution 10000 estimations of equation 2 were good soil status is randomly assigned. χ refers to the coefficient of the effect of each placebo instrument on violent crime rates in the city center. The red line indicates the first stage estimate of our instrument.

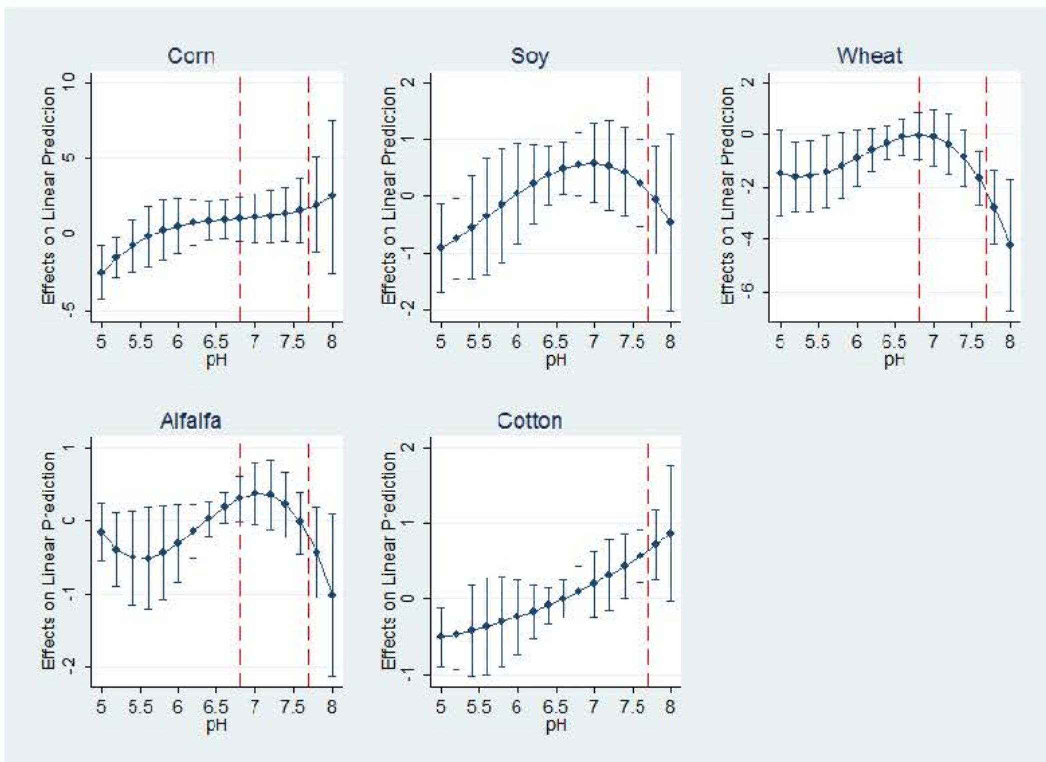


Figure H2: Average marginal effect of pH on potential yield of different crops

Marginal effects (with 95 % confidence interval) derived after regressing potential yield for specific crop on fourth polynomial of pH and Census division dummies. Robust standard errors have been used. Red lines indicates pH equal to 6.8 and 7.7

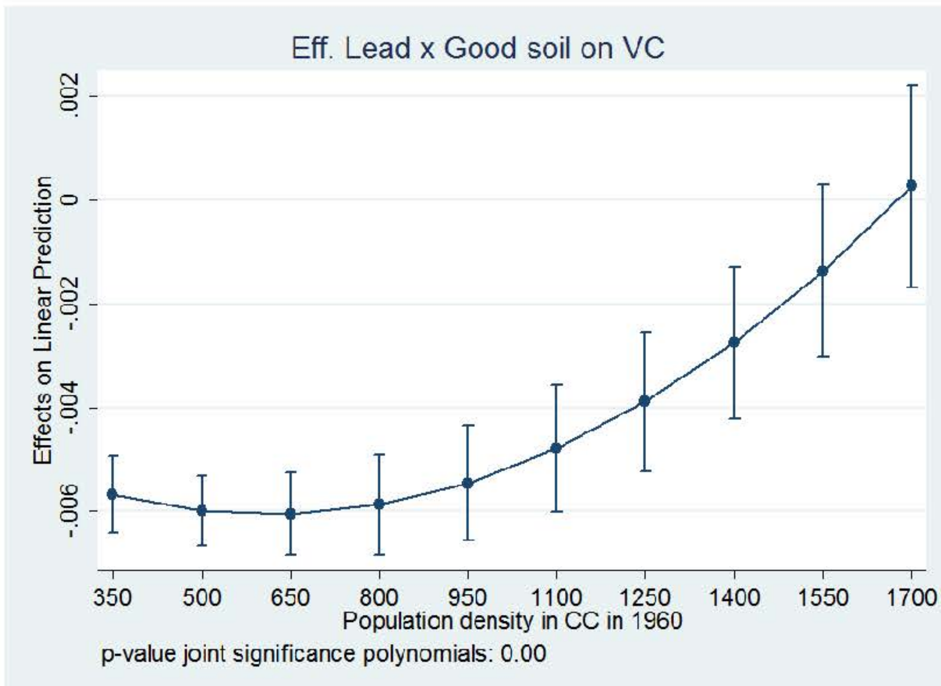


Figure H3: Heterogeneity of average marginal effect of the first stage with respect to population density values in 1960 in CC

VC: Violent crime per capita in the city center. Tetraethyl lead: Tonnes of lead consumed in U.S. as gasoline additive 19 years before, normalized by the maximum level of tetraethyl lead consumption. Good soil: dummy taking 1 if pH in the city center is between 6.8 and 7.7. $dens_{CC,60}$: population density in city center in 1960. Marginal effects derived after regressing violent crime per capita in CC on tetraethyl lead x good soil, tetraethyl lead x good soil x $dens_{CC,60}$, and tetraethyl lead x good soil x $dens^2_{CC,60}$. Robust standard errors have been used. Marginal effects reported for value of population density in 1960 between the 20th and 80th percentile. p-value for the test of joint significance of the coefficients of the following regressors: tetraethyl lead x good soil x $dens_{CC,60}$, and tetraethyl lead x good soil x $dens^2_{CC,60}$.

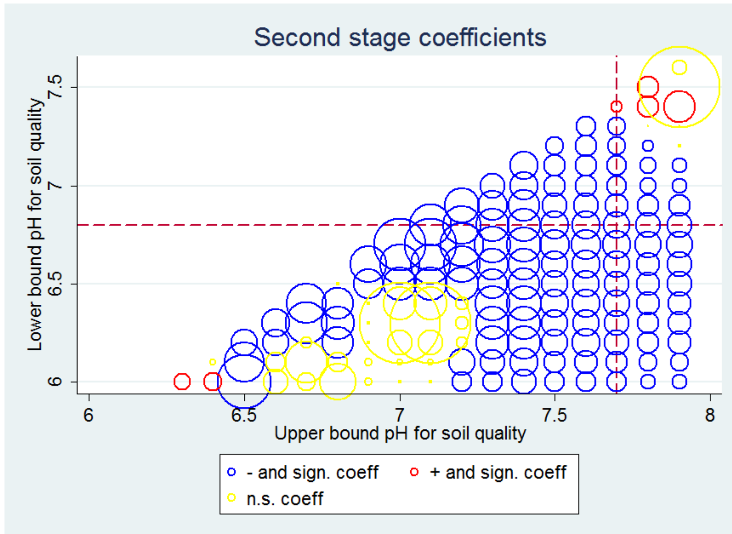


Figure H4: Coefficients of second stage regression

CoefïňĄcients derived after regressing suburbanization on city and year ïňĄxed effects instrumenting violent crime per capita by the interaction between tetraethyl lead 19 years before and the soil quality index. Every different circle refers to a different regression for every possible minimum and maximum level of pH. Robust standard errors have been used. The size of the circles refer to the absolute value of the coefïňĄcient with respect to the coefïňĄcient in the same category (- and sign.: negative and signiïňĄcant, + and sign: positive and signiïňĄcant, n.s.: non signiïňĄcant). Dashed lines indicate our chosen soil quality index: pH between 6.8 and 7.7

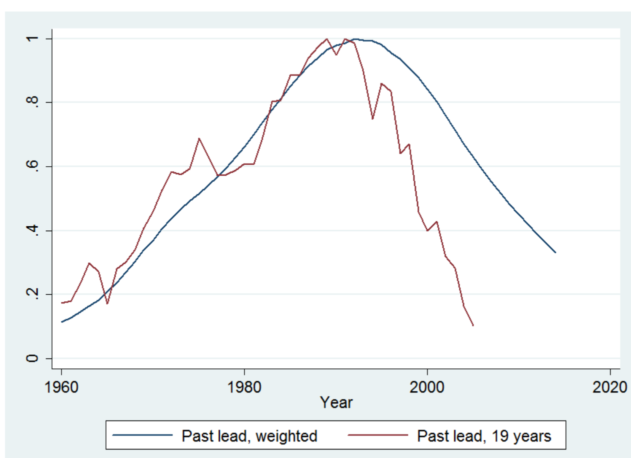


Figure H5: Time series of national past levels of lead poisoning

Past lead, weighted: sum of tonnes of past lead consumed in U.S. as gasoline additive weighted by propensity of committing crime at that particular age, normalized by the maximum level of tetraethyl lead consumption. Past lead, 19 years: tonnes of lead consumed in U.S. as gasoline additive 19 years before, normalized by the maximum level of tetraethyl lead consumption.

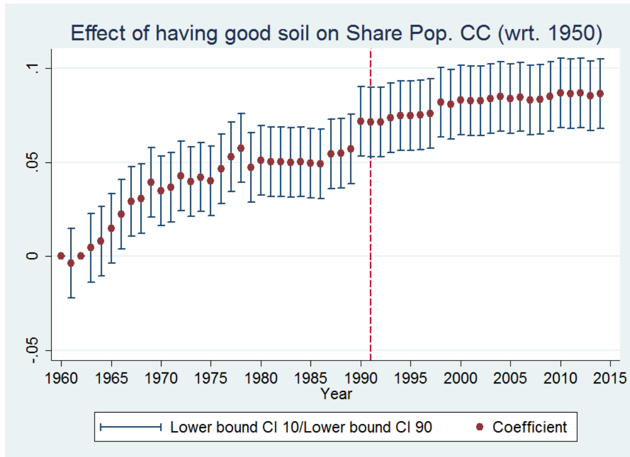


Figure H6: Time series of the effect of the good soil index on suburbanization

Share Pop CC: Proportion of MSA population living in city center. City centers with good soil: pH between 6.8 and 7.7. Lower bound CI: lower bound confidence interval at 10 % significance level. Higher bound CI: higher bound confidence interval at 10 % significance level. 1960 year dummy has been omitted. Robust standard errors have been used.

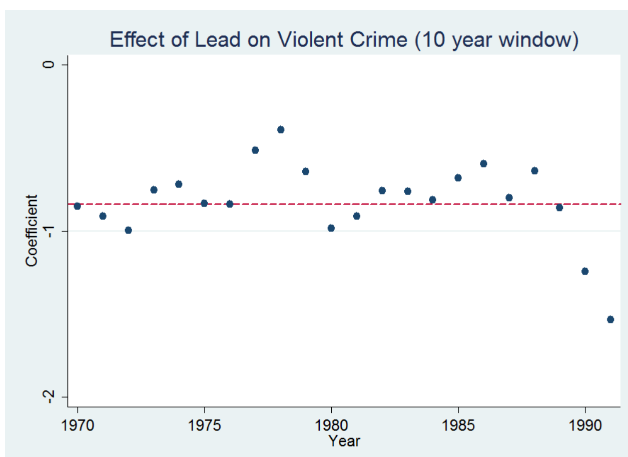
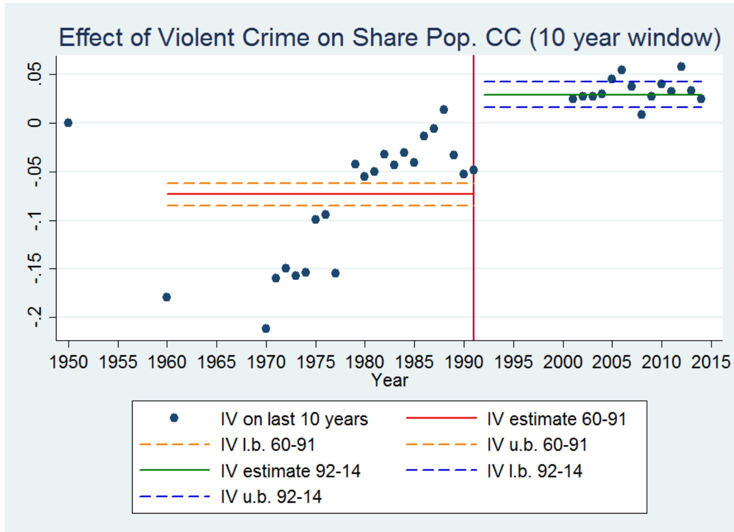
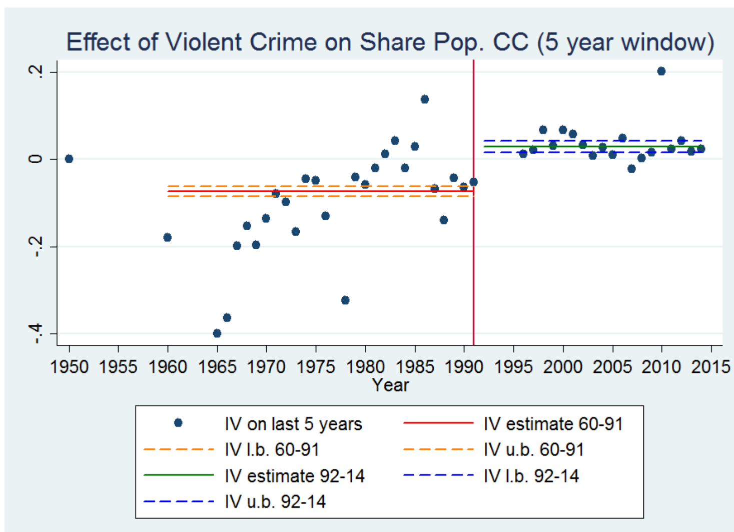


Figure H7: Time series of the effect of the good soil index on violent crime with respect to 10 years before

City centers with good soil: pH between 6.8 and 7.7. Lead: sum of tonnes of past lead consumed in U.S. as gasoline additive weighted by propensity of committing crime at that particular age, normalized by the maximum level of tetraethyl lead consumption. The coefficients refers to the ratio between the difference in the effect of good soil on violent crime in that year and 10 years before over the difference between weighted past lead poisoning in that year and 10 years before. 1960 year dummy has been omitted. Red-dashed line indicates the first stage effect of the interaction between good soil and national past lead poisoning on violent crime. Robust standard errors have been used.



(a) With respect to 10 years before



(b) With respect to 5 years before

Figure H8: Time series of the effect of the violent crime on suburbanization

Panel a): Time series of the effect of the violent crime on suburbanization with respect to 10 years before. Panel b): Time series of the effect of the violent crime on suburbanization with respect to 5 years before. IV on last X years: ratio between the difference in the effect of good soil on suburbanization in that year and X years over the difference in the effect of good soil on violent crime in that year and X years before. 1950 year dummy has been omitted. Data for share of population living in city centers in 1950 comes from Baum-Snow (2007), while data from 1960 onwards comes from F.B.I. Uniform Crime Reporting (UCR) Program Data. The effect of good soil on violent crime for the period 1950-1960 has been obtain by interpolation. All coefficient in 1960 refers to the 10 years difference with respect to 1950. Red line indicates the IV effect of violent crime on suburbanization estimated in the period 1960 to 1991. Orange dashed line refers to the corresponding confidence interval lower bound (l.b.) and higher bound (h.b.). Green line indicates the IV effect of violent crime on suburbanization estimated in the period 1992 to 2014. Blue dashed line refers to the corresponding confidence interval lower bound (l.b.) and higher bound (h.b.). Outlier coefficients have been omitted. Robust standard errors have been used.

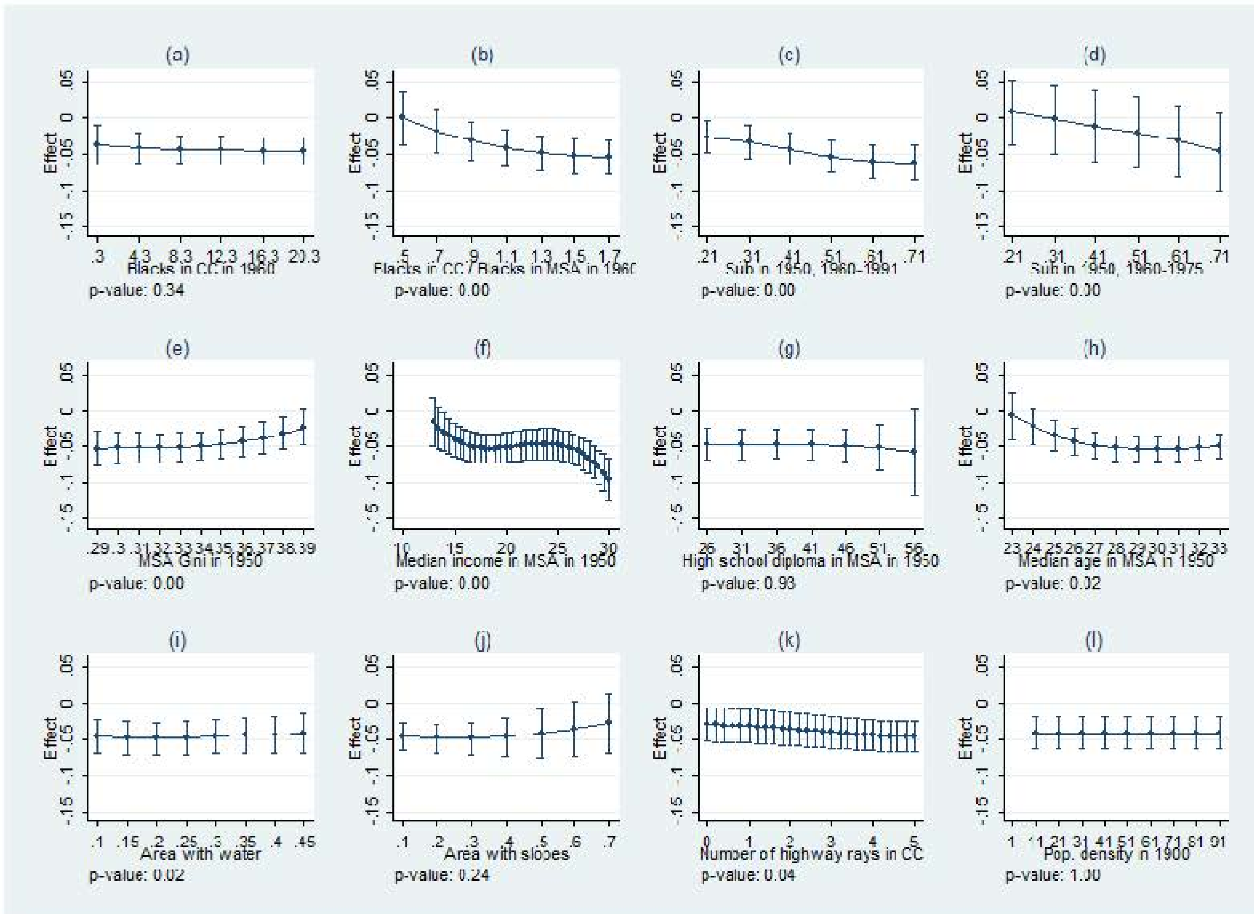


Figure H9: Heterogeneity in the effect of violent crime on suburbanization

Heterogeneity of average marginal effect of violent crime on suburbanization with respect to: Panel a): percent of black population in the city center in 1960. Panel b): ratio between the percent of black population in the city center and in the MSA in 1960. Panel c): suburbanization in 1950 for the time period from 1960 to 1991. Panel d): suburbanization in 1950 for the time period from 1960 to 1975. Panel e): simulated Gini in the MSA in 1950. Panel f): median income (in thousands) in MSA in 1950. Panel g): proportion of people with high school diploma in 1950. Panel h): median age in the MSA in 1950. Panel i): percent of 50 km radius around centroid of city center with water. Panel j): percent of 50 km radius around centroid of city center with slope higher than 15 %. Panel k): number of highway rays passing through the city center. Panel l): population density in 1900 of the corresponding county. Suburbanization (Sub): proportion of population in MSA living in city center. Violent crime (VC): Violent crime per capita in the city center. For each variable under consideration in the heterogeneity (het), marginal effects (with 95 % confidence interval) estimated regressing Sub on VC , VC x het, VC x het^2, VC x het^3, city and year fixed effects, and Census division x year. Standard errors have been bootstrapped. p-value for the test of joint significance of the coefficients of the following regressors: VC x het, VC x het^2, VC x het^3.

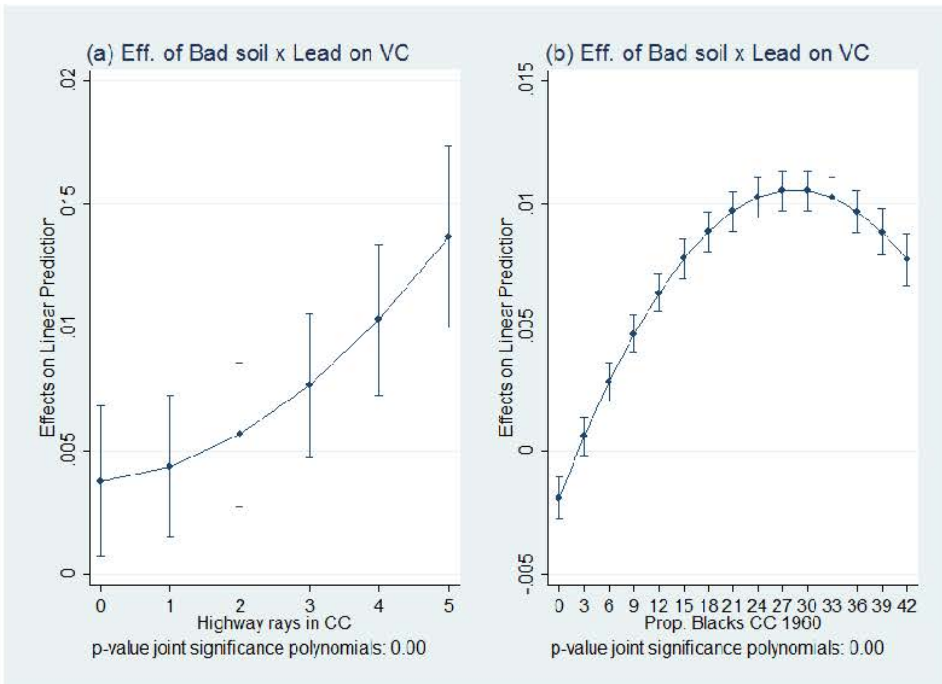


Figure H10: Heterogeneity of average marginal effect of the first stage with respect to number of highways and black population in 1960

Heterogeneity of average marginal effect of bad soil interacted with lead on violent crime with respect to Panel a): number of highways rays passing through city center. Panel b): proportion of black population in city center in the MSA in 1960. Violent crime (VC): Violent crime per capita in the city center. City centers with bad soil: pH not between 6.8 and 7.7. Lead: tonnes of lead consumed in U.S. as gasoline additive 19 years before, normalized by the maximum level of tetraethyl lead consumption. For each variable under consideration in the heteogeneity (het), marginal effects (with 95 % confidence interval) estimated regressing VC on bad soil x Lead, bad soil x Lead het, bad soil x Lead het^2, city and year fixed effects, and Census division x year. Robust standard errors have been used. p-value for the test of joint significance of the coefficients of the following regressors: bad soil x Lead het, bad soil x Lead het^2.

H.2 Tables

Table H1: Summary statistics

Variable	Mean	s.e.	Change 60-91
Pop. CC	1.9e+05	5.3e+05	22120.16***
Pop. NCC	3.3e+05	5.8e+05	1.8e+05***
Prop. Pop. CC	0.39	0.20	-0.05***
Pop. MSA	5.2e+05	9.8e+05	2.0e+05***
Area CC (sqkm)	182.22	334.76	.
Area NCC (sqkm)	332.91	592.69	.
Area MSA (sqkm)	503.59	727.35	.
pH CC	6.18	0.88	.
pH NCC	6.15	0.87	.
Tetraethyl Lead (in tonnes)	1.6e+05	72043.52	2.3e+05
Violent Crime Rate CC (per 10000)	57.80	58.08	104.52***
Murder Rate CC (per 10000)	1.10	0.91	0.86***
Rape Rate CC (per 10000)	3.83	3.44	6.72***
Robbery Rate CC (per 10000)	21.41	26.20	34.58***
Agg. Assault Rate CC (per 10000)	31.99	34.72	62.25***
Burglary Rate CC (per 10000)	158.73	98.92	135.07***
Larceny Rate CC (per 10000)	382.00	210.78	363.44***
Vehicle Theft Rate CC (per 10000)	48.96	45.66	57.79***
Total crimes CC (per 10000)	697.73	406.14	813.05***

CC: city center. MSA: metropolitan statistical area. sqkm: Squared km. per 10000: value per capita per 10,000 inhabitants. Pop.: Population. Prop.: Proportion. Aggr. assault: aggravated assaults. s.e.: Standard error. Change 60-93: Change between 1960 and 1993 obtained regressing variable under interest on dummy for year 1993 and city fixed effects for the sample of years 1960 and 1993. Since the variable tetraethyl lead does not change over cities, Change 60-93 for this variable is just the difference between the national tetraethyl lead level in 1993 and in 1960.

Table H2: Correlation between violent crime, tetraethyl lead and fixed effects

VARIABLES	(1) Violent crime	(2) Violent crime	(3) Violent crime	(4) Violent crime	(5) Violent crime
Tetraethyl lead				0.0109***	0.0110***
				(0.000216)	(0.000124)
Observations	9,515	9,515	9,515	9,515	9,515
R-squared	0.488	0.259	0.751	0.234	0.465
MSA FE	YES	NO	YES	NO	YES
Year FE	NO	YES	YES	NO	NO
Year	60-91	60-91	60-91	60-91	60-91
Estimation	OLS	OLS	OLS	OLS	OLS

Violent crime: Violent crime per capita in the city center. Tetraethyl lead: Tonnes of lead consumed in U.S. as gasoline additive 19 years before, normalized by the maximum level of tetraethyl lead consumption. Robust standard errors have been used. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table H3: Estimation of the differential effect of lead on violent crime

VARIABLES	(1) Violent crime
pH between 5 and 5.5 x Lead	0.00191
	(0.00157)
pH between 5.5 and 6 x Lead	0.00230
	(0.00182)
pH between 6 and 6.5 x Lead	0.000766
	(0.00179)
pH between 6.5 and 7 x Lead	-0.00152
	(0.00155)
pH between 7 and 7.5 x Lead	-0.00460***
	(0.00139)
pH between 7.5 and 8 x Lead	-0.000631
	(0.00215)
pH between 8 and 8.5 x Lead	-0.00230
	(0.00148)
Observations	9,515
R-squared	0.529
MSA FE	YES
Year FE	YES
Year	60-91
Estimation	OLS

Violent crime: Violent crime per capita in the city center. Lead: Tonnes of lead consumed in U.S. as gasoline additive 19 years before, normalized by the maximum level of tetraethyl lead consumption. pH category omitted: 4.5-5. Robust standard errors have been used. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table H4: Control function estimates

VARIABLES	(1) Pop CC	(2) Pop NCC	(3) Pop CC	(4) Pop NCC	(5) Pop NCC	(6) Pop NCC	(7) Pop NCC
Violent crime	-72,122***	60,800***				59,748***	73,460***
	(6,088)	(11,707)				(1,981)	(7,436)
Share high school MSA			-10,363***	4,332			1,568
			(2,449)	(5,473)			(1,832)
Good soil x Lead					-58,400***	-956.3	65,355
					(11,170)	(10,795)	(53,436)
Residuals F.S. VC	60,590***	-1,052					
	(6,174)	(11,875)					
Residuals F.S. Edu			8,274***	-3,815			
			(2,622)	(5,858)			
Observations	9,515	9,481	939	939	9,716	9,481	921
MSA FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES
Year	60-91	60-91	CY 60-90	CY 60-90	60-91	60-91	CY 60-90
Estimation	CF	CF	CF	CF	CF	CF	CF

Violent crime: Violent crime per capita in the city center. Share high school: share of population with high school diploma in the MSA. Good soil x Lead (X years before): dummy taking 1 if pH in the city center is between 6.8 and 7.7 multiplied by tonnes of lead consumed in U.S. as gasoline additive X years before, normalized by the maximum level of tetraethyl lead consumption. Residuals F.S.: residuals obtained after regressing violent crime (VC) or share high school in the MSA (Edu) on MSA and year fixed effects and the variable Good soil x Lead. Pop CC (NCC): Population in the city center (suburbs). CY: Census years. CY: census years. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table H5: Effect of good soil on potential yield of several crops

VARIABLES	(1) Corn	(2) Soy	(3) Wheat	(4) Alfalfa	(5) Cotton
pH 6.8-7.7	0.325	0.275	0.109	0.120	-0.0316
	(0.641)	(0.216)	(0.344)	(0.0735)	(0.111)
Observations	305	305	305	305	305
R-squared	0.207	0.033	0.224	0.100	0.184
C. div FE	YES	YES	YES	YES	YES
Estimation	OLS	OLS	OLS	OLS	OLS

C. div FE: Census division fixed effects. Robust standard errors have been used

Table H6: Effect of different good soils definition on violent crime

VARIABLES	(1) Violent crime	(2) Violent crime
Good soil surface x Lead	-0.551***	-0.609***
	(0.0388)	(0.0500)
Good soil all layers x Lead		0.0913
		(0.0629)
Observations	9,200	9,200
R-squared	0.772	0.772
MSA FE	YES	YES
Year FE	YES	YES
C. div x Year	YES	YES
Year	60-91	60-91
Estimation	OLS	OLS
F	201.67	148.17

Violent crime: Violent crime per capita in the city center. Good soil surface: dummy taking 1 if pH of the soil surface in the city center is between 6.8 and 7.7. Good soil all layers: dummy taking 1 if average pH of between all soil layers in the city center is between 6.8 and 7.7. Lead: tonnes of lead consumed in U.S. as gasoline additive 19 years before, normalized by the maximum level of tetraethyl lead consumption. C. div FE: Census division fixed effects. F: F-statistics of the coefficient of Good soil surface x Lead. Robust standard errors have been used. Standard errors clustered at Census division times year level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table H7: First stage for violent crime per capita in the suburbs

VARIABLES	(1) pH CC	(2) Violent crime NCC	(3) Violent crime NCC	(4) Violent crime NCC	(5) Violent crime NCC
pH NCC	0.946***				
	(0.0173)				
Good soil NCC x Lead		0.000491***	0.000491	0.000682***	0.000682
		(0.000138)	(0.000532)	(0.000193)	(0.000868)
Good soil CC x Lead				-0.000276	-0.000276
				(0.000195)	(0.000872)
Observations	294	9,518	9,518	9,518	9,518
R-squared		0.337	0.337	0.338	0.338
MSA FE	NO	YES	YES	YES	YES
Year FE	NO	YES	YES	YES	YES
Year	NO	60-91	60-91	60-91	60-91
Estimation	OLS	OLS	OLS	OLS	OLS
s.e. cluster	NO	NO	MSA	NO	MSA

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Violent crime NCC: Violent crime per capita in the suburb (NCC). Good soil CC (NCC) x Lead: dummy taking 1 if pH in the city center (suburbs) is between 6.8 and 7.7 multiplied by tonnes of lead consumed in U.S. as gasoline additive 19 years before, normalized by the maximum level of tetraethyl lead consumption. s.e.: Standard error. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table H8: First stage for violent crime per capita in the city center

VARIABLES	(1) Violent crime CC	(2) Violent crime CC	(3) Violent crime CC	(4) Violent crime CC	(5) Violent crime
Good soil CC x Lead	-0.00528***	-0.00528***	-0.00636***	-0.00636***	
	(0.000326)	(0.000819)	(0.000454)	(0.00120)	
Good soil NCC x Lead			0.00155***	0.00155	
			(0.000455)	(0.00134)	
Dummy for CC					0.00363***
					(0.000236)
Observations	9,515	9,515	9,515	9,515	19,033
R-squared	0.527	0.527	0.527	0.527	
MSA FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Year	60-91	60-91	60-91	60-91	60-91
Estimation	OLS	OLS	OLS	OLS	OLS
s.e. cluster	NO	MSA	NO	MSA	NO

Violent crime CC: Violent crime per capita in the city center (CC). Good soil CC (NCC) x Lead: dummy taking 1 if pH in the city center (suburbs) is between 6.8 and 7.7 multiplied by tonnes of lead consumed in U.S. as gasoline additive 19 years before, normalized by the maximum level of tetraethyl lead consumption. Dummy for CC: dummy taking value 1 if observation is city center. s.e.: Standard error. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table H9: Heterogeneity effect in first stage with respect to population density in 1960

VARIABLES	(1) Violent crime CC
Good soil CC x Lead	-0.00494***
	(0.000553)
Lead x Density CC 1960	1.83e-06***
	(8.02e-08)
Good soil CC x Lead x Density CC 1960	1.91e-06**
	(7.77e-07)
Observations	9,515
R-squared	0.553
MSA FE	YES
Year FE	YES
Year	60-91
Estimation	OLS

Violent crime CC: Violent crime per capita in the city center. Good soil CC : dummy taking 1 if pH in the city center is between 6.8 and 7.7. Lead: tonnes of lead consumed in U.S. as gasoline additive 19 years before, normalized by the maximum level of tetraethyl lead consumption. Density CC 1960: population density in city center in 1960. Robust standard errors have been used. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table H10: First stage regression on different crimes

VARIABLES	(1) Violent	(2) Aggr. assaults	(3) Total	(4) Murder	(5) Rape	(6) Robbery	(7) Assaults	(8) Burglary	(9) Larceny
Good soil x Lead	-0.00451**	-0.00262**	-0.0171*	-2.71e-05	-0.000234***	-0.00175**	-0.00683**	-0.00306	-0.00301
	(0.000998)	(0.000795)	(0.00544)	(1.34e-05)	(2.71e-05)	(0.000326)	(0.00181)	(0.00197)	(0.00186)
Observations	9,484	8,338	9,484	9,484	9,428	9,484	9,484	9,484	9,484
MSA FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
C. reg x Year	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year	60-91	60-91	60-91	60-91	60-91	60-91	60-91	60-91	60-91
Estimation	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
F	20.37	10.84	9.88	4.12	74.64	28.95	14.17	2.41	2.63

All crime refers to per capita crime rate in city center. All crimes are per capita in the city center. Tot: total of the category under analysis. Good soil : dummy taking 1 if pH in the city center is between 6.8 and 7.7. Lead: tonnes of lead consumed in U.S. as gasoline additive 19 years before, normalized by the maximum level of tetraethyl lead consumption. Aggr. assault: aggravated assaults. C. reg: Census region fixed effects. F: F-statistics on the excluded instruments. Standard errors clustered at Census region level in parentheses . *** p<0.01, ** p<0.05, * p<0.1

Table H11: Correlation between highways and interaction between lead and good soil

VARIABLES	(1) Good soil x Lead	(2) Good soil x Lead	(3) Rays in plan	(4) Rays
Rays in plan x Prop. hgw	-0.000486			
	(0.00785)			
Rays		-0.00527		
		(0.00715)		
Good soil			-0.0150	0.0212
			(0.251)	(0.184)
Observations	939	939	235	1,174
MSA FE	YES	YES	NO	NO
Year FE	YES	YES	NO	YES
Year	CY 60-90	CY 60-90	CY 60	CY 60-90
Sample	Restricted	Restricted	Restricted	Restricted
Estimation	OLS	OLS	OLS	OLS

For the notes see Table H12

Table H12: Results controlling for instrumented highways construction

VARIABLES	(1) Sh pop CC	(2) Rays	(3) Sh pop CC	(4) V. crime	(5) Sh pop CC	(6) Rays	(7) V. crime
Rays	-0.0490***				0.0137		
	(0.0139)				(0.0312)		
Rays in plan x Prop. hgw		0.598***				0.334***	0.238***
		(0.0466)				(0.0529)	(0.0599)
Violent crime			-0.0759***		-0.0826**		
			(0.0198)		(0.0334)		
Good soil x Lead				-1.242***		-0.158	-1.082***
				(0.174)		(0.258)	(0.240)
Observations	939	1,174	1,184	1,189	921	939	921
MSA FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES
Year	CY 60-90	CY 60-90	CY 60-90	CY 60-90	CY 60-90	CY 60-90	CY 60-90
Sample	Restricted	Restricted	Restricted	Restricted	Restricted	Restricted	Restricted
Estimation	IV	OLS	IV	OLS	IV	OLS	OLS
F	140.25	.	49.12	.	7.02	.	.

Sh Pop CC: Proportion of MSA population living in city center. Violent (V.) crime: Violent crime per capita in the city center standardized. Good soil x Lead: dummy taking 1 if pH in the city center is between 6.8 and 7.7 multiplied by tonnes of lead consumed in U.S. as gasoline additive 19 years before, normalized by the maximum level of tetraethyl lead consumption. hgw: highways. Rays: actual highway rays built in the city center. Rays in plan x Prop. hgw: planned highway rays in the city center multiplied by number of highways built in that year normalized by total number of highways built in 2000. CY: Census Year. Sample restricted: all MSAs of at least 100,000 people with central cities of at least 50,000 people in 1950. F: F-statistics on the excluded instruments. Stock-Yogo weak ID test critical value at 10% maximal IV size: 7.03. Standard errors clustered at MSA level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table H13: Results controlling for MSA socioeconomic characteristics

VARIABLES	(1) Share pop CC	(2) Share pop CC	(3) Share pop CC	(4) Share pop CC	(5) Share pop CC	(6) Share pop CC	(7) Share pop CC
Violent crime	-0.0686***	-0.108***	-0.0756***	-0.0865***	-0.0673**	-0.0690**	-0.0685*
	(0.0206)	(0.0368)	(0.0242)	(0.0301)	(0.0294)	(0.0298)	(0.0370)
Share pop. CC 1950 x Trend		-0.00880***					
		(0.00241)					
% Blacks MSA			0.00309	0.00422	0.00288	0.00296	0.00354
			(0.00255)	(0.00308)	(0.00274)	(0.00279)	(0.00334)
Median family income				3.12e-06*	1.66e-06	2.01e-06	-2.46e-06
				(1.60e-06)	(1.14e-06)	(1.25e-06)	(3.19e-06)
% people over 25					0.0836	0.0783	0.0865
					(0.0871)	(0.0892)	(0.0878)
% people over 65						0.00230	0.00213
						(0.00345)	(0.00355)
Median gross rent							0.000318
							(0.000239)
Observations	1,184	922	921	921	691	691	691
MSA FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES
C. reg x Year FE	YES	YES	YES	YES	YES	YES	YES
Year	CY 60-90	CY 60-90	CY 60-90	CY 60-90	CY 70-90	CY 70-90	CY 70-90
Estimation	IV	IV	IV	IV	IV	IV	IV
F	17.27	8.90	13.04	10.03	6.31	6.37	4.18

Share pop CC: proportion of population in MSA living in city center. Violent crime: Violent crime per capita in the city center standardized. C. reg: Census Region fixed effects. CY: Census Year. F: F-statistics on the excluded instruments. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table H14: Second stage results using other functional form of pH

VARIABLES	(1) Share pop CC	(2) Share pop CC	(3) Share pop CC	(4) Share pop CC	(5) Share pop CC	(6) Share pop CC
Violent crime	-0.0717***	-0.0575***	-0.0486***	-0.0487***	-0.0435***	-0.0424***
	(0.00594)	(0.00506)	(0.00644)	(0.00644)	(0.00543)	(0.00540)
Observations	9,481	9,481	9,481	9,481	9,481	9,481
MSA FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Year	60-91	60-91	60-91	60-91	60-91	60-91
Estimation	IV	IV	IV	IV	IV	IV
pH used	6.8-7.7	0.5 bounds	1st poly	2nd poly	3rd poly	4th poly

Share pop CC: proportion of population in MSA living in city center. Violent crime: Violent crime per capita in the city center standardized. The pH dummies used in the first stage in the category 0.5 bounds are: pH between 5 and 5.5, pH between 5.5 and 6, pH between 6 and 6.5, pH between 6.5 and 7, pH between 7 and 7.5, pH between 7.5 and 8, pH between 8 and 8.5. The variables used in the first stage in the category 1st poly are: pH x past national lead poisoning. The variables used in the first stage in the category 2nd poly are: pH x past national lead poisoning, pH^2 x past national lead poisoning. The variables used in the first stage in the category 3rd poly are: pH x past national lead poisoning, pH^2 x past national lead poisoning, pH^3 x past national lead poisoning. The variables used in the first stage in the category 4th poly are: pH x past national lead poisoning, pH^2 x past national lead poisoning, pH^3 x past national lead poisoning, pH^4 x past national lead poisoning. Robust standard errors have been used. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table H15: Results using different lags of lead as instrument

VARIABLES	(1) Sh pop CC	(2) Sh pop CC	(3) Sh pop CC	(4) Sh pop CC	(5) Sh pop CC	(6) Sh pop CC	(7) Sh pop CC	(8) Sh pop CC	(9) Sh pop CC
Violent crime	-0.0836***	-0.0821***	-0.0779***	-0.0747***	-0.0717***	-0.0677***	-0.0651***	-0.0626***	-0.0575***
	(0.00790)	(0.00769)	(0.00720)	(0.00652)	(0.00594)	(0.00591)	(0.00594)	(0.00607)	(0.00600)
Observations	8,329	8,602	8,861	9,158	9,481	9,149	8,854	8,594	8,296
MSA FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year	60-91	60-91	60-91	60-91	60-91	60-91	60-91	60-91	60-91
Estimation	IV	IV	IV	IV	IV	IV	IV	IV	IV
Lag lead	15 years	16 years	17 years	18 years	19 years	20 years	21 years	22 years	23 years
F	192.30	197.18	209.77	235.96	264.55	250.90	236.26	216.00	204.64
FS Beta	-0.798	-0.799	-0.814	-0.861	-0.909	-0.925	-0.926	-0.927	-0.953
FS s.e.	0.058	0.057	0.056	0.056	0.056	0.059	0.061	0.064	0.067

Sh pop CC: proportion of population in MSA living in city center. Violent crime: Violent crime per capita in the city center standardized. Lag lead X years: instrument used is dummy taking 1 if pH in the city center is between 6.8 and 7.7 multiplied by tonnes of lead consumed in U.S. as gasoline additive X years before, normalized by the maximum level of tetraethyl lead consumption. F: F-statistics on the excluded instruments. FS beta: first stage coefficient of the instrument. FS s.e: first stage standard error of the instrument. Robust standard errors have been used. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table H16: Results using different past variables of lead poisoning as instrument

VARIABLES	(1) Share pop CC	(2) Violent crime	(3) Share pop CC	(4) Violent crime
Violent crime	-0.0717***		-0.0711***	
	(0.00594)		(0.00573)	
Good soil x Lead 19 years		-0.909***		
		(0.0561)		
Good soil x Lead weighted				-0.866***
				(0.0517)
Observations	9,481	9,515	9,481	9,515
MSA FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Year	60-91	60-91	60-91	60-91
Estimation	IV	OLS	IV	OLS
F	264.55	.	283.02	.

Share pop CC: proportion of population in MSA living in city center. Violent crime: Violent crime per capita in the city center standardized. Good soil x Lead 19 years: dummy taking 1 if pH in the city center is between 6.8 and 7.7 multiplied by tonnes of lead consumed in U.S. as gasoline additive 19 years before, normalized by the maximum level of tetraethyl lead consumption. Good soil x Lead weighted: dummy taking 1 if pH in the city center is between 6.8 and 7.7 multiplied by sum of tonnes of past lead consumed in U.S. as gasoline additive weighted by propensity of committing crime at that particular age, normalized by the maximum level of tetraethyl lead consumption. F: F-statistics on the excluded instruments. Robust standard errors have been used. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table H17: Results controlling for different geographical fixed effect

VARIABLES	(1) Share pop CC	(2) Share pop CC	(3) Share pop CC	(4) Share pop CC	(5) Share pop CC	(6) Share pop CC
Violent crime	-0.422***	-0.223***	-0.0717***	-0.0573***	-0.0843***	-0.0803***
	(0.0471)	(0.0154)	(0.00594)	(0.00713)	(0.0120)	(0.0148)
Observations	9,481	9,481	9,481	9,481	9,481	9,360
MSA FE	NO	NO	YES	YES	YES	YES
Year FE	NO	YES	YES	YES	YES	YES
C. region X Year FE	NO	NO	NO	YES	NO	NO
C. division X Year FE	NO	NO	NO	NO	YES	NO
State x Year FE	NO	NO	NO	NO	NO	YES
Year	60-91	60-91	60-91	60-91	60-91	60-91
Estimation	IV	IV	IV	IV	IV	IV
F		.	264.55	154.13	74.90	47.68

Share pop CC: proportion of population in MSA living in city center. Violent crime: Violent crime per capita in the city center standardized. C. region: Census region. C. division: Census division. F: F-statistics on the excluded instruments. Robust standard errors have been used. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table H18: Results controlling for different physical fixed effect

VARIABLES	(1) Share pop CC	(2) Share pop CC	(3) Share pop CC	(4) Share pop CC
Violent crime	-0.0843***	-0.0825***	-0.0740***	-0.125***
	(0.0120)	(0.0102)	(0.0173)	(0.0316)
Observations	9,481	8,579	8,597	6,630
MSA FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
C. division x Year FE	YES	YES	YES	YES
Slope pctl X Year FE	NO	YES	NO	NO
Precipitation pctl X Year FE	NO	NO	YES	NO
Distance pctl X Year FE	NO	NO	NO	YES
Year	60-91	60-91	60-91	60-91
Estimation	IV	IV	IV	IV

Share pop CC: proportion of population in MSA living in city center. Violent crime: Violent crime per capita in the city center standardized. C. division: Census division. Slope pctl: Percentile of distribution of MSA in terms of slope. Precipitation pctl: Percentile of distribution of MSA in terms of atmospheric precipitation. Distance pctl: Percentile of distribution of MSA in terms of distance to water or a border of the U.S. F: F-statistics on the excluded instruments. Robust standard errors have been used. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table H19: Results with different subsamples in function of population size

VARIABLES	(1) Share pop CC	(2) Share pop CC	(3) Share pop CC
Violent crime	-0.0717***	-0.123***	-0.167***
	(0.00594)	(0.0133)	(0.0177)
Observations	9,481	6,826	9,481
MSA FE	YES	YES	YES
Year FE	YES	YES	YES
Year	60-91	60-91	60-91
Sample	All	More than 100000	All
Weight	NO	NO	Pop. MSA
Estimation	IV	IV	IV
F	264.55	95.67	90.31

Share pop CC: proportion of population in MSA living in city center. Violent crime: Violent crime per capita in the city center standardized. More than 100,000: Sample using cities with population higher than 100,000 inhabitants. Weight: variable for weighting used. F: F-statistics on the excluded instruments. Robust standard errors have been used. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table H20: Results using different clustered standard errors

VARIABLES	(1) Share pop CC	(2) Share pop CC	(3) Share pop CC	(4) Share pop CC	(5) Share pop CC	(6) Share pop CC	(7) Share pop CC
Violent crime	-0.0843***	-0.0843*	-0.0843***	-0.0843***	-0.0843***	-0.0843***	-0.0843***
	(0.0120)	(0.0462)	(0.00605)	(0.0131)	(0.0134)	(0.0109)	(0.0167)
Observations	9,481	9,481	9,481	9,481	9,481	9,481	9,481
R-squared	0.934	0.934	0.934	0.934	0.934	0.934	0.934
MSA FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES
C. div x Year FE	YES	YES	YES	YES	YES	YES	YES
Year	60-91	60-91	60-91	60-91	60-91	60-91	60-91
Estimation	IV	IV	IV	IV	IV	IV	IV
s.e. cluster	NO	MSA	Year	MSA x Year	State x Year	C. reg x Year	C. div x Year
F	74.90	77.47	77.32	74.90	74.90	77.06	77.06

Share pop CC: proportion of population in MSA living in city center. Violent crime: Violent crime per capita in the city center standardized. C. div: Census division. C. reg: Census region. F: F-statistics on the excluded instruments. s.e. cluster: cluster level of standard errors. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table H21: The effect of crime on suburbanization using different year samples

VARIABLES	(1) Share Pop CC	(2) Share Pop CC	(3) Share Pop CC
Violent crime	-0.0869***	0.0304***	-0.0564***
	(0.0172)	(0.00730)	(0.0180)
Observations	9,481	4,028	13,513
MSA FE	YES	YES	YES
Year FE	YES	YES	YES
C. div x Year	YES	YES	YES
Year	60-91	92-14	60-14
Estimation	IV	IV	IV
F	77.06	13.67	70.63

Share Pop CC: Proportion of MSA population living in city center. Violent crime: Violent crime per capita in the city center standardized. F: Cragg-Donald Wald F-statistics on the excluded instruments. Robust standard errors have been used. C. div: Census Division fixed effects. Standard errors clustered at Census division times year level in parentheses. *** p<0.01, ** p<0.05, * p<0.1