All roads lead to Rome ... and to sprawl? Evidence from European cities

Miquel-Àngel Garcia-López[*][†]
Universitat Autònoma de Barcelona and *Institut d'Economia de Barcelona*

August 2019

ABSTRACT: Over the last 15 years, we have learned that highways cause suburbanization and population growth. However, little is known about the resulting residential land use patterns. This paper aims to fill this gap by being the first to comprehensively analyze the effects of highways on sprawl. Using data from the Corine Land Cover project and the highway network for 579 European cities in 1990, 2000 and 2012, I find that a 10% increase in the stock of highways (km) causes a 1.1% growth in the residential land area, a 2% growth in the number of residential lots, and an increase of 25 p.p. in the share of undeveloped land surrounding residential land over this 20-year period. Overall, these results confirm that highways also cause residential sprawl by expanding cities with new, more fragmented and more isolated land developments.

---

[*] I am very grateful to Gabriel Ahlfeldt, Nathaniel Baum-Snow, Jorge de la Roca, Gilles Duranton, Laurent Gobillon, Hans Koster, Alexander C. Lembcke, Marta MJ, Joel and Maria MG, Henry Overman, Diego Puga, Albert Solé-Ollé, Matt Turner, Josh van Ommeren, Paolo Veneri, Elisabet Viladecans-Marsal, and seminar and conference participants for helpful comments and discussions. Financial support from the Ministerio de Ciencia e Innovación (research projects ECO2014-52999-R and RTI2018-097401-B-I00), Generalitat de Catalunya (research project 2017SGR1301), and the "Xarxa de Referència d'R+D+I en Economia Aplicada" is gratefully acknowledged.

[†] Corresponding author. Department of Applied Economics, Universitat Autònoma de Barcelona, Edifici B, Facultat d'Economia i Empresa, 08193 Cerdanyola del Vallès, Spain (e-mail: `miquelangel.garcia@uab.cat`; phone: +34 93 581 4584; website: `http://gent.uab.cat/miquelangelgarcialopez`).

1. Introduction

Sprawl and suburbanization are well-known concepts which, while related, differ. Suburbanization refers to changes in the location patterns of population within cities and, in particular, to a population shift from central cities into their suburbs. In terms of land development, this relocation can present three patterns. First, suburbanizing population can settle in the suburbs on previously developed residential land, increasing its intensity of use, but without changing the initial spatial pattern of residential land. Second, new residential land can be developed next to previous developments, expanding city (residential) size in a continuous and compact fashion. Finally, new residential land developments can be more fragmented, scattered and isolated throughout the suburbs. While the first two spatial patterns are related to compact land development, the second describes a pattern of sprawl.

As Ahlfeldt and Pietrostefani (2017, 2019) and OECD (2018) highlight, the aforementioned land development patterns and, in particular, that of sprawl are at the heart of policy debate because of their economic, environmental and social impacts. This debate is ongoing in Europe, where cities have traditionally been settlements with higher density and more continuous land developments than their American counterparts, and where sprawl has generally been considered a US phenomenon. However, European cities were more compact and sprawled less in the mid-1950s than they do today. Between 1990 and 2012, their overall residential land increased by 12%[1] and their density decreased by 3%. New residential land developments were more fragmented and the number of residential lots increased by 17%. Since these recent trends are increasing, concerns about the future compactness of Europe's cities have been expressed (Couch, Leontidou, and Petschel-Held, 2007, Arribas-Bel, Nijkamp, and Scholten, 2011). Since this has important implications from a policy perspective (European Environment Agency, 2006, 2010, OECD, 2018), understanding the causes of sprawl is important.

Over the last 15 years we have learned that highways cause both suburbanization and population growth. However, little is known about the resulting residential land use patterns. This paper aims to fill this gap by being the first to comprehensively analyze the effects of highways on sprawl. I empirically answer the following three questions: Do highways expand cities with new residential land developments? Do they encourage fragmented or compact residential developments? Do they foster more isolated residential developments, with more undeveloped surroundings? Using data from the Corine Land Cover (CLC) project and the highway network for 579 European cities in 1990, 2000 and 2012, I rely on an instrumental variables (IV) strategy that uses Europe's historical transportation networks as sources of exogenous variation. I find that a 10% increase in the stock of highways (km) causes a 1.1% growth in the residential land area, a 0.6% reduction in the city residential density, a 2% growth in the number of residential lots (0.9% in the ratio of lots per sq km), and an increase of 25 p.p. in the share of undeveloped land surrounding residential land. Overall, these results confirm that highways also cause residential sprawl by expanding cities with more disconnected residential land use patterns based on lower

---

[1] This increase is similar to the 17% increase in overall developed (urban) land found in the US between 1990 and 2007 (Nickerson, Ebel, Brochers, and Carriazo, 2011).

density developments.

Although this study is the first attempt to examine the effects of highways on sprawl, the scarce literature on this topic has previously analyzed some of its dimensions. For example, Brueckner and Fansler (1983), Deng, Huang, Roxell, and Uchida (2008), McGrath (2005) focus on the spatial sizes of cities in terms of developed land area. Burchfield, Overman, Puga, and Turner (2006) and Angel, Parent, and Civco (2012) analyze the isolation of land developments using an indicator that measures the percentage of undeveloped land surrounding developed land. Finally, Oueslati, Alvanides, and Garrod (2015) study the size of the developed land area and its fragmentation in different land lots. Since area, fragmentation and isolation are dimensions, which when taken together, characterize sprawl, in this study I analyze all three: the expansion of cities with new land developments in a scattered (compact) pattern and with an increase (decrease) in their undeveloped surroundings.

The aforementioned papers have studied the determinants of sprawl including variables related to highways. Brueckner and Fansler (1983), McGrath (2005) and Angel *et al.* (2012) use 'indirect' indicators such as the percentage of households owning automobiles and the CPI for private transportation. Burchfield *et al.* (2006) focus on the density of major suburban roads, neglecting the effect of central roads. Deng *et al.* (2008) and Oueslati *et al.* (2015) use the density of highways at spatial scales (region, county) bigger than their city boundaries (FUA, urban core) and, as a result, include additional information that might bias their results. In this paper, I use the length of the highway network at the metropolitan level to explain sprawl at the metropolitan level. With this variable, I consider not only commuting costs (as do Brueckner and Fansler, 1983, McGrath, 2005, Angel *et al.*, 2012), but also the size of the highway network (as do Burchfield *et al.*, 2006, Deng *et al.*, 2008, Oueslati *et al.*, 2015). Finally, while the above papers assume that their transportation variables are exogenous to land development, in this paper I address endogeneity concerns relying on IV techniques.

In general, this paper furthers our understanding of the effects of transportation. Recent research shows that highways shape cities. Besides fostering urban growth (Duranton and Turner, 2012) and population suburbanization (Baum-Snow, 2007, Garcia-López, Holl, and Viladecans-Marsal, 2015*a*, Garcia-López, Pasidis, and Viladecans-Marsal, 2018), highways also cause employment decentralization (Baum-Snow, Brandt, Henderson, Turner, and Zhang, 2017), spread suburban population and employment out along their ramps (Garcia-López, 2012, Garcia-López *et al.*, 2015*a*, Garcia-López, Hémet, and Viladecans-Marsal, 2017*a*,*b*), and modify local zoning policies (Garcia-López, Solé-Ollé, and Viladecans-Marsal, 2015*c*). However, little is known about the role of highways in causing cities to sprawl in relation to characteristics of land development.

The rest of the paper is structured as follows. In Section 2, I describe the residential sprawl phenomenon in Europe and its cities, and the highway network and other ancient (rail)roads. In Section 3, I review the theoretical and empirical literature. The empirical strategy is discussed in Section 4. In Sections 5.1, 5.2 and 5.3, I answer the three questions about the relationship between highways and sprawl. In Section 6, I test the heterogeneity and robustness of the main results. Finally, I present my conclusions in Section 7.

2. Sprawl and highways in Europe

I use the Functional Urban Area (FUA) (formerly known as Larger Urban Zone (LUZ)), defined by the European Commission (Urban Audit Project) and the OECD as the unit of observation. In common with the Metropolitan Statistical Area in the US, the FUA consists of a central city (with at least 50,000 inhabitants) and a commuting zone (made up of all municipalities with at least 15% of their employed residents working in the city[2].

My dataset includes 579 FUAs located in 29 European countries: 26 member states of the European Union (EU)[3] and Switzerland, Norway and Iceland. It is important to note that my sample of cities is quite heterogeneous (as I show below with summary statistics) in terms of their political, cultural and geographical characteristics. Of the FUAs, 171 are located in Mediterranean countries (Greece, Spain, Southern France, Italy, Malta and Portugal); 156 are located in Eastern Europe (East Germany, Bulgaria, Czech Republic, Estonia, Croatia, Hungary, Latvia, Poland, Romania, and Slovakia); and 252 are located in Northwestern countries (Austria, Belgium, Switzerland, Germany, Denmark, Finland, Northern France, Ireland, Iceland, Luxembourg, Latvia, The Netherlands, Norway, Sweden, West Germany and the United Kingdom).

2.1 *Sprawl in Europe*

To measure sprawl in terms of residential land developments, I use land data from the CLC project. Coordinated by the European Environment Agency, the project integrates CLC databases from 27 to 39 European countries for various years. The CLC is produced for the majority of countries by visual interpretation of high resolution satellite imagery, using a minimum mapping unit of 25 ha for areal phenomena (5 ha for changes in land cover layers) and a minimum width of 100 m for linear phenomena[4].

The CLC database and related GIS maps are available for several years starting in 1990 and classify land in 44 classes. Eleven classes are labeled as 'Artificial surfaces' and include man-made land covers related to residential settlements, economic activity and artificial (non-agricultural) vegetated areas. Jointly, all these land cover types define and measure 'Developed land'.

Among the eleven classes of developed land, I focus on those primarily related to 'Residential land'. Specifically, I consider the two 'Urban fabric' classes, which refer to land mainly occupied by dwellings and buildings, including their connected areas (associated lands, streets, parking lots, etc.). These classes include both continuous and discontinuous settlements. The former refer to land with more than 80% of the surface covered by buildings and other impermeable features and where areas of vegetation and bare soil are exceptional. The latter refer to land with 30 to 80% of the surface covered by buildings and other artificial surfaces and with a significant coverage of vegetated areas and bare surfaces in a discontinuous spatial pattern.

---

[2] See `https://ec.europa.eu/eurostat/web/cities/spatial-units` and https://ec.europa.eu/eurostat/statistics-explained/index.php/Archive:European_cities_–_the_EU-OECD_functional_urban_area_definition for more details.

[3] Slovenia and Lithuania are not included because their population datasets were not available.

[4] Further details on the method and labeling processes can be found at `https://land.copernicus.eu/user-corner/technical-library/corine-land-cover-nomenclature-guidelines/docs/pdf/CLC2018_Nomenclature_illustrated_guide_20170930.pdf`.

Besides 'Residential land', in some additional analyses I also consider the aforementioned total 'Developed land' and, in particular, the 'Industrial and commercial units' land, which is directly related to economic activity and includes all kinds of industrial and commercial buildings and connected areas. I ignore the other eight 'Artificial' classes because they are highly heterogeneous (e.g., 'Mineral extraction sites', 'Dump sites', and 'Construction sites') and also consider land directly related to transportation infrastructure as 'Road and rail networks and associated land', 'Port areas', and 'Airports' classes.

In line with the empirical literature (Brueckner and Fansler, 1983, Burchfield *et al.*, 2006, Angel *et al.*, 2012), residential sprawl, in common with any other spatial pattern of land development, can be characterized in relation to three (spatial) dimensions: The 'size of land developments', the 'degree of fragmentation', and the 'degree of undeveloped surroundings'. This first dimension allows me to determine city size in terms of residential land increases or decreases. I use the CLC vectors maps to measure this and its evolution for 1990, 2000 and 2012. Since residential land consumption is directly related to city population size (and its evolution), I alternatively measure the size dimension with the residential density (computed as the ratio between population and residential land area), that is, the number of inhabitants per sq km of residential land. To do so, I use population data from 1990, 2000 and 2010 Population Censuses.

The second dimension, the 'degree of fragmentation', shows just how fragmented (new) land developments are. It is measured using the number of residential land lots which, in the CLC vector maps, are identified as discontinuous polygons. Since the number of land lots is directly related to the overall size of the residential land area and, moreover, since the size of residential lots also matters, I alternatively measure the fragmentation dimension with the ratio between residential lots and their area, that is, the number of residential land lots per sq km of residential land.

Finally, the third dimension indicates just how scattered (new) land developments are and how isolated they are from other developments. To measure this, I compute the sprawl index proposed by Burchfield *et al.* (2006): The percentage of undeveloped land surrounding residential land. To do so, I use the CLC raster maps (100 m resolution) for 1990, 2000 and 2012. For each residential cell I compute the percentage of undeveloped land in the surrounding sq km. The index for each FUA is computed averaging across all residential cells in each FUA[5].

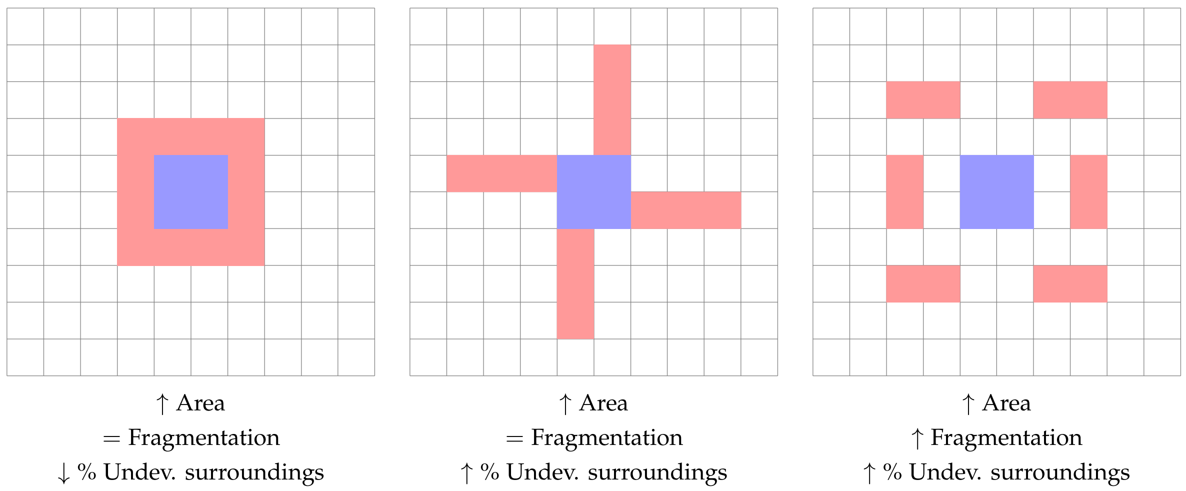
Figure 1 shows three alternative spatial patterns of land development that might emerge from a hypothetical population suburbanization process: Compact (Figure 1a), linear (Figure 1b) and leapfrog (Figure 1c). The indicators of each dimension change according to the type of spatial development. This means they can be distinguished one from another and the degree of respective sprawl measured. For the same increase in residential land area (size dimension), residential sprawl is greatest in the case of leapfrog development (Figure 1c) because both the number of residential lots and the percentage of undeveloped surroundings increase (higher degrees of fragmentation and isolation). It is followed by the case of linear development (Figure 1b) because, although the number of residential lots does not change, the percentage of undevel-

---

[5] Compared to Burchfield *et al.* (2006), our cell resolution is three times larger (100 m vs. 30 m), and, as a result, our measure is less accurate.

oped surroundings increases (higher degree of isolation). Finally, although the case of compact development (Figure 1a) presents the least sprawl[6], we may wish to compare its initial and final patterns. To do so, we also need to consider the evolution of city population by computing the initial and final residential density: If it decreases, the final pattern presents a greater degree of sprawl than the initial one.

Figure 1: Spatial patterns of land developments and the three dimensions


(a) Compact development

(b) Linear development

(c) Leapfrog development

Notes: The blue areas refer to the initial land development (a continuous and compact pattern). The red areas refer to the new developed land and, jointly with the blue areas, the resulting spatial pattern of land development. Figure 1a shows that new developments take place next to previous ones, expanding the city in a continuous and compact fashion. In Figure 1b the city expands in a linear fashion following, for example, a new transportation infrastructure. Finally, Figure 1c depicts a case of leapfrog development, that is, a pattern in which new developments are discontinuous and isolated from earlier developments.

Table 1 reports the computations of these indicators for Europe (first four columns) and for the whole sample of 579 European cities (last four columns). These cities cover 20% of the almost 5 million sq km of Europe's total land area (first row). Although developed land (second row) increased by 15% between 1990 and 2012, it represented only 4 and 10% of Europe and the FUA's total land in 2012.

Residential land (third row) is the most abundant type of developed land (more than 70%) and it grew at a similar rate (12-13%) in both Europe and its cities. Figure 2 shows additional details of the spatial pattern of residential land in Europe in 2012 (Figure 2a) and its evolution in Barcelona between 1990 and 2012 (Figure 2b). Compared to the US, Europe shows higher residential density levels (fourth row): more than 3,000 inhabitants per sq km. But, despite the growth in population in this period, the increase in residential land was more marked and, as a result, density fell in Europe and its cities by 200 and 130 inhabitants per sq km (-6 and -3%), respectively.

---

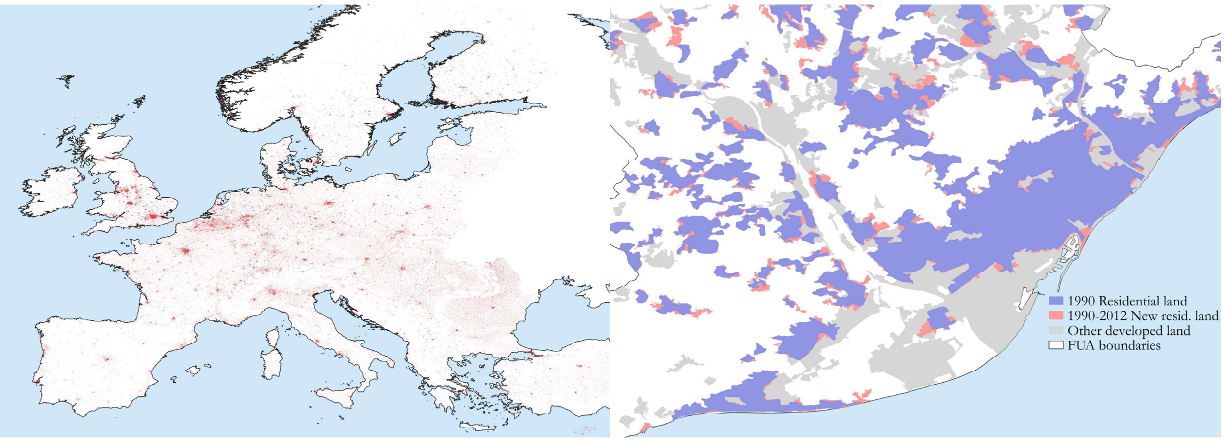
[6] It is important to note that, in this case, the percentage of undeveloped surroundings decreases by construction. That is, since the land area expands in a continuous and compact way, the number of residential cells without undeveloped surroundings increases, and, as a result, the average value (across all residential cells) decreases.

Table 1: Residential land in Europe and its cities

	Europe (29)				Functional Urban Areas			
	1990	2000	2012	1990–2012	1990	2000	2012	1990–2012
Total land area	4,851,351	4,851,351	4,851,351		976,178	976,178	976,178	
Developed area	185,458	193,837	213,309	27,851 (15%)	87,953	92,212	101,043	13,090 (15%)
Residential area	139,334	144,634	156,691	17,357 (13%)	63,622	66,230	71,162	7,540 (12%)
Pop/Res area	3,473	3,360	3,271	-201.9 (-6%)	4,344	4,300	4,212	-131.3 (-3%)
Fragmentation	121,270	125,376	142,794	21,524 (18%)	43,000	44,974	50,106	7,106 (17%)
Fragm/Res area	0.87	0.87	0.91	0.04 (5%)	0.68	0.65	0.71	0.03 (4%)
Surroundings	40.7	38.9	40.6		36.9	36.2	36.5	

Notes: 'Total land area', 'Developed land area' and 'Residential area' are measured in square kilometers. 'Population/Res area' is the number of inhabitants per sq km of residential land, 'Fragmentation' refers to the number of residential land lots and 'Fragm/Res area' is the number of residential land lots per sq km, and 'Surroundings' refers to the % of undeveloped land surrounding residential land.

Figure 2: CORINE Residential land


(a) Europe, 2012

(b) Barcelona, 1990–2012

The fragmentation of residential land (fifth and sixth rows) is quite high, with more than 143,000 and 50,000 lots in Europe and its cities, respectively, in 2012, that is, 0.9 and 0.7 lots per sq km of residential land, respectively. Furthermore, compared to 1990, the degree of fragmentation increased during the period.

Finally, the degree of isolation (last row) is also high in Europe and its cities: The percentage of undeveloped surroundings was 36.5% in 2012. In other words, in a square km surrounding an average residential development, 0.365 km^2 were undeveloped and the remaining 0.635 km^2 were developed. Although this share decreased slightly (down 0.4 p.p. i.e. 0.004 km^2 less undeveloped land in each square km) between 1990 and 2012, it is important to note that is behavior between 1990 and 2000 (decrease) and between 2000 and 2012 (increase) differed.

Table 2 reports the main summary statistics of the variables in 1990, 2000 and 2012. An average European city in 2012 had an area of 123 sq km of residential land (4,300 inhabitants per sq km), which was made up of 87 residential land lots (0.8 lots per sq km) and surrounded by 36.5% of undeveloped land. Between 1990 and 2012, residential area and the number of residential lots (per sq km) grew, while residential density and the percentage of undeveloped land decreased. In Appendix Table A.1, I report individual computations for the 60 largest FUAs, i.e. those with populations over one million in 2010.

The high standard deviations (SD) and the minimum and maximum values of all variables show that the FUAs are quite different and heterogeneous. For example, the average residential area grew between 1990 and 2012, but there were FUAs with negative growth (73). Similarly, the average residential density decreased, but 228 cities increased the number of inhabitants per sq km. The average fragmentation and fragmentation per sq km of residential land grew, but there were cities that experienced a reduction in their values (90 and 292, respectively). As mentioned, the surroundings indicators decreased by 0.4 p.p., but there were 182 cities that increased the percentage of undeveloped land surrounding residential land.

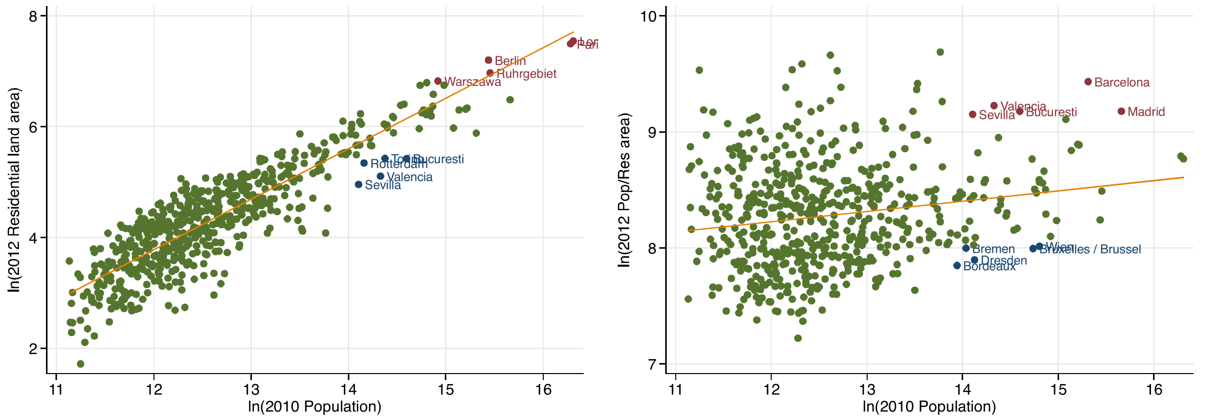
Table 2: Residential land in FUAs: Summary statistics

	Mean	S.D.	Min	Max
Land area				
2012 Residential area (km^2)	122.9	175.8	5.6	1,887
2000 Residential area (km^2)	116.0	169.9	7.0	1,863
1990 Residential area (km^2)	109.9	166.2	5.6	1,867
Population/Land area				
2012 Population/Res area ($Inhab/km^2$)	4,373	2,225	1,369	16,105
2000 Population/Res area ($Inhab/km^2$)	4,410	2,287	1,2	19,913
1990 Population/Res area ($Inhab/km^2$)	4,798	2,584	1,222	20,429
Fragmentation				
2012 Residential fragmentation (*Lots*)	86.5	105.1	2	997
2000 Residential fragmentation (*Lots*)	78.8	97.67	1	980
1990 Residential fragmentation (*Lots*)	74.3	94.6	1	982
Fragmentation/Land area				
2012 Fragmentation/Res area ($Lots/km^2$)	0.82	0.38	0.09	2.0
2000 Fragmentation/Res area ($Lots/km^2$)	0.78	0.37	0.09	2.1
1990 Fragmentation/Res area ($Lots/km^2$)	0.78	0.37	0.06	2.1
Surroundings				
2012 Undeveloped surroundings (%)	36.5	10.7	6.0	62.6
2000 Undeveloped surroundings (%)	36.2	10.4	9.4	60.9
1990 Undeveloped surroundings (%)	36.9	10.3	9.5	61.6

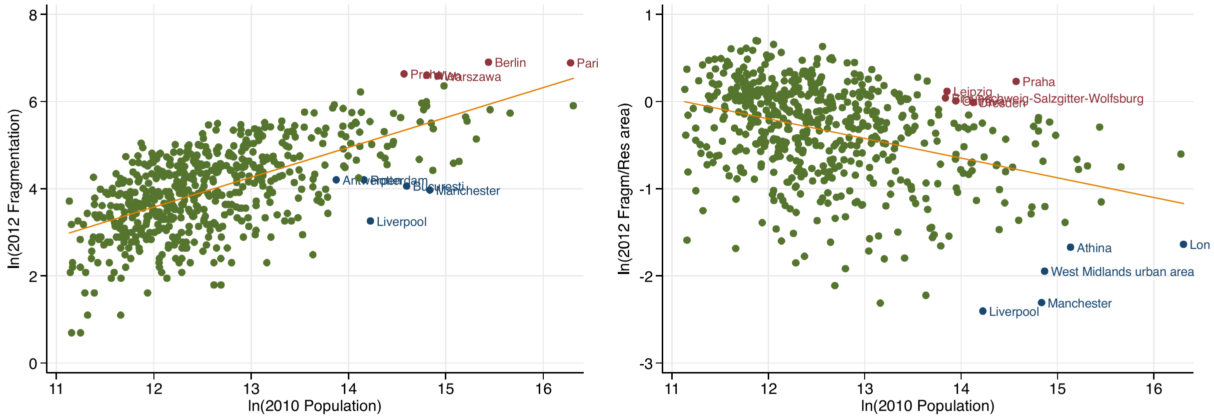
Overall, Tables 1 and 2 show that European cities are, on average, undergoing a process of sprawl that (1) increases their residential size with new, more fragmented and more isolated land developments and (2) reduces their population residential density. However, these tables also show a high degree of heterogeneity between cities.

Figure 3 shows the relationships between each residential land variable and the level of population in 2012. As expected, the larger a city in terms of population, the larger the city's residential land area (Figure 3a). This positive link is also evident between residential density (Figure 3b) and the number of residential lots (Figure 3c). In contrast, the larger the population, the lower the number of lots per sq km (Figure 3d) and the lower the share of undeveloped surroundings (Figure 3e).

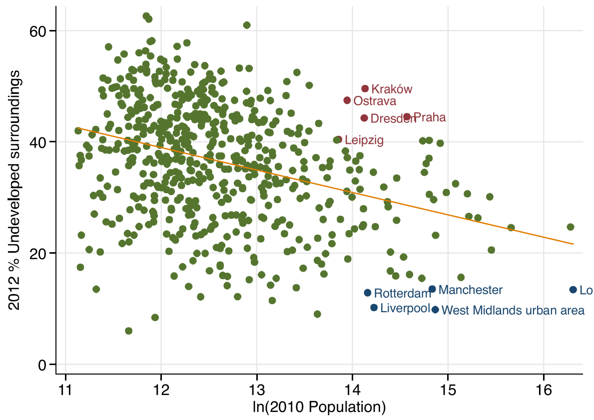
Figure 3: Scatterplots of residential land variables, 2012


(a) Residential area vs. Population

(b) Pop/Res area vs. Population


(c) Fragmentation vs. Population

(d) Fragm/Res area vs. Population


(e) % Undev. surroundings vs. Population

Notes: The red and blue dots refer to the 5 highest and the 5 lowest values of each variable, respectively, among the 60 largest cities in Table A.1; the green dots refer to the other cities.

2.2 *Highways in Europe*

2.2.1 *Modern highways*

The study of highways in Europe is important for, at least, three reasons. First, highways are a short- and long-term priority for the European Union. The goal of the Trans-European Transport Network (TEN-T) program and, in particular, of the Trans-European Road Network (TERN) project is to improve the internal road infrastructure of the EU (see Council Decision 93/629/EEC). The TERN includes highways and high-quality roads, whether existing, new or to be adapted, which play an important role in long-distance traffic, bypass urban centers, provide interconnection with other modes of transport, or link landlocked and peripheral regions to central regions of the Union (see Article 9 of Decision 661/2010/EU).

Second, as Duranton and Turner (2012) highlight for the US case, highways represent large segments of the economy and large amounts of money are devoted to road transportation. According to the ERF 2010 European Road Statistics, 53% of EU structural funds were allocated to roads between 2007 and 2013. Furthermore, the new EU transportation policy was approved in 2014 with a budget of €24 billion up to 2020[7]. Given the magnitude of these investments, it is important that the impact of the EU policy on cities' outcomes be carefully evaluated.

Finally, highway construction in Europe was important during the 20th century and it continues to be so in the 21st century. The first European highways date back to the early twentieth century and were built in Italy (83 km in 1925). Up to 1940, two other countries built highways, Germany (with its *Reichsautobahn* program) and the Netherlands. By 1960, there were around 259 km of highways concentrated in the aforementioned countries, but also a few kilometers had been built in Belgium, Croatia and Poland. Between 1960 and 1980 highways spread unevenly across Europe with almost 28,000 new kilometers of highways. Specifically, some Northwestern countries built national networks of remarkable length and achieved high levels of infrastructural density. Between 1980 and 1990, the EU partially funded highway construction in Mediterranean countries, in particular in Spain, Greece and Portugal, and the network increased with 16,000 km. Similarly, with the latest enlargements of the European Union, the EU Regional policy targeted the new members from Eastern Europe and funded the expansion of the European highway network up to 68,000 km in 2010. The aforementioned TERN project aims to reach 90,000 km of highways and high-quality roads by 2020.

Table 3: Highways and old (rail)roads in Europe and its cities

Km of	Highways							
	1990	2000	2010	1990–2010	200 Roman	15th Trade	1810 Postal	1870 Rail
Europe (29)	43,502	57,763	67,779	24,227 (46%)	103,090	19,615	128,000	81,151
FUAs	22,834	25,627	32,270	9,436 (41%)	7,721	2,051	10,784	13,616

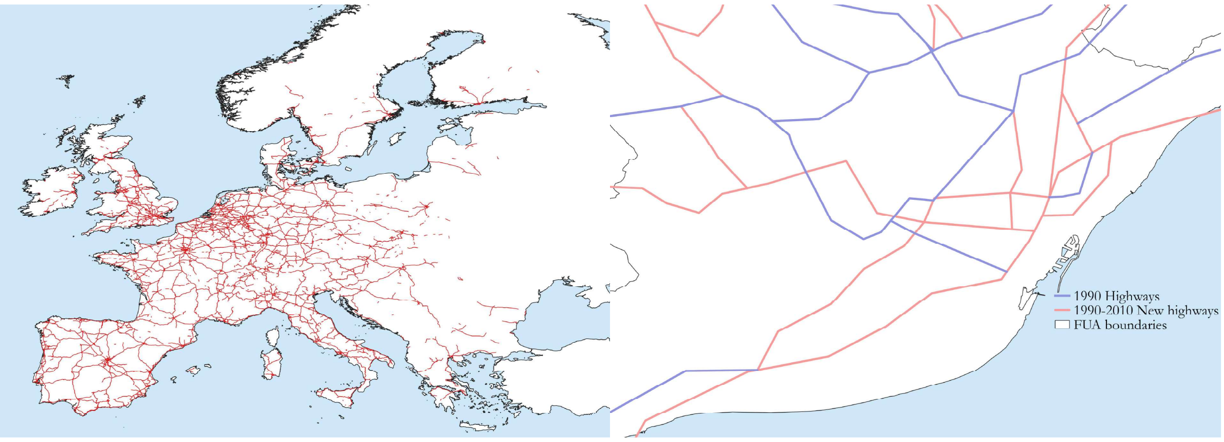
To measure highways, I use the dataset developed by Garcia-López, Pasidis, and Viladecans-Marsal (2015*b*) using Eurostat data at the country level and the RRG GIS Database at the FUA

---

[7] See `http://ec.europa.eu/transport/themes/infrastructure/index_en.htm` for further information.

level. Table 3 reports aggregate values of the length of the highway network in 1990, 2000 and 2010, and their 1990–2010 change and growth rate, for the whole of Europe (first row) and my sample of 579 FUAs (second row). The computations clearly show that half of the European highway network is located in FUAs. Furthermore, between 1990 and 2010 both the European and FUA networks increased by more than 40% (24,000 and 9,000 kilometers, respectively). Figure 4 shows the location of the highway network in Europe in 2010 (Figure 4a) and its evolution in Barcelona between 1990 and 2010 (Figure 4b).

Figure 4: Highways and (old)rail roads



(a) Europe: 2010 highways (b) Barcelona: 1990–2010 highways

(c) Europe: Old (rail)roads (d) Barcelona: Old (rail)roads

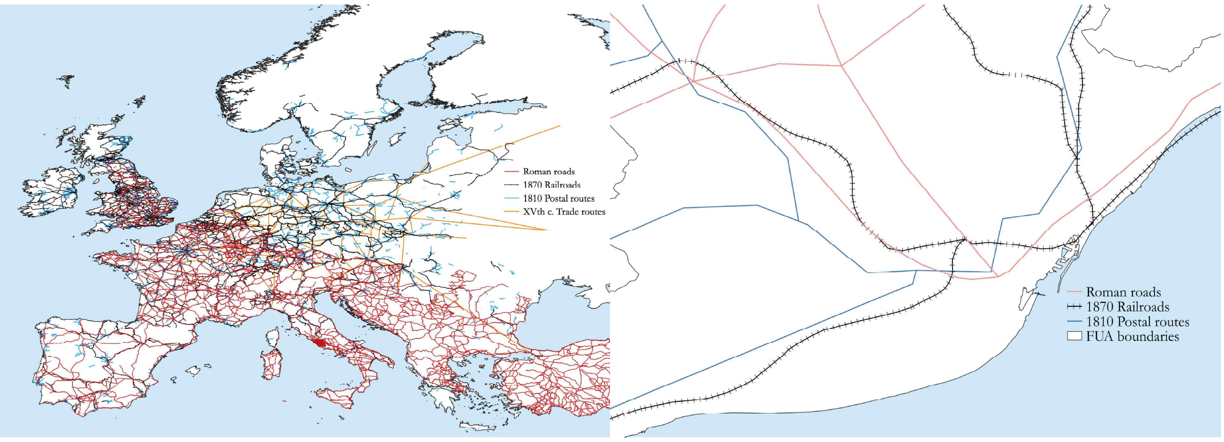


Table 4: Highways and old (rail)roads in FUAs: Summary statistics

	Mean	S.D.	Min	Max
Highways				
2010 Km of highways	55.7	74.8	0	678.6
2000 Km of highways	44.3	66.6	0	562.9
1990 Km of highways	39.4	62.8	0	562.9
Old (rail)roads and instruments				
Km of Roman roads	13.3	52.0	0	1,162
Km of Trade routes	3.5	10.0	0	63.0
Km of 1810 Postal roads	18.6	28.7	0	437.9
Km of 1870 Railroads	23.5	48.0	0	783.6

Table 4 shows the summary statistics (mean, SD, and minimum and maximum values) of the length of the highway network in 1990, 2000 and 2010 (in FUAs). Similar to population and residential variables, the length variables show a high degree of heterogeneity. With an average length of 56 km and a SD of 75 km in 2010, there are cities with 0 km of highways and a maximum of 563 and 679 km in 1990, 2000 and 2010, respectively. Appendix Table A.2 shows the evolution of the different transportation networks in the 60 most populated European cities.

Overall, FUA sample computations in Tables 3 and 4 show that highways (and their construction) are unevenly distributed in Europe and between its cities.

2.2.2 *Ancient 'highways'*

Tables 3 and 4 (and Appendix Table A.2) also report length computations for transportation networks in Europe that were important in the past: The Roman roads, the 15th century Trade routes, the 1810 Postal roads, and the 1870 Railroads. Figure 4 also shows the location of these networks in Europe (Figure 4c) and in Barcelona (Figure 4d). A visual comparison between these old (rail)roads and the modern highway network (Figures 4a and 4b) shows a positive connection between their size and their location. In Section 4, I explore this relationship in detail and test whether instruments built using these past networks predict the 1990–2010 changes in the highway network.

I first consider the Roman road network using the GIS map created by McCormick, Huang, Zambotti, and Lavash (2013). The Romans were the first to build an extensive and sophisticated network of paved and crowned roads. These roads radiated from Rome and connected the different parts of the Empire, from Britain to Syria (O'Flaherty, 1996). Overall , there were more than 100,000 km of main and secondary roads in Europe. At the city level, 7,700 km of Roman roads were built in 285 FUAs.

Based on Ciolek (2005)'s digital map, I compute the length of the main trade routes in the Holy Roman Empire in the 15th century. As its name indicates, the map includes the main routes between Central and Eastern European cities, but also with some other leading European cities in Northwestern and Mediterranean countries. Overall, there were around 20,000 km of routes in Europe.

For reasons of communication and, in particular, military motives, postal roads and postal stations were built in Europe and contributed to the rise of absolute monarchies during the 17th and 18th centuries. Although the first postal roads were relatively primitive, they were improved in the last quarter of the 18th century and permitted the use of wheeled coaches and wagons for carrying letters, goods and people. According to Crew, Kleindofer, and Campbell (2008), postal stations were located every 10 to 15 miles. Across the whole of Europe, there were around 8,000 postal stations in 1799 (Elias, 1982). As a result, I estimated the total length of European postal roads at 128,000 km (=8,000 stations × 10 mi/station × 1.6 km/mi). At the city level, I use a digital vector map that I created from the Map exhibiting the great postal roads, physical and political divisions of Europe by A. Arrowsmith in 1810 and downloaded from the David Rumsey Historical Map Collection (`http://www.davidrumsey.com`). Almost 11,000 km of the 1810 postal roads were built in 487 of the 579 FUAs.

Finally, I consider the 1870 railroad network because, as Duranton and Turner (2012) point out, old railroads are easily converted to roads reducing the construction costs of leveling and grading. In fact, while the network kept expanding, linking up much of Europe between 1870 and 1900, many lines were closed down and some converted to roads between 1900 and 1960 and, in particular, during the great highway expansion between 1960 and 2010 (Garcia-López *et al.*, 2015*b*). To compute the European and FUA lengths, I create a digital vector map based on the online maps built by Historical GIS for European Integration Studies (`http://www.europa.udl.cat/hgise`). Overall, 81,000 km of railroads had been built in Europe by 1870, with 13,000 km of them in 441 FUAs.

3. A brief literature review of urban spatial structure, sprawl and transportation

3.1 *Theory*

Transportation plays a crucial role in the spatial distribution of residences and firms within cities. The classical monocentric city model developed by citetalonso1964, Mills (1967) and Muth (1969) shows that transportation (accessibility) is the main factor determining urban land use (Duranton and Puga, 2015). Transportation is characterized as a non-limited, radial-type infrastructure covering the whole city in the same way and, as such, allowing the same access to the city's unique center or CBD from any point located at the same distance from this CBD. This homogeneous and continuous spatial distribution of transportation infrastructure leads to (1) a continuous (and non-fragmented) development of land for urban uses and (2) an homogeneous reduction in land use intensity (i.e., population density) as population moves away from the CBD. The monocentric city model also predicts that transportation improvements foster both the physical expansion of the city with new and continuous residential land developments and the increase in city population. When the first effect is greater than the second, transportation improvements also foster the (relative) suburbanization of population and may reduce the city's overall density (Duranton and Puga, 2015).

The monocentric city model also predicts that transportation improvements foster both the physical expansion of the city with new and continuous residential land developments and the increase of city population. When the former effect is larger than the latter, transportation improvements also foster the (relative) suburbanization of population and might reduce the overall city density (Duranton and Puga, 2015).

Anas and Moses (1979) and, in particular, Baum-Snow (2007) extended the monocentric model by considering two competing transportation infrastructures: First, the classical transportation infrastructure based on a dense network of radial streets, and second, a high speed transit system (Anas and Moses, 1979)) or a highway network (Baum-Snow, 2007) both based on sparse radial corridors. Depending on the cost of alternative transportation modes, the authors find that population spreads out along the sparse corridors, increasing land rents and densities adjacent to them and decreasing them elsewhere. As Anas and Moses (1979) show through several graphical examples, the total residential land area of the city, its size and its shape depend on the size of its transportation networks.

In the aforementioned studies, residential land area is continuous (and non-fragmented) as the authors assume that highways and railroads can be accessed from any point of the network. In contrast, if we assume that these infrastructures can only be accessed via their access points (highway ramps and railroad stations), population and residential land developments would tend to locate around them (and not along the whole infrastructure). As a result, the number of residential land lots and the percentage of undeveloped surroundings also depend on the size of the transportation networks.

In summary, the theoretical literature on urban land use inspired by the monocentric model shows that transportation influences (1) the size of the city in terms of residential land, (2) the degree of fragmentation in terms of residential land lots, and (3) the degree of undeveloped surroundings.

3.2 *Empirics*

Although several studies have documented the phenomenon of sprawl in terms of land development in the US (Brueckner and Fansler, 1983, Burchfield *et al.*, 2006, McGrath, 2005, Paulsen, 2012), China (Deng *et al.*, 2008), Europe (Oueslati *et al.*, 2015) and even around the world (Angel *et al.*, 2012), the literature on the determinants of sprawl and, in particular, on the effects of transportation remains scarce.

If we focus on the three dimensions of sprawl discussed above, we find most papers study the impact of transportation (costs) on the spatial size of cities in terms of the area of land developed. For a sample of 40 US urbanized areas in 1970, Brueckner and Fansler (1983) do not find any significant effect related to two alternative proxies for commuting costs: the percentage of households owning automobiles and the percentage of commuters using public transit. In contrast, McGrath (2005) uses panel data techniques with a sample of 33 large US cities between 1950 and 1990 and finds a significant negative effect of the CPI for private transportation. Focusing on Chinese counties, Deng *et al.* (2008) find significant positive effects of the density of highways on the built-up area of urban cores in 1988, 1995 and 2000. Similarly, Oueslati *et al.* (2015) find significant positive effects of the regional density of highways in 282 European cities in 1990, 2000 and 2006.

Only Oueslati *et al.* (2015) study the impact of transportation on the degree of land fragmentation measured as the ratio between the number of urban land lots and the total developed area. In contrast to their developed land size findings, they report no significant effect of regional highway density.

Finally, Burchfield *et al.* (2006) and Angel *et al.* (2012) study the impact of transportation on the average percentage of undeveloped land surrounding developed land. Using the density of suburban roads as their transportation variable, Burchfield *et al.* (2006) do not find any significant effect in 275 US metropolitan areas between 1976 and 1992. In contrast, Angel *et al.* (2012) find that greater automobile ownership encourages compact developments and, as a result, reduces the percentage of undeveloped surroundings in 120 cities in the world. They suggest that this result emerges when private transport complements public transit (railroads, buses) and development concentrates around their access points (railroad stations, bus stops).

The above studies do not present any clear evidence of the effect of transportation on the sprawl phenomenon. A possible explanation for this may be the transportation variables they use. Brueckner and Fansler (1983), McGrath (2005) and Angel *et al.* (2012) use 'indirect' proxies for commuting costs (private transportation ownership and price index). In contrast, Burchfield *et al.* (2006), Deng *et al.* (2008) and Oueslati *et al.* (2015) use 'more direct' proxies (density of highways) that allow them to measure the effect of the network size (while including commuting costs). However, by focusing on suburban roads, Burchfield *et al.* (2006) neglect the effect of central roads on population and employment suburbanization and, as a result, on sprawl. citetdengetal2008 and Oueslati *et al.* (2015) use the density of highways computed at the county and regional levels, respectively. Since both measurements exceed the spatial boundaries used in the dependent variables (FUAs and county urban cores, respectively), their results might be biased.

Another reason for these inconclusive results may be attributed to problems of endogeneity (Duranton and Puga, 2015). While a greater number of cars or highways can lead to more (and more fragmented) land development, cities that sprawl for other reasons can also produce an increase in car ownership and the availability of highways. Unfortunately, most of the aforementioned empirical papers assume that their transportation variables are exogenous to land development. While this is true in the case of Burchfield *et al.* (2006), because of the way in which they construct their dependent variable, endogeneity is not addressed in the other five papers and, as result, their estimated coefficients (and their statistical significance) may be biased. As I seek to show in the following section, concerns regarding endogeneity in my transportation explanatory variable can be addressed using IV techniques.

4. The empirical strategy

4.1 *Empirical model*

To study the role played by highways in residential sprawl in 579 FUAs, I empirically answer three questions: Do highways expand cities with new residential land developments? Do they encourage fragmented or compact residential developments? Do they foster more isolated residential developments, with more undeveloped surroundings?

Indexing cities by i and years by t, the general empirical equation is as follows:

$$\begin{aligned} \ln(\text{Sprawl}_{it}) = \beta_0 + \beta_1 \times \ln(\text{Km of highways}_{it}) \\ + \beta_2 \times \text{Socioeconomy}_{it} + \beta_3 \times \text{Geography}_i + \beta_4 \times \text{History}_i + \epsilon_{it} \end{aligned} \tag{1}$$

where Sprawl_{it} refers to the three variables characterizing the different dimensions of sprawl: (a) the residential land area (in sq km), (b) the number of residential land lots, and (c) the percentage of undeveloped land surrounding residential land.

The main explanatory variable is the length of the highway network (in km), indicating the size of the network.

Socioeconomy_{it} is a vector of socioeconomic characteristics including income, proxied by GDP; unemployment rate, proxied by ((active population - employment)/active population); and industrial composition, proxied by the share of employment in manufacturing. Since there are no data available at the FUA level, all three variables are computed using data from the NUTS3 in which the FUA is located.

Geography$_i$ includes controls for physical geography such as total land area (km^2), altitude (m), elevation range (m), and the terrain ruggedness index computed *a la* Riley, DeGloria, and Elliot (1999).

Finally, History$_i$ adds two types of historical controls: First, three dummy variables to consider the cities that were important during three historical periods: The Roman Empire, the Medieval Ages, and the XVI–XIX centuries[8]; and, second, three controls for more recent history (the 20th century): The decennial kilometers of both highways and railroads, and the decennial levels of population between 1960 and 1980.

Summary statistics for all control variables are shown in Appendix Table B.1.

With data describing a panel of cities, I can partition ϵ into permanent (δ_i) and time-varying (η_{it}) components. By so doing, I can estimate Equation (1) using city fixed-effects or alternatively I can estimate the following first-difference equation:

$$\Delta \ln(\text{Sprawl}_{it}) = \beta_1 \times \Delta \ln(\text{Km of highways}_{it}) + \beta_2 \times \Delta\text{Socioeconomy}_{it} + \Delta\eta_{it} \tag{2}$$

where Δ is the first-difference operator.

Since all time invariant factors drop out of the above equation and, as discussed in more detail in the following subsection, I am interested in including some of them, finally I opt to estimate the following two alternative equations:

$$\begin{aligned}\Delta \ln(\text{Sprawl}_{it}) = \beta_1 \times \Delta \ln(\text{Km of highways}_{it}) &+ \beta_2 \times \Delta\text{Socioeconomy}_{it} \\ &+ \beta_3 \times \text{Geography}_i + \beta_4 \times \text{History}_i + \Delta\eta_{it}\end{aligned} \tag{3a}$$

where I include Geography$_i$ and History$_i$ variables, and country fixed-effects.

$$\begin{aligned}\Delta \ln(\text{Sprawl}_{it}) = \beta_1 \times \Delta \ln(\text{Km of highways}_{it}) &+ \beta_2 \times \Delta\text{Socioeconomy}_{it} \\ &+ \delta_i + \Delta\eta_{it}\end{aligned} \tag{3b}$$

which includes city-fixed effects (δ_i).

In some additional specifications, I also include the ten-year lagged value of the sprawl variable analyzed.

4.2 Method

Assuming that the random element of land developments is uncorrelated with transportation, I can estimate Equations (3a) and (3b) by ordinary least squares (OLS). However, as pointed out in the previous sections, highway length is expected to be endogenous to land developments because of reverse causation (e.g., land developments fostering the construction of new highways), measurement error (e.g., the stock of highways mismeasured because some may have just

---

[8] These dummies are based on Garcia-López *et al.* (2018), who study the impact of historical amenities on cities. To do so, the authors geolocated all well-preserved historical buildings in Europe belonging to pan-European architectural styles: The Classical Antiquity, the Romanesque and the Gothic, and the Renaissance and the Baroque.

opened or are about to be opened) and omitted variables (e.g., geography, amenities or economic structure leading to more highways).

To address concerns of endogeneity, I rely on IV estimations (two-stage least squares, TSLS), which use instruments built on the ancient (rail)roads in Europe. Since, by definition, the 'pure' historical instruments are *time-invariant*, I follow Baum-Snow (2007) and Garcia-López (2012) and adopt a 'shift-share' approach *a la* Bartik (1991) using each historical (rail)road as the 'share' component and the evolution of the highway network as the 'shift' component. Specifically, I compute each *time-variant* historical instrument by multiplying its historical length (in km) by the fraction of the highway network kilometrage in each country completed at each decade and excluding each city's own contribution. Since my empirical strategy is based on first-difference regressions, I can now compute the first-difference version of each *time-variant* historical instrument.

As common sense suggests, historical transportation networks may be relevant because modern networks are not built in isolation from them. On the contrary, it is easier and cheaper to build new infrastructures close to old infrastructures. Duranton and Turner (2011, 2012), Garcia-López (2012) or Garcia-López *et al.* (2015*a*), among others, show that the stocks of historical (rail)roads are indeed highly correlated with the stocks of modern transportation networks in the US, Barcelona and Spain, respectively.

I test econometrically the relevance of each *time-variant* historical (rail)road in Appendix C. I run first-stage regressions in which I regress changes in the length of the highway network (km) on changes in the length of each ancient (rail)road (km). Valid instruments should have positive and significant effects on modern highways and high first-stage statistical values. I also run reduced-form regressions in which I separately regress changes in each residential land variable (area, fragmentation and surroundings) on changes in the length of each historical network. As Murray (2006) points out, valid instruments should also have positive and significant effects on the dependent variable of interest. First-stage results in Table C.1 columns 1-2 and 6-7 show that only the Roman roads predict changes in the length of highways (first-stage) and show first-stage statistics that are near or above Stock and Yogo (2005)'s critical values. According to the reduced-form results, the Roman roads also predict residential area growth (columns 3 and 8), fragmentation growth (columns 4 and 9), and the increase in the percentage of undeveloped surroundings (columns 5 and 10).

The next requirement for my *time-variant* historical instrument is that it needs to be exogenous, in particular its 'share' component[9]: The length of Roman roads. In general any historical transportation network may be exogenous because of the length of time that has passed since it was built and the significant changes undergone by society and economy in the intervening years (Duranton and Turner, 2012). In this case, Roman roads were not built to anticipate the current land developments in European cities thousands of years later. They were built to achieve military, administrative, and commercial goals between different parts of the Roman Empire (Garcia-López

---

[9] By construction, the 'shift' element of all potential instruments is exogenous because it allows us to compute the length of the highway network that would have existed in each decade had governments allocated highway construction uniformly across historical (rail)roads within the countries.

et al., 2015*a*)[10]. However, it is true that, first, the geography could have influenced the construction of both Roman roads and modern highways, and, second, Roman roads have undoubtedly shaped the historical development of European cities in other ways (e.g., FUAs with more historical networks tend to be larger than other cities). As a result, following Duranton and Turner (2011, 2012) and Duranton and Puga (2014), the exogeneity of my Roman instrument hinges on having an appropriate set of geographical and historical controls. As discussed in the previous section, Equation (3a) includes Geography and History variables (and country fixed-effects). Another alternative is to add city fixed-effects as in Equation (3b).

In summary, I estimate Equations (3a) and (3b) using dependent variables related to three dimensions of sprawl: Changes in residential land area, fragmentation and undeveloped surroundings in 1990, 2000 and 2012. Since my main explanatory variable (changes in the stock of highways) is endogenous, I rely on IV estimations using changes in the length of (*time-variant*) Roman roads as my instrument. According to its first-stage and reduced-form results, and the above comments, I believe that the instrument is relevant and, conditional on controls, exogenous.

5. Main results

5.1 *Do highway improvements expand cities with new land developments?*

To study the impact of highways on residential sprawl, I first investigate whether they foster new land developments increasing the residential land area (size) of the city as theory suggests.

Column 1 of Table 5 reports OLS results when I estimate Equation (3a), which includes historical and geographical variables as controls. Column 2 adds the ten-year lagged value of residential land area. In columns 3 and 4, I estimate Equation (3b) with city fixed-effects. The estimated coefficient for the changes in highways is positive and significantly different from zero in all specifications and shows that a 10% increase in the stock of highways (km) expands residential land area by 0.2%. TSLS counterparts of column 1 and 3 are in columns 5 and 6 and show a higher impact of highways on residential city size (1.1%).

Following Duranton and Turner (2012), I consider the effect of highways on population by estimating my preferred specification in column 6 using the changes in population as dependent variable. OLS and TSLS results in columns 7 and 8 confirm that highways also foster population growth in European cities. Specifically, after addressing concerns of endogeneity, I find that a 10% increase in the highway network causes a 0.5% increase in city population (column 7).

Finally, in columns 9 and 10, I study the impact of highways on residential density. In other words, I simultaneously consider the highway effects on both population and residential land area. Results show that, as theory suggests, effects on residential land are greater than on population and, as a result, residential density decreases with transportation improvements (Duranton

---

[10] Similarly, the 15th century trade routes were built for commercial purposes (Garcia-López *et al.*, 2018). The 18th and 19th centuries postal roads were designed as a central government tool for nation building (military and communication purposes) (Garcia-López, 2012). Finally, similar to the US, most of the 1870 railroad network was built for profit by private companies at the beginning of the second industrial revolution, when cities' economy and industrial specialization were quite different than today (Duranton and Turner, 2012, Garcia-López, 2012).

and Puga, 2015). Specifically, a 10% increase in the length of highways reduces residential density by 0.6% (column 10).

Table 5 also reports first-stage statistics for the aforementioned TSLS regressions (which use changes in the *time-variant* Roman roads as instrument), and all of them are near or above Stock and Yogo (2005)'s critical values.

Table 5: The effect of highways on residential land area in European cities

Dependent variable:	Δln(Km^2 of residential land area)						Δln(Pop)		Δln(Pop/Area)	
Method:	OLS [1]	OLS [2]	OLS [3]	OLS [4]	TSLS [5]	TSLS [6]	OLS [7]	TSLS [8]	OLS [9]	TSLS [10]
Δln(Km of highways)	0.0129^{b}	0.0121^{b}	0.0170^{a}	0.0171^{a}	0.1609^{a}	0.1103^{a}	0.0040^{b}	0.0501^{a}	-0.0130^{b}	-0.0601^{b}
	(0.0056)	(0.0055)	(0.0061)	(0.0062)	(0.0512)	(0.0303)	(0.0019)	(0.0148)	(0.0060)	(0.0241)
Δln(Socioeconomy)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Geography	Y	Y	N	N	Y	N	N	N	N	N
History	Y	Y	N	N	Y	N	N	N	N	N
Initial ln(Resid. area)	N	Y	N	Y	N	N	N	N	N	N
Country FE	Y	Y	N	N	Y	N	N	N	N	N
FUA FE	N	N	Y	Y	N	Y	Y	Y	Y	Y
Adjusted R^2	0.44	0.46	0.35	0.35			0.11		0.30	
First-stage F-statistic					13.23	17.39		17.39		17.39
Instrument:										
Δln(Km Roman roads)					✓	✓		✓		✓

Notes: All regressions include a constant and decade effects. 1,158 observations (579 FUAs × 2 decades) in each regression. Robust standard errors clustered by FUA are in parentheses. ^{a}, ^{b}, and ^{c} indicates significant at 1, 5, and 10 percent level, respectively.

The difference between the TSLS coefficient in column 6 (0.110) and its OLS counterpart (0.012) suggests that the construction of highways in Europe is endogenous. Why? It may be the result of a classical measurement error, but, as I find similar OLS-LIML differences when I use a different measure of highways based on ramps, I can rule out this possibility.

It may also be the result of a negative correlation between changes in highways and the error term because of missing variables or reverse causation. Despite controlling for geography, socioeconomics, history and country fixed-effects, and, in my preferred specification, by city fixed-effects, the possibility remains that the OLS-LIML differences might be explained by missing variables such as consumption amenities or local planning regulations, which could be associated with higher residential area growth and with fewer changes in highways.

Alternatively, it could be that, conditional on controls, cities subject to less expansion on average experience positive shocks to their stock of highways. Although not reported for reasons of space, I test this hypothesis by regressing changes in highways on changes in residential area. Similar to Duranton and Turner (2012), I find a negative estimated coefficient for the residential area variable, confirming that FUAs with lower area growth are assigned more highways.

5.2 *Do highways encourage fragmented land developments?*

After establishing that highways foster new residential land developments, I now turn my attention to study their impact on the degree of residential fragmentation. To do so, I estimate Equations (3a) and (3b) using changes in the number of residential land lots as dependent variable.

Table 6 presents results for the degree of fragmentation, its format being similar to the previous table: OLS results (columns 1 to 4) and TSLS results (columns 5 and 6) when controlling for geography and history (columns 1, 2 and 5) and when including city fixed-effects (columns 3, 4 and 6). As for residential area, the OLS results indicate that highways have a positive and significant effect on the number of residential land lots. After instrumenting the highway variable with the *time-variant* Roman road instrument, TSLS results confirm that highways foster residential fragmentation. Specifically, the estimated coefficient of my preferred specification in column 6 shows that a 10% increase in the stock of highways (km) causes a 2.0% growth in the number of residential lots.

Table 6: The effect of highways on residential land fragmentation in European cities

Dependent variable:	Δln(Number of residential land lots)						Δln(Lots/Area)	
Method:	OLS [1]	OLS [2]	OLS [3]	OLS [4]	TSLS [5]	TSLS [6]	OLS [7]	TSLS [8]
Δln(Km of highways)	0.0344^a	0.0329^a	0.0425^a	0.0424^a	0.2013^b	0.1988^a	0.0255^b	0.0886^b
	(0.0115)	(0.0109)	(0.0121)	(0.0121)	(0.0880)	(0.0626)	(0.0101)	(0.0431)
Δln(Socioeconomy)	Y	Y	Y	Y	Y	Y	Y	Y
Geography	Y	Y	N	N	Y	N	N	N
History	Y	Y	N	N	Y	N	N	N
Initial ln(Resid. lots)	N	Y	N	Y	N	N	N	N
Country FE	Y	Y	N	N	Y	N	N	N
FUA FE	N	N	Y	Y	N	Y	Y	Y
Adjusted R^2	0.38	0.40	0.29	0.29			0.09	
First-stage F-statistic					13.23	17.39		17.39
Instrument:								
Δln(Km Roman roads)					✓	✓		✓

Notes: All regressions include a constant and decade effects. 1,158 observations (579 FUAs × 2 decades) in each regression. Robust standard errors clustered by FUA are in parentheses. ^{a}, ^{b}, and ^{c} indicates significant at 1, 5, and 10 percent level, respectively.

A comparison of the estimated coefficients for the number of residential lots (0.199) and for residential land area in Table 5 (0.110) shows the former to be higher than the latter. The positive difference between the two coefficients indicates that new land developments are more fragmented than earlier ones. By using changes in the ratio of lots per residential area as dependent variable, in column 8, I confirm the significance of this result: a 10% increase in the stock of highways causes a 0.09% increase in the number of lots per sq km of residential land.

Once again, the difference between OLS and TSLS estimates can be related to a missing variable such as local land use regulations. This difference may also be due to the construction of highways in less sprawling cities in terms of their fragmentation (as shown by a negative estimated coefficient when regressing changes in highways on changes in residential lots).

5.3 *Do highways foster isolated land developments?*

Now that we know that highways encourage new and more fragmented residential developments, I investigate whether highways influence their degree of undeveloped surroundings (isolation): Do highways promote residential lots that are much more isolated from other artificial lands? To answer this question, I estimate Equations (3a) and (3b) using changes in the percentage of undeveloped land surrounding residential land as dependent variable.

Table 7 reports OLS results when controlling by geography and history (columns 1 and 2) and when adding city fixed-effects (columns 3 and 4). In both cases, the estimated coefficient of highways is positive and significant and shows that a 10% increase in the stock of highways (km) increases the percentage of undeveloped land by 6.0 p.p. TSLS results in columns 5 and 6 show that, after addressing endogeneity concerns, the effect of highways is higher: A 10% increase in the highway network causes an increase of 24.6 p.p. in the share of undeveloped land in the sq km surrounding each residential development.

Table 7: The effect of highways on residential land isolation in European cities

Dependent variable:	Δ(% of undeveloped land surrounding residential)					
Method:	OLS [1]	OLS [2]	OLS [3]	OLS [4]	TSLS [5]	TSLS [6]
Δln(Km of highways)	0.4482^a (0.1521)	0.3857^a (0.1410)	0.5705^a (0.1582)	0.4416^a (0.1499)	1.7929^a (1.1448)	2.4556^a (0.8650)
Δln(Socioeconomy)	Y	Y	Y	Y	Y	Y
Geography	Y	Y	N	N	Y	N
History	Y	Y	N	N	Y	N
Initial ln(% undev. land)	N	Y	N	Y	N	N
Country FE	Y	Y	N	N	Y	N
FUA FE	N	N	Y	Y	N	Y
Adjusted R^2	0.24	0.30	0.13	0.17		
First-stage F-statistic					13.23	17.39
Instrument:						
Δln(Km Roman roads)					✓	✓

Notes: All regressions include a constant and decade effects. 1,158 observations (579 FUAs × 2 decades) in each regression. Robust standard errors clustered by FUA are in parentheses. ^{a}, ^{b}, and ^{c} indicates significant at 1, 5, and 10 percent level, respectively.

As previously commented, the difference between my preferred OLS estimated coefficient (0.570) and its TSLS counterpart (significant 2.456) might be explained by a missing variable (local land use regulations) and/or by a positive shock on highway construction on more compact cities (as shown by a negative estimated coefficient when separately regressing changes in highway length on changes in the share of undeveloped surroundings).

Overall, the results in Sections 5.1, 5.2 and 5.3 confirm that highways expand cities with more residential land developments (area) and reduce their overall residential density. Because of highways, these new land developments are also more discontinuous and fragmented than earlier ones. Furthermore, they are surrounded by more undeveloped land and, as a result, are more isolated from other artificial lands.

6. Heterogeneity and robustness

6.1 Heterogeneity and additional results

As shown in Section 2 when describing sprawl and highways in Europe, my sample of cities (and countries) is large and diverse. Although the regressions in Section 5 control for unobserved differences in the development of cities (by including city fixed-effects in Equation (3b)) and countries (by including country fixed-effects in Equation (3a)), some differences may still emerge between groups of cities/countries with common features. In this section, I analyze changes in the results when considering different groups of cities based on (1) cultural and institutional characteristics, (2) the importance of highway investments, and (3) the existence of EU funding. To do so, I run new regressions based on interactions between the highway variable and dummies for the different groups of cities. These interacted endogenous variables are instrumented with the interactions of the time-variant Roman instrument and the dummies.

In Panel A of Table 8, I study whether the effect is heterogeneous between cities located in Mediterranean (171 FUAs), Northwestern (252 FUAs), and Eastern European countries (156 FUAs). Results show that the average estimated coefficients for residential land area and fragmentation (columns 6 of Tables 5 and 6) are related to Mediterranean cities (columns 1 and 2). However , while these estimated coefficients are non-significant for Northwestern cities, they are positive and higher than the average values for cities in Eastern Europe. Results for the share of undeveloped surroundings (column 3) only show significant effects in Eastern European cities and they are higher than average (column 6 of Table 7). Overall, this heterogeneous analysis shows that highway improvements only foster a more sprawled pattern of residential land developments in Mediterranean and, more particularly, in Eastern European cities.

I also explore whether there are differences according to the amount of highway investment cities receive. Specifically, in Panel B of Table 8, I compare cities that expanded their highway network above/below the average (30% between 1990 and 2010). Results in columns 4, 5 and 6 show slightly higher effects for cities expanding their network above the average. However, neither the difference with the *below* effects nor the difference with the average effects (columns 6 of Tables 5, 6 and 7) are significant. Although not reported for reasons of space, similar results are obtained when grouping cities according to the degree of development of the highway network in 1990.

Since EU structural funds are critical for highway construction in Europe, I investigate the existence of differences between cities with high/low EU funding. In particular, in Panel C of Table 8 I group cities located in regions that received funds from the Objective 1 Program (Obj. 1) and I compare them with the others (No Obj.1). Interestingly, columns 7, 8 and 9 report highway effects that are lower for Obj. 1 cities. However, these effects are not statistically different from those obtained for the No Obj. 1 cities. In fact, the estimated coefficients for each group of cities are not statistically different from the average values in columns 6 of 5, 6 and 7.

Table 8: Highways and sprawl in European cities: Heterogeneity of the results

	Panel A: Regions				Panel B: Improvements				Panel C: EU Funding		
	Δln(A)	Δln(F)	Δ%Und		Δln(A)	Δln(F)	Δ%Und		Δln(A)	Δln(F)	Δ%Und
	TSLS [1]	TSLS [2]	TSLS [3]		TSLS [4]	TSLS [5]	TSLS [6]		TSLS [7]	TSLS [8]	TSLS [9]
Δln(hwy)				Δln(hwy)				Δln(hwy)			
×Medit	0.0925^a	0.1632^a	0.9560	×Above	0.1344^a	0.2547^a	3.1055^b	×Obj.1	0.0661^b	0.1683^a	1.3846^c
	(0.0336)	(0.0610)	(0.8038)		(0.0460)	(0.0985)	(1.2918)		(0.0264)	(0.0532)	(0.7315)
×Northw	-0.0172	-0.0599	0.9570	×Below	0.0856^b	0.1419^b	1.7924^b	×No Obj.1	0.1727^c	0.2739^c	3.2114
	(0.0851)	(0.1172)	(1.5025)		(0.0342)	(0.0685)	(0.8639)		(0.1033)	(0.1659)	(2.2616)
×Eastern	0.6303^a	0.8443^a	10.8831^a								
	(0.2104)	(0.2862)	(3.5036)								
F-S F-stat	9.05	9.05	9.05	F-S F-stat	6.39	6.39	6.39	F-S F-stat	6.08	6.08	6.08
Instruments:	Interactions between Δln(Km of Roman roads) and the different heterogenity groups										

Notes: All regressions include a constant, decade effects, socioeconomic variables and FUA effects. 1,158 observations (579 FUAs × 2 decades) in each regression. Robust standard errors clustered by FUA are in parentheses. ^{a}, ^{b}, and ^{c} indicates significant at 1, 5, and 10 percent level, respectively.

6.2 *Robustness checks*

In this section I check the robustness of the average results in Section 5 (columns 6 of Tables 5, 6 and 7) by (1) considering other types of transportation networks, (2) estimating alternative empirical specifications, and (3) using an alternative empirical strategy to estimate the effect of highway improvements on sprawl.

In Table 9 , I gradually include changes in the length of railroads, other main roads and secondary roads as explanatory variables. Since I have three additional endogenous variables, I use three additional historical instruments that I build using the 'shift-share' approach explained in Section 4[11]. Results show a slightly lower effect of highways on residential area (0.08), the number of lots (0.17) and the share of undeveloped surrounding (1.92). However, these new coefficients are not statistically different from their average counterparts.

As for other types of infrastructure, only railroad improvements affect sprawl by increasing residential land area, the number of lots and the percentage of undeveloped surroundings. Despite the higher estimated coefficients for railroads, they are not statistically different from those for the highways. Interestingly, while other types of roads only positively affect the number of residential lots, secondary roads do not show any significant effect on any sprawl dimension. As a result, it seems that only highways and railroads foster more sprawled developments.

---

[11] 11The length of each historical (rail)road is the 'share' component and the evolution of the country transportation network (excluding each city's own contribution) is the 'shift' component.

Table 9: Robustness to other transportation networks

	Δln(Area)			Δln(Number of lots)			Δ%Undeveloped		
	TSLS [1]	TSLS [2]	TSLS [3]	TSLS [4]	TSLS [5]	TSLS [6]	TSLS [7]	TSLS [8]	TSLS [9]
Δln(Km of highways)	0.0764^b	0.0774^b	0.0760^b	0.1698^b	0.1724^b	0.1682^b	1.9440^b	1.9744^b	1.9227^c
	(0.0337)	(0.0340)	(0.0371)	(0.0697)	(0.0691)	(0.0727)	(0.9591)	(0.9640)	(1.0183)
Δln(Km of railroads)	0.1843^a	0.1988^a	0.2003^a	0.1589^c	0.1959^c	0.2005^c	2.8886^b	3.3258^b	3.3824^b
	(0.0635)	(0.0735)	(0.0744)	(0.0892)	(0.1018)	(0.1041)	(1.1660)	(1.3337)	(1.3533)
Δln(Km of other main roads)		0.0081	0.0081		0.0208^c	0.0209^c		0.2455	0.2464
		(0.0093)	(0.0092)		(0.0115)	(0.0115)		(0.1710)	(0.1713)
Δln(Km of secondary roads)			-0.0044			-0.0128			-0.1584
			(0.0306)			(0.0346)			(0.4916)
First-stage F-statistic	6.32	3.51	2.67	6.32	3.51	2.67	6.32	3.51	2.67
First-stage F-stat highways	13.02	13.36	12.38	13.02	13.36	12.38	13.02	13.36	12.38
First-stage F-stat railroads	13.68	11.98	11.69	13.68	11.98	11.69	13.68	11.98	11.69
First-stage F-stat other main		32.10	41.23		32.10	41.23		32.10	41.23
First-stage F-stat secondary			27.61			27.61			27.61
Instruments:									
Δln(Km Roman roads)	✓	✓	✓	✓	✓	✓	✓	✓	✓
Δln(Km 1870 Rails)	✓	✓	✓	✓	✓	✓	✓	✓	✓
Δln(Km 1810 Postal main)		✓	✓		✓	✓		✓	✓
Δln(Km 1810 Postal second)			✓			✓			✓

Notes: All regressions include a constant, decade effects, socioeconomic variables and FUA effects. 1,158 observations (579 FUAs × 2 decades) in each regression. Robust standard errors clustered by FUA are in parentheses. ^{a}, ^{b}, and ^{c} indicates significant at 1, 5, and 10 percent level, respectively.

Table 10: Robustness to alternative empirical specifications

	ln(Area)	ln(Fragm)	%Undev
	TSLS [1]	TSLS [2]	TSLS [3]
Level on level pooled regression with time-invariant instrument	0.1433^c	0.1856^c	0.2614^c
Level on level pooled regression with time-variant instrument	0.1377^c	0.1771^b	0.0778^c
Level on level FUA FE regression with time-variant instrument	0.0706^a	0.1268^b	0.2169^c
Change on initial level *a la* Duranton and Turner (2012)	0.0196^b	0.0291^c	0.1174^c

In Table 10, I check the robustness of my preferred results by estimating alternative empirical specifications. First, I estimate Equation (1), that is, a 'level on level' pooled regression with geographical and historical controls and country fixed-effects. Alternatively, I use two related instruments: The original and *time-invariant* length of Roman roads (first row) and my computed *time-variant* length of Roman roads (second row). In both cases, the estimated coefficients show that highways cause sprawl.

Second, after partitioning the error term into permanent and time-varying components, I can estimate Equation (1) using city fixed-effects and instrumenting with the *time-variant* length of

Roman roads. Results (third row of Table 10) confirm the effect of highways on sprawl.

Finally, I estimate a 'changes on level' specification that only considers the effect of 1990 conditions on 1990–2012 changes in residential sprawl (fourth row). Although the estimated coefficients are smaller, they point in the same direction: Highway improvements lead to more sprawled cities.

Table 11: Robustness to alternative empirical strategy: Inconsequential units approach

	Δln(Area)		Δln(Fragm)		Δ%Undev	
	OLS [1]	TSLS [2]	OLS [3]	TSLS [4]	OLS [5]	TSLS [6]
Excluding 264 FUAs with 1990 pop $\geq$ 250.000 inhab.	0.0124^{c}	0.1178^{a}	0.0457^{a}	0.2103^{b}	0.6725^{a}	2.4908^{b}
Excluding 249 FUAs with 1960 pop $\geq$ 250.000 inhab.	0.0120^{c}	0.1105^{a}	0.0441^{a}	0.1806^{b}	0.6157^{a}	2.0746^{b}
Excluding 205 FUAs with 1850 pop $\geq$ 50.000 inhab.	0.0114^{c}	0.1116^{a}	0.0437^{a}	0.2514^{a}	0.6104^{a}	2.7973^{b}
Excluding 230 FUAs that were major towns in 1450	0.0194^{b}	0.0975^{a}	0.0537^{a}	0.1793^{b}	0.6805^{a}	2.2677^{b}
Excluding 116 FUAs that were major towns in 814	0.0165^{a}	0.1302^{a}	0.0386^{a}	0.2479^{a}	0.5036^{a}	3.2324^{b}
Excluding 225 FUAs that were Roman settlements	0.0226^{b}	0.1589^{a}	0.0555^{a}	0.2102^{b}	0.7765^{a}	2.5992^{c}

The empirical literature analyzing the impact of transportation infrastructures has also dealt with endogeneity issues using the so-called 'inconsequential units approach (Redding and Turner, 2015). Since transportation networks are planned to connect large cities/regions, the basic idea of this approach proposed by Chandra and Thompson (2000) is to focus the analysis on those cities/regions that accidentally (and exogenously) lie on the route between two larger cities/regions. To implement this approach, I follow Möller and Zierer (2018) and, from my original sample of 579 cities, I exclude some of them based on their population size and historical importance. With the remaining cities, I repeat the analysis by estimating Equation (3b) (with city fixed-effects) by OLS and TSLS.

Table 11 reports results when applying the inconsequential units strategy. First, I exclude those FUAs with more than 250,000 inhabitants in 1990 (first row) and in 1960 (second row). Using population data from Bairoch (1988), in row three I exclude cities with more than 50,000 inhabitants in 1850. Using information from the Digital Atlas of Roman and Medieval Civilizations (DARMC, `http://darmc.harvard.edu`), I exclude cities that were major towns in 1450 (fourth row) and in 814 (fifth row). Finally, since my instrument is based on Roman roads, I exclude those cities that were Roman settlements according to the DARMC (sixth row). In all six subsamples considered OLS and TSLS results confirm the positive and significant impact of highways on all three sprawl dimensions (area, fragmentation and isolation). Furthermore, these estimated coefficients are not statistically different from their counterparts in columns 3 and 6 of Tables 5, 6 and 7. As a result, I can confirm the robustness of my preferred IV results.

7. Conclusions

In this paper, I investigate the effect of highways on sprawl in Europe between 1990 and 2012. My results confirm the causal effect and, in particular, show that a 10% increase in the stock of highways (km) causes a 1.1% growth in the residential land area, a 0.6% reduction in the overall residential density, a 2% growth in the number of residential lots, and a 25 p.p. rise in the share of undeveloped land surrounding residential land.

Since highways on average grew by around 30% (in length) and European cities expanded their residential size by 12%, decreased their residential density by 3% and increased their number of land lots by 17%, highways are responsible for at least a quarter of the growth in the residential land area ((30%×0.110)/12%), two thirds of the reduction in residential density ((30%×-0.060)/3%), and a third of the growth in the number of residential lots ((30%×0.199)/17%).

With these computations in one hand and the concerns of the future compactness of Europe's cities in the other, the policy implications of my results are clear. If policy makers want to control the sprawl phenomenon, they need to increase the cost of using highways by, for example, introducing road pricing mechanisms targeting specific types of users (e.g., cars, motorcycles) and/or by setting higher fuel taxes. Furthermore, highway construction can also be used to control sprawl. Since suburbanizing population relocates near new highway ramps (Garcia-López, 2012, Garcia-López *et al.*, 2015*a*), their number and their location would allow policy makers to influence the spatial pattern of new land developments.

A qualifier is important here. The above computations also show that there are other factors that might be responsible for sprawl (the 75% of residential area growth, a thrid of residential density reduction, and two thirds of the growth in the number of residential lots). Differences in land and housing prices between central cities and their suburban areas (OECD, 2018), local planning policies and local taxes (Burchfield *et al.*, 2006), and the evolution of the mortgage market (Hilbers, Banerji, Shi, and Hoffmaister, 2008) are some determinants suggested by the literature. Unfortunately, none of these variables are available at the local or city level at this moment. I leave them for future research.

The contributions made by this paper are relevant. First, it is the first paper to undertake a comprehensive analysis of the effects of highways on sprawl. To do so, I have addressed endogeneity concerns using IV techniques that include instruments built on ancient (rail)roads.

Second, the paper uses three measures to capture one or more dimensions of the sprawl phenomenon: (1) Land area and population density, which are simple and well-known measures of size; (2) Number of land lots and lots per area, which are more explicit measures of disconnected land use; (3) Share of surrounding undeveloped land, which measures isolation.

Finally, the paper furthers our understanding of how transportation and, in particular, highways shape cities. Besides their impact on urban growth, suburbanization-decentralization processes, and zoning policies, highways also affect land developments.

References

Ahlfeldt, Gabriel and Elisabetta Pietrostefani. 2017. The compact city in empirical research: A quantitative literature review. Serc discussion papers, Spatial Economics Research Centre, LSE.

Ahlfeldt, Gabriel and Elisabetta Pietrostefani. 2019. The economic effects of density: A synthesis. *Journal of Urban Economics* 111: 93–107.

Anas, Alex and Leon N. Moses. 1979. Mode choice, transport structure and urban land use. *Journal of Urban Economics* 6(2): 228–246.

Angel, Shlomo, Jason Parent, and Daniel L Civco. 2012. The fragmentation of urban landscapes: global evidence of a key attribute of the spatial structure of cities, 1990–2000. *Environment and Urbanization* 24(1): 249–283.

Arribas-Bel, Daniel, Peter Nijkamp, and Henk J. Scholten. 2011. Multidimensional urban sprawl in Europe: A self-organizing map approach. *Computers, Environment and Urban Systems* 35(4): 263–275.

Bairoch, Paul. 1988. *Cities and Economic Development: From the Dawn of History to the Present*. University of Chicago Press.

Bartik, Timothy J. 1991. *Who Benefits from State and Local Economic Development Policies?* W.E. Upjohn Institute for Employment Research.

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

Baum-Snow, Nathaniel, Loren Brandt, J. Vernon Henderson, Matthew A. Turner, and Qinghua Zhang. 2017. Roads, railroads and decentralization of chinese cities. *Review of Economics and Statistics* 99(3): 435–448.

Brueckner, Jan K. and David A. Fansler. 1983. The economics of urban sprawl: Theory and evidence on the spatial sizes of cities. *Review of Economics and Statistics* 65(3): 479–482.

Burchfield, Marcy, Henry G. Overman, Diego Puga, and Matthew A. Turner. 2006. Causes of sprawl: A portrait from space. *The Quarterly Journal of Economics* 121(2): 587–633.

Chandra, Amitabh and Eric Thompson. 2000. Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system. *Regional Science and Urban Economics* 30(4): 457–490.

Ciolek, T. Matthew. 2005. Georeferenced data set (Series 1 - Routes): Main trade routes in the Holy Roman Empire and nearby countries, c. 1500 ce. OWTRAD Dromographic Digital Data Archives (ODDDA). Old World Trade Routes (OWTRAD) Project. Canberra: Asia Pacific Research Online. `www.ciolek.com/OWTRAD/DATA/tmcZCEm1500.html`.

Couch, Chris, Lila Leontidou, and Gerhard Petschel-Held. 2007. *Urban Sprawl in Europe: Landscapes. Land-Use Change and Policy*. Blackwell Publishing Ltd.

Crew, Michael A., Paul R. Kleindofer, and James I. Campbell (eds.) . 2008. *Handbook of Worldwide Postal Reform*. Edward Elgar.

Deng, Xhiangzheng, Jikun Huang, Scott Roxell, and Emi Uchida. 2008. Growth, population and industrialization, and urban land expansion in China. *Journal of Urban Economics* 63: 96–115.

Duranton, Gilles and Diego Puga. 2014. The growth of cities. In *Handbook of Economic Growth*, volume 2, 1 edition, chapter 05. Elsevier, 781–853.

Duranton, Gilles and Diego Puga. 2015. Urban land use. In Gilles Duranton, J. Vernon Henderson, and William C. Strange (eds.) *Handbook of Regional and Urban Economics Volume 5*. Amsterdam: North-Holland, 467–560.

Duranton, Gilles and Matthew A. Turner. 2011. The fundamental law of road congestion: Evidence from US cities. *American Economic Review* 101(6): 2616–2652.

Duranton, Gilles and Matthew A. Turner. 2012. Urban growth and transportation. *Review of Economic Studies* 79(4): 1407–1440.

Elias, Werner. 1982. Maps and road books of Europe's mail coach era 1780-1850. *The Map Collector* 20: 24.

European Environment Agency. 2006. Urban sprawl in Europe: The ignored challenge. Technical report no 10, Luxembourg, Office for Official Publications of the European Communities.

European Environment Agency. 2010. Land in Europe: Prices, taxes and use pattens. Technical report no 4, Luxembourg, Office for Official Publications of the European Communities.

Garcia-López, Miquel-Àngel. 2012. Urban spatial structure, suburbanization and transportation in Barcelona. *Journal of Urban Economics* 72(2–3): 176–190.

Garcia-López, Miquel-Àngel, Camille Hémet, and Elisabet Viladecans-Marsal. 2017*a*. How does transportation shape intrametropolitan growth? An answer from the Regional Express Rail. *Journal of Regional Science* 57(5): 758–780. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/jors.12338`.

Garcia-López, Miquel-Àngel, Camille Hémet, and Elisabet Viladecans-Marsal. 2017*b*. Next train to the polycentric city: The effect of railroads on subcenter formation. *Regional Science and Urban Economics* 67(C): 50–63.

Garcia-López, Miquel-Àngel, Adelheid Holl, and Elisabet Viladecans-Marsal. 2015*a*. Suburbanization and highways in Spain when the Romans and the Bourbons still shape its cities. *Journal of Urban Economics* 85(C): 52–67.

Garcia-López, Miquel-Àngel, Ilias Pasidis, and Elisabet Viladecans-Marsal. 2015*b*. Express deliver to the suburbs? Transportation effects in heterogeneous European cities. Working Paper 2015/22, Institut d'Economia de Barcelona.

Garcia-López, Miquel-Àngel, Ilias Pasidis, and Elisabet Viladecans-Marsal. 2018. Amphitheaters, cathedrals and operas: The role of historic amenities on suburbanization. Manuscript.

Garcia-López, Miquel-Àngel, Albert Solé-Ollé, and Elisabet Viladecans-Marsal. 2015*c*. Does zoning follow highways? *Regional Science and Urban Economics* 53(C): 148–155.

Hilbers, Paul Louis Ceriel, Angana Banerji, Haiyan Shi, and Willy A Hoffmaister. 2008. House Price Developments in Europe; A Comparison. IMF Working Papers 08/211, International Monetary Fund. URL `https://ideas.repec.org/p/imf/imfwpa/08-211.html`.

McCormick, Michael, Guoping Huang, Giovanni Zambotti, and Jessica Lavash. 2013. Roman road network (version 2008). DARMC Scholarly Data Series, Data Contribution Series 2013-5. DARMC, Center for Geographic Analysis, Harvard University, Cambridge MA 02138. `http://darmc.harvard.edu`.

McGrath, Daniel T. 2005. More evidence on the spatial scale of cities. *Journal of Urban Economics* 58(1): 1–10.

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

Möller, Joachim and M. Zierer. 2018. Autobahns and jobs: A regional study using historical instrumental variables. *Journal of Urban Economics* 103(C): 18–33.

Murray, Michael P. 2006. Avoiding invalid instruments and coping with weak instruments. *Journal of Economic Perspectives* 20(4): 111–132.

Muth, Richard F. 1969. *Cities and Housing: The Spatial Pattern of Urban Residential Land Use*. Chicago, IL: University of Chicago Press.

Nickerson, Cynthia, Robert Ebel, Allison Brochers, and Fernando Carriazo. 2011. Major uses of land in the United States, 2007. Economic information bulletin EIB-89, United States Department of Agriculture.

OECD. 2018. *Rethinking Urban Sprawl*. URL https://www.oecd-ilibrary.org/content/publication/9789264189881-en.

O'Flaherty, Coleman A. 1996. *Transport Planning and Traffic Engineering*. Taylor & Francis.

Oueslati, Walid, Seraphim Alvanides, and Guy Garrod. 2015. Determinants of urban sprawl in European cities. *Urban Studies* 52(9): 1594–1614.

Paulsen, Kurt. 2012. Yet even more evidence on the spatial size of cities: Urban spatial expansion in the US, 1980–2000. *Regional Science and Urban Economics* 42(4): 561–568.

Redding, Stephen J. and Matthew A. Turner. 2015. *Transportation Costs and the Spatial Organization of Economic Activity*, *Handbook of Regional and Urban Economics*, volume 5, chapter 0. Elsevier, 1339–1398.

Riley, Shawn J., Stephen D. DeGloria, and Robert Elliot. 1999. A terrain ruggedness index that quantifies topographic heterogeneity. *Intermountain Journal of Science* 5(4): 23–27.

Stock, Jame H. and Motohiro Yogo. 2005. Testing for weak instruments in linear iv regression. In Donald W.K. Andrews and James H. Stock (eds.) *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge: Cambridge University Press, 80–108.

Appendix A. Residential land and transportation networks in the largest 60 FUAs

In Table A.1, I report individual computations for the 60 largest FUAs, i.e. those with populations over one million in 2010. For each variable, I highlight in red and blue the 5 highest and the 5 lowest values, respectively. According to their residential area, the 5 largest cities in 2012 were London (UK), Paris (FR), Berlin (DE), Essen (Ruhrgebiet, DE) and Warszawa (PL), and the 5 smallest were Sevilla (ES), Valencia (ES), Rotterdam (NL), Bucuresti (RO), and Torino (IT). Most cities increased their residential area and the FUAs with the highest growth were Warszawa (PL), Kraków (PL), Athina (EL), Madrid (ES) and Hamburg (DE). In fact, there were 5 cities that reduced their size: Barcelona (ES), Birmingham (West Midlands, UK), Helsinki (FI), Manchester (UK), and Sofia (BF). If we consider residential density, inhabitants per sq km of residential land, the densest cities in 2012 were Barcelona (ES), Valencia (ES), Madrid (ES), Bucuresti (RO), and Sevilla (ES), and the least dense were Bordeaux (FR), Dresden (DE), Bruxelles (BE), Bremen (DE), and Wien (AT). In absolute values, the cities that increased their residential density most were Barcelona (ES), London (UK), Helsinki (FI), Oslo (NO), and Zürich (CH), and those that reduced their density most were Kraków (PL), Gdansk (PL), Athina (EL), Warszawa (PL), and Bucuresti (RO).

In 2012, the most compact cities were Liverpool (UK), Manchester (UK), Bucuresti (RO), Rotterdam (NL), and Antwerpen (BE), and the most fragmented were Warszawa (PL), Wien (AT), Praha (CZ), Paris (FR), and Berlin (DE). Between 1990 and 2012, while there were cities that became more compact decreasing the number of residential lots, the case of Helsinki (FI), Napoli (IT), Torino (IT), Lille (FR) and København (DK), most FUAs became more fragmented, the case of Hamburg (DE), Wien (AT), Leipzig (DE), Kraków (PL) and Warsawa (PL). If we also consider the size of the residential area (lots per sq km), Liverpool (UK) and Manchester (UK) are still the least fragmented cities (per area), followed by West Midlands (UK), Athina (EL), and London (UK). Similarly, the FUAs with most lots per area were Praha (CZ), Leipzig (DE), Braunschweig (DE), Ostrava (CZ), and Dresden (DE).

Finally, this heterogeneity can also be observed in the surroundings indicator computed textita la Burchfield *et al.* (2006). In 2012, the lowest rates of undeveloped surroundings were in Birmingham (West Midlands, UK), Liverpool (UK), Rotterdam (NL), London (UK) and Manchester (UK), and the cities with most sprawl were Leipzig (DE), Dresden (DE), Praha (CZ), Ostrava (CZ) and Kraków (PL). Between 1990 and 2012, a quarter of the 60 FUAs (14) increased their undeveloped surroundings, the case of Gdansk (PL), Krakóv (PL), Warszawa (PL), Leipzig (DE) and Dresden (DE), while this index decreased in the other 46 cities, including Sevilla (ES), Valencia (ES), Porto (PT), Amsterdam (NL) and Rotterdam (NL).

Table A.1: Residential land in 60 FUAs with population over one million, 1990–2012

		Area						Fragmentation						Surroundings		
FUA (Country)	FUA Km^2	1990 Km^2	2012 Km^2	90–12 ↑↓	1990 $\frac{Pop}{Res}$	2012 $\frac{Pop}{Res}$	90–12 ↑↓	1990 *Num*	2012 *Num*	90–12 ↑↓	1990 $\frac{Lots}{Res}$	2012 $\frac{Lots}{Res}$	90–12 ↑↓	1990 %	2012 %	90–12 ↑↓
Amsterdam (NL)	2,914	346	388	↑	5,922	6,404	↑	101	107	↑	0.29	0.28	↓	19.9	15.5	↓
Antwerpen (BE)	1,191	308	314	↑	3,192	3,380	↑	66	67	↑	0.21	0.21	=	27.9	26.1	↓
Athina (EL)	3,030	386	543	↑	9,030	6,896	↓	85	102	↑	0.22	0.19	↓	16.5	15.6	↓
Barcelona (ES)	1,794	384	358	↓	10,323	12,482	↑	169	171	↑	0.44	0.48	↑	29.2	26.3	↓
Berlin (DE)	17,484	1,329	1,336	↑	3,663	3,790	↑	972	997	↑	0.73	0.75	↑	31.3	30.1	↓
Bordeaux (FR)	5,543	381	444	↑	2,453	2,559	↑	202	229	↑	0.53	0.52	↓	32.3	31.1	↓
Braunschweig (DE)	4,128	297	333	↑	3,394	3,064	↓	322	348	↑	1.08	1.04	↓	43.0	40.3	↓
Bremen (DE)	5,895	375	419	↑	3,124	2,972	↓	288	313	↑	0.77	0.75	↓	39.7	36.5	↓
Bruxelles (BE)	3,266	843	848	↑	2,587	2,961	↑	248	253	↑	0.29	0.30	↑	41.4	40.2	↓
Bucuresti (RO)	1,078	203	226	↑	11,061	9,657	↓	60	58	↓	0.30	0.26	↓	18.1	16.1	↓
Budapest (HU)	6,077	660	721	↑	4,343	3,986	↓	215	216	↑	0.33	0.30	↓	25.0	23.2	↓
Dresden (DE)	5,835	463	507	↑	3,198	2,686	↓	416	502	↑	0.90	0.99	↑	41.7	44.3	↑
Dublin (IE)	6,991	327	395	↑	4,112	4,534	↑	117	144	↑	0.36	0.36	=	20.0	20.2	↑
Düsseldorf (DE)	1,202	277	285	↑	5,401	5,362	↓	107	151	↑	0.39	0.53	↑	27.0	24.7	↓
Frankfurt (DE)	4,303	471	517	↑	5,039	4,937	↓	356	389	↑	0.76	0.75	↓	36.3	34.5	↓
Gdansk (PL)	2,630	140	245	↑	7,366	4,559	↓	80	191	↑	0.57	0.78	↑	29.4	38.3	↑
Glasgow (UK)	3,373	383	413	↑	4,552	4,326	↓	76	95	↑	0.20	0.23	↑	16.6	16.8	↑
Grad Zagreb (HR)	3,903	242	249	↑	4,523	4,576	↑	130	136	↑	0.54	0.55	↑	37.5	37.1	↓
Hamburg (DE)	7,343	744	851	↑	3,865	3,771	↓	458	578	↑	0.62	0.68	↑	31.8	30.9	↓
Hannover (DE)	2,973	321	345	↑	3,829	3,749	↓	230	251	↑	0.72	0.73	↑	36.5	35.0	↓
Helsinki (FI)	3,822	471	442	↓	2,255	3,055	↑	232	200	↓	0.49	0.45	↓	26.6	25.1	↓
Katowice (PL)	3,945	483	544	↑	5,882	4,782	↓	281	335	↑	0.58	0.62	↑	34.3	36.1	↑
Kraków (PL)	3,757	172	421	↑	7,530	3,259	↓	161	315	↑	0.94	0.75	↓	45.1	49.6	↑
Köln (DE)	1,626	335	360	↑	5,352	5,305	↓	131	161	↑	0.39	0.45	↑	25.8	25.9	↑
København (DK)	2,789	470	483	↑	3,624	3,917	↑	155	151	↓	0.33	0.31	↓	17.0	15.9	↓
Leeds (UK)	1,494	231	252	↑	4,579	4,605	↑	64	89	↑	0.28	0.35	↑	21.7	23.5	↑
Leipzig (DE)	3,979	298	332	↑	3,775	3,130	↓	250	374	↑	0.84	1.12	↑	37.5	40.4	↑
Lille (FR)	1,443	265	289	↑	4,835	4,674	↓	148	144	↓	0.56	0.50	↓	30.0	28.0	↓
Lisboa (PT)	3,901	387	488	↑	6,477	5,767	↓	206	246	↑	0.53	0.50	↓	32.5	29.6	↓
Liverpool (UK)	725	287	288	↑	5,297	5,235	↓	27	26	↓	0.09	0.09	=	10.7	10.2	↓
London (UK)	8,024	1,867	1,887	↑	5,430	6,415	↑	325	367	↑	0.17	0.19	↑	14.4	13.4	↓
Lyon (FR)	3,670	443	474	↑	3,693	4,028	↑	290	308	↑	0.65	0.65	=	36.3	33.5	↓
Madrid (ES)	8,025	545	654	↑	9,222	9,669	↑	234	309	↑	0.42	0.47	↑	27.5	24.6	↓
Manchester (UK)	1,817	560	532	↓	4,601	5,209	↑	42	53	↑	0.08	0.10	↑	14.7	13.5	↓
Mannheim (DE)	2,045	223	243	↑	5,050	4,826	↓	168	180	↑	0.75	0.74	↓	36.1	34.9	↓
Marseille (FR)	4,235	360	421	↑	4,221	4,120	↓	167	202	↑	0.46	0.48	↑	30.2	29.2	↓
Milano (IT)	2,638	521	562	↑	7,435	7,230	↓	257	258	↑	0.49	0.46	↓	28.1	26.6	↓
München (DE)	5,499	485	512	↑	4,785	5,247	↑	313	365	↑	0.65	0.71	↑	30.7	30.5	↓
Napoli (IT)	1,552	362	392	↑	9,217	9,017	↓	112	98	↓	0.31	0.25	↓	35.7	32.5	↓
Newcastle (UK)	5,437	227	248	↑	4,905	4,618	↓	76	92	↑	0.34	0.37	↑	19.3	18.9	↓
Nürnberg (DE)	2,934	278	287	↑	4,374	4,530	↑	199	246	↑	0.72	0.86	↑	36.4	37.5	↑
Oslo (NO)	7,428	345	351	↑	2,763	3,509	↑	152	154	↑	0.44	0.44	=	24.4	24.4	=
Ostrava (CZ)	3,878	340	342	↑	3,471	3,331	↓	340	344	↑	1.00	1.01	↑	47.9	47.4	↓
Paris (FR)	12,098	1,742	1,793	↑	6,073	6,566	↑	982	981	↓	0.56	0.55	↓	25.6	24.7	↓
Porto (PT)	952	209	242	↑	5,693	5,380	↓	87	105	↑	0.42	0.43	↑	36.5	31.5	↓
Praha (CZ)	6,980	548	604	↑	3,607	3,525	↓	722	761	↑	1.32	1.26	↓	45.9	44.5	↓
Roma (IT)	5,744	506	551	↑	7,430	7,269	↓	280	284	↑	0.55	0.51	↓	32.0	30.6	↓
Rotterdam (NL)	1,518	188	209	↑	6,956	6,761	↓	65	67	↑	0.35	0.32	↓	16.2	12.9	↓
Ruhrgebiet (DE)	4,440	1,011	1,060	↑	5,338	4,858	↓	301	335	↑	0.30	0.32	↑	23.0	20.5	↓
Sevilla (ES)	3,076	125	142	↑	9,037	9,415	↑	68	70	↑	0.55	0.49	↓	30.4	24.7	↓
Sofia (BG)	5,717	332	327	↓	4,440	4,591	↑	243	255	↑	0.73	0.78	↑	34.6	34.5	↓
Stockholm (SE)	7,093	582	592	↑	2,823	3,471	↑	269	279	↑	0.46	0.47	↑	19.1	19.3	↑
Stuttgart (DE)	3,654	482	499	↑	5,151	5,366	↑	360	402	↑	0.75	0.80	↑	39.2	37.0	↓
Torino (IT)	1,781	218	227	↑	8,134	7,706	↓	115	110	↓	0.53	0.48	↓	28.3	28.3	=
Toulouse (FR)	5,246	306	389	↑	2,919	3,184	↑	192	232	↑	0.63	0.60	↓	36.2	32.9	↓
Valencia (ES)	1,443	130	165	↑	10,897	10,146	↓	104	118	↑	0.80	0.72	↓	37.2	31.9	↓
Warszawa (PL)	8,615	561	917	↑	4,944	3,288	↓	368	725	↑	0.66	0.79	↑	35.5	39.7	↑
West Midlands (UK)	2,075	613	581	↓	4,369	4,920	↑	62	83	↑	0.10	0.14	↑	9.5	9.8	↑
Wien (AT)	9,205	822	890	↑	2,891	3,020	↑	616	740	↑	0.75	0.83	↑	39.1	40.3	↑
Zürich (CH)	1,090	253	254	↑	4,041	4,726	↑	117	116	↓	0.46	0.46	=	30.2	30.1	↓

Notes: 'Land area' variables are measured in square kilometers. '$\frac{Pop}{Res}$' is the number of inhabitants per sq km of residential land. 'Fragmentation' refers to the number of residential land lots and '$\frac{Lots}{Res}$' is the number of residential land lots per sq km of residential land. 'Surroundings' refers to the % of undeveloped land surrounding residential land. Red and blue numbers show the 5 highest and the 5 lowest values, respectively.

Table A.2: Highways and old (rail)roads in 60 FUAs with population over one million

Km of	Highways						
FUA (Country)	1990	2010	1990–2010	Roman	Trade	1810 Post	1870 Rail
Amsterdam (NL)	228	275	↑	0	0	0	24
Antwerpen (BE)	126	167	↑	1	12	6	25
Athina (EL)	91	177	↑	64	0	25	0
Barcelona (ES)	205	338	↑	72	0	68	116
Berlin (DE)	395	395	=	0	10	135	186
Bordeaux (FR)	159	179	↑	36	0	23	20
Braunschweig (DE)	184	184	=	0	47	53	59
Bremen (DE)	117	134	↑	0	28	82	54
Bruxelles (BE)	267	267	=	34	0	38	152
Bucuresti (RO)	10	28	↑	0	0	26	1
Budapest (HU)	129	246	↑	25	20	35	45
Dresden (DE)	140	179	↑	0	13	30	88
Dublin (IE)	39	216	↑	0	0	169	100
Düsseldorf (DE)	223	249	↑	8	3	10	114
Frankfurt (DE)	293	293	=	30	53	77	129
Gdansk (PL)	0	18	↑	0	2	36	6
Glasgow (UK)	88	162	↑	15	0	98	237
Grad Zagreb (HR)	95	192	↑	19	0	18	16
Hamburg (DE)	259	345	↑	0	20	110	59
Hannover (DE)	171	171	=	0	30	25	57
Helsinki (FI)	147	180	↑	0	0	0	15
Katowice (PL)	83	122	↑	0	21	20	138
Kraków (PL)	43	62	↑	0	27	12	14
Köln (DE)	149	149	=	14	59	25	99
København (DK)	141	141	=	0	0	35	15
Leeds (UK)	27	89	↑	44	0	43	157
Leipzig (DE)	94	124	↑	0	36	32	116
Lille (FR)	123	123	=	13	0	24	30
Lisboa (PT)	162	288	↑	40	0	56	16
Liverpool (UK)	59	59	=	8	0	94	142
London (UK)	460	460	=	220	0	438	784
Lyon (FR)	235	235	=	52	0	44	33
Madrid (ES)	105	679	↑	23	0	141	45
Manchester (UK)	146	146	=	71	0	124	343
Mannheim (DE)	132	132	=	25	10	50	97
Marseille (FR)	286	286	=	57	0	32	22
Milano (IT)	193	205	↑	119	60	138	103
München (DE)	281	281	=	52	6	88	53
Napoli (IT)	105	110	↑	87	0	33	53
Newcastle (UK)	2	4	↑	29	0	78	94
Nürnberg (DE)	172	172	=	0	60	84	64
Oslo (NO)	139	139	=	0	0	49	25
Ostrava (CZ)	0	54	↑	0	0	46	59
Paris (FR)	530	535	↑	131	16	200	260
Porto (PT)	70	198	↑	18	0	47	7
Praha (CZ)	130	149	↑	0	17	52	53
Roma (IT)	225	271	↑	1162	0	48	29
Rotterdam (NL)	97	97	=	0	0	0	23
Ruhrgebiet (DE)	563	563	=	20	12	10	343
Sevilla (ES)	41	151	↑	23	0	45	8
Sofia (BG)	84	121	↑	43	0	18	0
Stockholm (SE)	250	278	↑	0	0	15	13
Stuttgart (DE)	89	89	=	86	2	33	82
Torino (IT)	135	151	↑	61	0	16	80
Toulouse (FR)	113	204	↑	28	0	20	48
Valencia (ES)	19	75	↑	12	0	15	9
Warszawa (PL)	0	0	=	0	15	18	49
West Midlands (UK)	180	180	=	45	0	144	267
Wien (AT)	214	277	↑	41	19	77	99
Zürich (CH)	145	145	=	21	2	44	77

Notes: Red and blue numbers show the 5 highest and the 5 lowest values, respectively.

Table A.2 shows the evolution of the different transportation networks in the 60 most populated

European cities. For each network, I highlight in red and blue the 5 highest and the 5 lowest values, respectively. In the case of highways, the largest networks in 2012 were in Madrid (ES), Essen (Ruhrgebiet, DE), Paris (FR), London (UK) and Berlin (DE), and the smallest were in Warszawa (PL), Newcastle (UK), Gdansk (PL), Bucuresti (RO) and Ostrava (CZ). Between 1990 and 2010, 23 cities, mainly Northwestern, did not change their highway length, the case of London (UK), Berlin (DE), Bruxelles (BE), Lille (FR), Rotterdam (NL) and Oslo (NO), among others. The remaining 556 FUAs increased their highway network, most notably Madrid (ES), Dublin (IE), Barcelona (ES), Porto (PT) and Lisboa (PT).

Table A.2 also report length computations for old transportation networks in Europe. The largest Roman networks were mainly built in Mediterranean cities (e.g., Rome (IT), Barcelona (ES), Athina (EL), Marseille (FR) and Lisboa (PT)), but also in Northwestern cities (e.g., London (UK), Paris (FR), Stuttgart (DE), Manchester (UK) and München (DE)) and East European cities (Zagreb (HR), Budapest (HU) and Sofia (BF)). The main trade routes in the Holy Roman Empire in the 15th century linked Central and Eastern European cities (e.g., Berlin (DE), Wien (AT), Warszawa (PL), Budapest (HU) and Zelenogradsk (RU)), but also some other leading European cities (e.g., Paris (FR), Basel (CH), Bruxelles (BE), Genova (IT) or Milano (IT)). Postal roads (in 1810) were built mostly in Northwestern cities (e.g., London (UK), Paris (FR), Dublin (IE) and Birmingham (UK)), but also in some Mediterranean and East European cities (e.g., Madrid (ES) and Praha (CZ)). Finally, Railroad construction clearly benefited Northwestern cities such as London (UK) (780 km) and Manchester (UK) (343 km), but also some Mediterranean cities such as Barcelona (ES) (116 km) and Milano (IT) (103 km), and East European cities such as Katowice (PL) (137 km) and Leipzig (DE) (116 km).

Appendix B. Some descriptive statistics

Table B.1: Summary statistics for geography, history and socieconomy variables in FUAs

	Mean	S.D.	Min	Max
Population				
2010 Population (*'000*)	518	971	68	12,100
2000 Population (*'000*)	492	901	62	10,980
1990 Population (*'000*)	477	869	54	10,580
Railroad				
2010 Km of railroads	90.2	129	0	1,381
2000 Km of railroads	82.6	122.4	0	1,368
1990 Km of railroads	54.9	70.8	0	595.5
Socioeconomy				
2010 GDP (*€bn 2005 prices*)	16.3	22.8	0.19	200.0
2000 GDP (*€bn 2005 prices*)	14.1	19.3	0.220	163.1
1990 GDP (*€bn 2005 prices*)	11.4	15.1	0.514	129.4
2010 Unemployment rate proxy (%)	13.5	8.3	0.07	47.4
2000 Unemployment rate proxy (%)	13.0	8.9	0.00	45.0
1990 Unemployment rate proxy (%)	13.1	10.3	0.14	51.8
2010 Employment in manufacturing (%)	16.5	7.6	2,0	40.3
2000 Employment in manufacturing (%)	19.6	8.1	3.5	43.6
1990 Employment in manufacturing (%)	23.4	9.2	2.8	54.3
Geography				
Total area (Km^2)	1,686	1,814	57.4	17,484
Terrain ruggedness index	82.5	239.1	0	3,756
Altitude (*m*)	62.4	41.6	0.9	169.7
Elevation range (*m*)	78.3	44.8	2	186
City history				
Roman cities	0.14	0.35	0	1
Medieval cities	0.35	0.48	0	1
16th–19th c. cities	0.25	0.43	0	1
Population history				
1980 Population (*'000*)	460	843	49	9,985
1970 Population (*'000*)	432	827	37	10,720
1960 Population (*'000*)	385	766	24	10,620
Highway history				
1980 Km of highways	29.4	50	0	428.7
1970 Km of highways	16.0	33.9	0	358.7
1960 Km of highways	4.5	19.6	0	310.7
Railroad history				
1980 Km of railroads	54.7	70.1	0	580.6
1970 Km of railroads	54.6	70.0	0	580.6
1960 Km of railroads	54.6	70.0	0	580.6

Appendix C. First-stage and reduced-form results

Table C.1: Modern highways, historical (rail)roads and residential land in European cities

Specification:	FD regressions with FUA characteristics					FD regressions with FUA fixed-effects				
Regression type:	First-Stage		Reduced-Form			First-Stage		Reduced-Form		
Dependent variable:	Δln(Km hwys)		Δln(A)	Δln(F)	Δ%Und	Δln(Km hwys)		Δln(A)	Δln(F)	Δ%Und
Method:	OLS [1]	OLS [2]	OLS [3]	OLS [4]	OLS [5]	OLS [6]	OLS [7]	OLS [8]	OLS [9]	OLS [10]
Δln(Km Roman roads)	1.0029^{a}	0.8754^{a}	0.1408^{a}	0.1762^{a}	1.5696^{c}	1.4521^{a}	1.2018^{a}	0.1325^{a}	0.2390^{a}	2.9511^{a}
	(0.2548)	(0.2406)	(0.0293)	(0.0651)	(0.9343)	(0.3004)	(0.2882)	(0.0284)	(0.0625)	(0.8898)
Δln(Km Trade routes)	0.4123					0.4397				
	(0.3064)					(0.3086)				
Δln(Km 1810 Postal roads)	-0.2840					-0.4472				
	(0.2468)					(0.3208)				
Δln(Km 1870 Railroads)	0.1509					0.2793				
	(0.2448)					(0.3220)				
Δln(Socioeconomy)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Geography	Y	Y	Y	Y	Y	N	N	N	N	N
History	Y	Y	Y	Y	Y	N	N	N	N	N
Country FE	Y	Y	Y	Y	Y	N	N	N	N	N
FUA FE	N	N	N	N	N	Y	Y	Y	Y	Y
First-stage F-statistic	6.32	13.23	13.23	13.23	13.23	6.45	17.39	17.39	17.39	17.39

Notes: All regressions include a constant and decade effects. 1,158 observations (579 FUAs × 2 decades) in each regression. Historical instruments are time-variant because they are computed by multiplying their kilometers by the fraction of the highway network kilometrage in each country completed at each decade (excluding each city's own contribution). Robust standard errors clustered by FUA are in parentheses. ^{a}, ^{b}, and ^{c} indicates significant at 1, 5, and 10 percent level, respectively.