Institutional Fragmentation and Urbanisation in the EU Cities

Federica Cappelli, Gianni Guastella, Stefano Pareglio
Institutional Fragmentation and Urbanisation in the EU Cities

By Federica Cappelli, Roma Tre University
Gianni Guastella, Università Cattolica del Sacro Cuore and Fondazione Eni Enrico Mattei
Stefano Pareglio, Università Cattolica del Sacro Cuore and Fondazione Eni Enrico Mattei

Summary

This article examines the relationship between institutional fragmentation and the spatial extent of cities in Europe’s Functional Urban Areas. European Union planning regulations vary across member states, but in most cases, local authorities determine land use within the more general regulatory frameworks set by national or subnational authorities. More decentralised and fragmented settings may favour urban sprawl, allowing developers to avoid land-use restrictions in one municipality by moving to adjacent ones and providing incentives for municipalities to adopt less strict land-conversion regulations to attract households and workers. The empirical results fully support this hypothesis and unveil significant differences between small and large cities, the effect of governance fragmentation being a substantial factor in the latter case.

Keywords: Urban Sprawl, Institutional Fragmentation, Threshold Regression

JEL Classification: R52, R58, C24

This is a preliminary version, not for quote. The final version of the paper is published in Regional Studies with doi: 10.1080/00343404.2020.1800625

Address for correspondence:
Gianni Guastella
Fondazione Eni Enrico Mattei
Corso Magenta, 63
20123 Milano
Italy
E-mail: gianni.guastella@feem.it

The opinions expressed in this paper do not necessarily reflect the position of Fondazione Eni Enrico Mattei
Corso Magenta, 63, 20123 Milano (I), web site: www.feem.it, e-mail: working.papers@feem.it
Introduction

During the last decades, European Union (EU) cities experienced unprecedented spatial growth. Europe’s urbanisation level in 2018 was 74%, and it is expected to increase to 83.7% by 2050 (UN-DESA, 2018). However, the urbanisation trend that has characterised Europe and the US in recent years is moving in the opposite direction with respect to the rest of the world, where populations are growing faster than or at the same rate of built-up areas (Alberti et al., 2019). This means that cities in Europe and the United States (US) are no longer growing in height (and, then, compactly); instead, more and more land is being converted to accommodate new citizens. In addition, evidence from European cities shows that the density of urban settlements decreased, and their discontinuity increased, especially in small and medium-sized cities due to shrinking populations (Guastella et al., 2019). Many consider this growth unnecessary and the result of market or political inefficiencies (Downs, 2005): market fails to price the negative externalities on the environment, and local policymakers pay low attention to the issue of efficient land use. Many others oppose this view (Brueckner, 2000), claiming that land-use restrictions limit people utility, lowering the average living space below desirable levels (Mills, 2006) and increasing the price of houses (Ehrlich et al., 2018). In fact, urban sprawl has multiple benefits, such as the availability of larger houses, easier air ventilation and lower pollution, greater availability of green spaces, etc. However, the costs are also high in relation to the transformation of agricultural and natural land at high rates, the loss of related ecosystem services, and the push towards car-based commuting and the related pollution. Appendix A provides a complete overview of the potential costs and benefits.

Several studies have investigated the drivers of urban sprawl by looking at single cases (Couch and Kerecha, 2006; Hewitt and Escobar, 2011; Salvati, 2016) and cross-country and cross-city comparisons in the US (Brueckner and Fansler, 1983; McGrath, 2005; Paulsen, 2012; Spivey, 2008) and Europe (Patacchini and Zenou, 2009; Schwartz, 2010; Siedentop and Fina, 2012; Oueslati et al., 2015). The increase in housing demand driven by demographic trends and income growth, alongside the increased use of cars as a prevalent mode of transportation and low farmland values, explain urban spatial expansion to the greatest extent, estimated at 80% of total city size variation by Paulsen (2012). Likewise, many studies have looked at the effectiveness of policy measures to contrast urban sprawl, usually focusing on single-case scenarios and largely on the effectiveness of containment policies and planning tools (Brueckner and Sridhar, 2012; Gennaio et al., 2009; Wassmer, 2006; Woo and Guldmann, 2011; Paulsen, 2013). The results vary from context to context, but, generally, the effectiveness of local containment policies is strongly related to the governance structure. Weak governance systems within a territory likely lead to ineffective land-use policies.
This paper proposes an empirical cross-city analysis of urbanisation and sprawl aimed at disentangling the specific role of governance and, in particular, administrative fragmentation.

The empirical framework of the paper grounds on the empirical model firstly proposed by Brueckner and Fansler (1983) to test the validity of the monocentric model (Alonso, 1964; Muth, 1969; Mills, 1972 [AMM]). In this model, four primary variables determine the equilibrium size of the city: the total population, the median household income, the cost of commuting, and the agricultural rents at the urban fringe. Several studies support the validity of the monocentric city model empirically (Brueckner and Fansler, 1983; Deng et al., 2008; Ke et al., 2009; McGrath, 2005; Spivey, 2008; Paulsen, 2012; Oueslati et al., 2015). Additionally, evidence suggests that specific factors including demographics, household preferences and emerging lifestyles, contribute to more dispersed urban forms (Anas et al., 1998; Brueckner, 2000; Pisman et al., 2011; Cirtautas, 2013; Jaeger and Schwick, 2014).

General aspects related to urban governance and, more specifically, institutional fragmentation have been studied extensively but rarely analysed empirically in the context of cross-city comparisons. In his historical review of urbanisation and suburbanisation in the US, Teaford (1979) traces the roots of political, cultural, and social processes that have driven the spatial decentralisation and institutional fragmentation of US metropolitan areas. Lewis (1996) proposed a politico-economic analysis of metropolitan growth patterns, pointing to the fundamental role of institutional organisations in the process of aggregating agents’ preferences and translating them into decision-making, especially when it comes to land use. Rusk (1998) also recognised the strict interconnection between the balkanisation of local governance and the spatial growth of US metropolitan cities, and suggested that only higher level, in particular, regional agendas may effectively respond to the many challenges of growth management. The same conclusion was reached by Orfield and Luce (2010) based on a case-study analysis of the Twin Cities. They argued that American metropolitan areas have evolved from the simple city-suburb scheme to more complex ‘mosaics’, with jobs and house units dispersed across very large spaces and governed by many local governments.

This general result suggests that more administratively fragmented metropolitan areas should also be characterised more by urban sprawl. Empirical analyses of the relationship between urban sprawl and governance fragmentation have been carried out previously at the county level in the US (Carruthers and Ulfarsson, 2002) and the EU (Ehrlich et al., 2018). Only two studies have considered the city as a unit of analysis. Paulsen (2014) considered institutional fragmentation among the causes of dynamic sprawl, while Carruthers (2003) focussed on its effects on urban growth but only considered the fringe of the cities; both studies considered metropolitan areas in the US. The present work aims to fill this gap by providing empirical evidence for EU cities.

This paper extends the Brueckner and Fansler (1983) empirical model to consider institutional fragmentation among the determinants of the spatial extent of built-up metropolitan areas in Europe. We employ 2012 Copernicus land cover data to measure the spatial extent of a
sample of EU Functional Urban Areas (FUAs) with more than 50,000 inhabitants and link this information to socio-economic indicators suggested by the empirical literature and to a measure of governance fragmentation, which is the number of municipal authorities within the FUA per thousand inhabitants. The definition of municipal authority corresponds to levels one or two of the European classification of Local Administrative Units (LAU), depending on the country. We hypothesise and empirically test that a higher number of local governments deciding autonomously on land use induces larger urban expansion ceteris paribus.

The study frames into the body of literature on the determinants of city size, to which it also contributes in several aspects. Firstly, it focuses on the EU, where considerably lower empirical evidence exists on the determinants of city size compared to the US. Secondly, it expands the scope of the empirical analysis to small cities, for which relatively less evidence exists about the functioning of the monocentric model. Finally, it allows for scale-related structural heterogeneity identifying endogenous threshold effects. This last contribution is relevant to disentangle the structural differences about the role of administrative fragmentation between small and large cities. Our research hypothesis is that the coordination of planning policies becomes progressively more difficult as the size of the metropolitan area increases and may be a relevant issue in large areas only. However, the threshold size at which fragmentation shows its impact is not known a priori and needs to be considered endogenously and, accordingly, estimated simultaneously with the other regression parameters.

The remaining paper is structured as follows. The following section reviews the relevant literature concerning the determinants of city size, the role of institutional fragmentation, urban spatial size and the institutional setting of European cities. Next two sections are dedicated to the description of the data used and the illustration of the empirical model and methodological approach. The empirical results are then presented and discussed, and a final section concludes and draws the resulting policy implications.
**The determinants of city size**

In the urban economics tradition of the AMM model (Alonso, 1964; Muth, 1969; Mills, 1972), the observed spatial size of cities is understood as the result of population, income, transportation costs and agricultural land values (Brueckner and Fansler, 1983). Combined, these factors determine the spatial distribution of housing prices and, consequently, the household’s location choice. The nexus of population and income is clear because an increase in both factors increases the demand for houses. A decline in transportation costs accelerates the expansion of cities in the urban fringe where houses are more affordable. High agricultural land values contrast the spatial expansion of cities, making land development relatively more costly. The empirical literature that tested empirically the validity of this hypothesis refers largely to the US (Brueckner and Fansler, 1983; McGrath, 2005; Wassmer, 2006; Spivey, 2008; Paulsen, 2012), and only a minority to developing countries (Deng et al., 2008; Song et al., 2014) and India (Brueckner and Sridhar, 2012; Lata et al., 2003).

In Europe, Patacchini and Zenou (2009) were the first to present a sprawl analysis that refers to cities in the whole EU. Arribas-Bel et al. (2011) presented an analysis of sprawl that matches different socioeconomic and morphological dimensions of the phenomenon in Europe. Schwartz (2010) analysed the urban form of 231 European cities, looking at several landscape metrics and their cross-correlation. Wolff et al. (2018) compared residential density changes between 1990–2000 and 2000–2010 in European urban morphological zones (UMZs), defined by the spatial continuity of urbanisation. They found that declining density was the main characteristic of urban spatial expansion in Europe in both periods, although to different extents. Guastella et al. (2019) reached the same conclusion based on an analysis of European FUAs for the period 1990–2014, adding that the sprawl characteristics of low density and high discontinuity are typical of small and shrinking cities in Europe. Salvati et al. (2018) used explanatory statistical analysis to profile land consumption patterns of 155 metropolitan regions, matching land-use data with information about the socio-economic characteristics of the cities.

All the previous literature about urban sprawl in Europe approaches the topic with descriptive analyses aimed at representing sprawl or, in some cases, classifying cities according to the identified sprawl patterns, without deepening the investigation into the determinants of urban spatial size. Conversely, Oueslati et al. (2015) employed two different indexes of sprawl to explain the determinants of urban sprawl. Their results are consistent with those that emerged in the US concerning the validity of the monocentric model.

Beyond the tradition of urban economics literature, other factors have been related...
to the spatial expansion of cities. Fujita and Thisse (2002) argued that cities specialised in sectors that benefit most from agglomeration economies tend to be less dispersed. Saiz (2010) provided evidence that geographical factors influencing the amount of land that can be developed also drive housing demand and, consequently, sub-urbanisation patterns. The average cost of fuel (Molloy and Shan, 2013) and the network infrastructure also shape households’ location decisions and, hence, the spatial development of cities (Baum-Snow, 2007). Cavailhès et al. (2004) and Coisnon et al. (2014) linked households’ living decisions with the value of rural amenities, especially agricultural amenities. This evidence is relevant to explain not only the fast suburbanisation of many modern cities but also the occurrence of so-called leapfrogging development.

Institutions, policies and sprawl

Institutions and policies play a substantial role in determining households’ location decisions, and the final effect on sprawl depends mostly on their interaction. Tiebout (1956) confuted the analysis on the level of national and federal expenditures on public goods provided by Musgrave (1939) and Samuelson (1954), claiming that such a theoretical framework does not apply to local expenditures, which entail a shift in the decision-maker agent from the national government to the consumer-voter.1 He proposed the idea of households ‘voting with their feet’ to describe location choices depending on local public policies, including fiscal and land development policies. Local authorities, in fact, finance the provision of public services, using resources raised with local taxation in addition to transfers from central authorities, thereby generating a heterogeneous supply of taxes and services that favours sorting among municipalities in the same metropolitan area. The institutional setting of a metro area, however, largely influences the outcome of this sorting (Ehrlich et al., 2018). Fragmented and decentralised settings, allowing for a higher degree of autonomy at the local level, are characterised by higher competition among municipalities, which can hardly be on tax policies.2 Instead, local governments compete using land-development policies. Sometimes, these policies are aimed at protecting the value of houses in a specific site by limiting growth but, together, originate a growth of artificial area that shifts from site to site and results in urban sprawl (Carruthers and Ulfarsson, 2002). In other circumstances, these policies are aimed at encouraging growth through lower land-use regulations and leveraging on the relatively low price of land.

Another channel through which institutional settings and local policies determine urban sprawl is zoning, whereby households self-select zones that are homogeneous with respect to residents’ socio-economic conditions (Shertzer et al., 2018). Although also very common in most European countries, the meaning of zoning differs significantly between the US and Europe (Hirt, 2007, 2012, 2013). Generally, zoning is applied to provide public regulations that limit private-sector activity in terms of building options in the territory’s subdivisions. However, in Europe, these regulations mainly concern building characteristics, (height, colour, architectural style, etc.), while in the US, zoning pertains exclusively to the functions those buildings perform (Hirt, 2012, 2013). Zoning with respect to function is also pursued in most European countries, but mixed uses are always allowed.
What matters in Europe is the proportion in which residential and commercial buildings are found (Hall, 2006).

Since the 1960s, zoning has been harshly criticised because of social and ecological concerns. Residents have an incentive to get policies enacted that emphasise land-use restrictions, limiting land use in the zone to preserve the value of houses (Albouy and Ehrlich, 2018). The limits to the development of new land together with the high prices of homes refrain lower income households from moving into the periphery, consequently characterising low-density housing (Hilber and Robert-Nicoud, 2013). This contributes to increased segregation of populations by class and race (Frug, 1996) and, by strongly encouraging car use, to favour sprawl and intensified pollution and land waste (Kleppel, 2002). These effects are more likely to appear in very fragmented and decentralised institutional settings: the smaller the size of the community, the higher the control on political groups and the easier the implementation of rules that increase residential values (Carruthers, 2003). Thus, institutional fragmentation facilitates zoning, and despite promoting the enforcement of strict land-use policies in single municipalities, it favours the spread of low densities in the suburbs at the aggregate level. This is confirmed by the differences in the allocation of land-use planning powers between the US and European countries, where the former is much more decentralised than the latter (Hirt, 2012). Additionally, huge differences exist among EU MS: for instance, the United Kingdom is the most centralised country in Europe while Mediterranean countries are the most decentralised. However, in all European countries, the national government plays a role in land-use planning, although marginal in some cases. Conversely, in the US, land use is not really planned, as the whole system is grounded on zoning at the local level (Hall, 2002).

The last consideration concerns the local spillover effects of land-use policies. As argued in Glaeser and Kahn (2004), very fragmented and decentralised settings facilitate eluding policies that limit land development as developers can easily move to neighbouring municipalities with looser regulations. Thus, policies that put boundaries on the growth of developed land are found to be effective if managed in aggregate metropolitan areas (Cunningham, 2007). Conversely, when urban containment policies are enacted at the local level, evidence suggests that land development concentrates outside the urban core, generating sprawl (Brueckner and Sridhar, 2012). Paradoxically, policies aimed at containing urban growth at the local level may produce the opposite effect in neighbourhoods that, depending on their magnitude, may even generate a positive net effect on urban development at the aggregate level (Irwin and Bockstael, 2004).

The institutional setting of cities in the EU

Decisions concerning land use at the city level, both explicitly and implicitly, also affect all neighbouring communities, because of the many implications on other sectoral issues (e.g. housing, transportation, environment, tourism, etc.). Furthermore, in cities or metropolitan areas where the number of administrative entities equipped with competence on land use is high, a lack of coordination leaves little
scope for efficient and sustainable spatial development. Most metropolitan areas have a corresponding metropolitan authority, but only a few are endowed with effective administrative powers (Ahrend et al., 2014). Metropolitan areas should be integrated into urban governance as functional bodies to allow an integrated evaluation of sectoral policies while staying close to citizens. In this regard, the formalisation of FUAs (OECD, 2012) in Europe is the first step, as they avoid confusion generated by different definitions of metropolitan cities, allowing direct comparisons between different areas. However, they are merely statistical entities, while the administrative functions, including land-use planning, are still subject to municipalities’ decision-making.

Land-use plans can be divided into three categories depending on their function: policy guidelines, strategic plans and zoning/boundary plans. Policy guidelines are very general, referring to vast territories (usually national ones) and lacking geographic references. Strategic plans identify main policy challenges but remain general enough to allow a high flexibility in the choice of policy instrument. Zoning/boundary plans specify in detail the type of land use that is intended or permitted for a given location and typically contain the only binding rules for landowners. However, all plans can have more than one function simultaneously. Notwithstanding the considerable differences across EU MS, higher governmental levels usually provide plans containing general guidelines, while the narrower the controlled area, the wider and more binding are the authority’s powers. Indeed, the national government and/or the federal states usually produce integrated plans that cover a broad range of policy fields and often provide only policy guidelines (OECD, 2017). Regional plans, where they exist, are more detailed than national ones and are usually oriented towards strategic planning. Metropolitan plans are very rare and prepared for single metropolitan areas, such as in Copenhagen, Auckland, Dublin, London and Budapest. The Territorial Coherence Scheme in France represents the only case in Europe where metropolitan plans are arranged for every metropolitan area of the country (OECD, 2017). Municipal plans are the most detailed and binding plans in all European countries. They contain elements of boundary plans and encompass a broad range of policy fields, aiming to define the intended use of a clear portion of land (OECD, 2017).

Although municipalities handle land-use planning in all European countries, different degrees of coordination with and supervision of cities or regions occur. For instance, in Austria, the national government prepares the Austrian Spatial Development Concept, which must be jointly approved by the nine federal states and 2,100 municipalities. The Austrian Conference for Spatial Planning coordinates land-planning decisions at all government levels. In addition to Austria, only another eight European countries (Denmark, Finland, Germany, Latvia, Luxembourg, Netherlands, Norway and Sweden) envisage a certain degree of coordination among municipalities when drawing up regional plans (Farínós Dasi, 2007). In contrast, in countries like Poland, where the hierarchical planning system is only on paper since the only binding zoning plans are the Local Spatial Development Plans, plans are drawn up at the municipal level by the 2,478 Gminas (OECD, 2017).
Differences in local governance typologies among EU MS are also dictated by the different spatial planning systems in place. CEC (1997) identified four prevailing spatial planning systems found in the then 15 EU countries: i) urbanism tradition, ii) land-use spatial planning, iii) regional economic approach and iv) comprehensive integrated approach. The ESPON project carried out by Farinós Dasí and Milder (2007) extended the analysis of the CEC (1997) to include all 27 EU countries as well as changes in spatial planning systems occurring in each country. This classification is neither static nor exclusive, as in many countries several systems coexist (Farinós Dasí and Milder, 2007). European countries have had, for decades, a long tradition of ‘urbanism’, characterised by a preventive binding zoning and assignment of land-use rights at the local level, where the only decision criterion is conformity with the plan (Janin Rivolin, 2017). However, the proliferation of laws and regulations has favoured private speculation and privatisation of vast urban soils to the detriment of public interest (Janin Rivolin, 2017) and effective control of urban development (CEC, 1997). Thus, this system has been abandoned by the vast majority of European countries but remains firmly in place in Mediterranean countries. Conversely, the land-use spatial planning system that originated in the UK is aimed at controlling the path of change in urban development, with a focus on sustainable growth (Farinós Dasí and Milder, 2007). Accordingly, any assignation of land-use right is discretionary (Booth, 2003) and subject to specific evaluations by local authorities, under the supervision of central administration. This guarantees coherence with the overall spatial development strategy and not merely conformity with the plan’s criteria (Cullingworth and Nadin, 2002; Janin Rivolin, 2017). The power exercised by local authorities is drastically reduced in the regional economic approach and in the comprehensive integrated approach. Countries belonging to the former usually have a hierarchical distribution of competencies, where territorial development is based on regional plans drawn up by either the regions or the national government. Through the pursuit of spatial justice, these plans address issues relevant for the whole region, including environmental, social and economic concerns, and are executed by authorities at the local level (CEC, 1997). Finally, in countries adopting the comprehensive integrated approach, both vertical and horizontal coordination can be found between all sectors and institutions dealing with spatial development (CEC, 1997; Farinós Dasí and Milder, 2007). Table 1 summarises the spatial planning system existing in each EU country (where, for each system, countries highlighted in bold have that system as their prevalent one):

<table>
<thead>
<tr>
<th>Urbanism</th>
<th>Land Use Planning</th>
<th>Regional Economic Approach</th>
<th>Comprehensive Integrated Approach</th>
</tr>
</thead>
</table>

Source: Farinós Dasí and Milder, 2007
In short, if decentralising land-use planning is advantageous in terms of flexibility and closeness to the community’s needs, decentralisation in the absence of effective coordination can produce adverse effects and create conflict with the objectives set at the national/regional levels (OECD, 2017) or in neighbouring communities. Although regulatory frameworks and the degree of interaction between different authorities vary across countries, the common tract of European planning policies is the municipalities’ direct power to decide on land-use change by identifying zones suitable for converting from natural or agricultural to urbanised areas. Since each municipality counts as a single decision centre, the number of municipalities within a city may effectively proxy the complexity of urban governance, making the measure suitable for our empirical analysis.
The dataset employed considers 359 European FUAs, our unit of observation, for which data is available. In Europe, a clear and harmonised definition of ‘city’ was lacking until very recently. Later, the EU-OECD classification was adopted, defining a city as a densely inhabited core surrounded by a less densely populated commuting zone, whose labour market is highly integrated with the city. The core centres are identified based on gridded population data and consist of continuous clusters of cells with population densities larger than 1,500 inhabitants. The hinterlands are defined based on commuting data: a municipality is part of the FUA if at least 15% of the residents work in the urban core. The sample is largely heterogeneous in terms of population size and covers a wide variety of city typologies, ranging from highly populated capitals to second-tier metropolitan areas and highly industrialised cities to smaller coastal or historical towns.

Table 2 describes the model variables and provides basic descriptive statistics. The dependent variable, (UA), is the total artificial area (artificial surface) in an FUA and was retrieved from the Corine Land Cover database 2012.4 The explanatory variables are divided into two groups. The first group includes the variables suggested by the related economic and econometric literature (Brueckner and Fansler, 1983; Wassmer, 2006; Spivey, 2008; Paulsen, 2012; Oueslati et al. 2015), namely population, median income, agricultural land values and transport costs. The population, sourced from the Eurostat city database, is measured in thousands of inhabitants (POP). We use per-capita gross domestic product (GDP), sourced from the same database, as a proxy for median income (INC). Eurostat does not provide city-level information on agricultural land values; hence, we use the agricultural value added per employee, a measure of average productivity in the agricultural sector that is only available at the Nomenclature Units for Territorial Statistics (NUTS) level III (sub-regional administrations) territorial detail, as the best proxy for agricultural land values (ALV). Following Glaeser and Kahn (2004), the inverse of the number of cars per inhabitant is included to proxy transport costs (TC), assuming that the diffusion of cars makes commuting easier, fostering urban spatial expansion. We are aware that this is an imperfect proxy for transportation costs, because it does not consider the availability of alternative transportation modes and, specifically, access to public transportation networks. Unfortunately, the lack of comparable data for the supply of public transportation services at the EU level impedes controlling the regression estimates. This misspecification likely causes an upward bias in the estimation of the TC parameter because in large metropolitan areas fewer people are willing to use private transportation simply because alternative modes are easier and more convenient to use. However, this should not affect the other parameter estimates.
The second group of variables covers our research hypothesis (administrative fragmentation) and control variables. The measure of administrative fragmentation is based on the official definition of boundaries provided by the geographical information system of the European Commission (GISCO-EUROSTAT). We construct this measure by superimposing the map of the FUA on that of the municipalities (communes) and computing the number of administrative units within the FUA (FRAG). The absolute measure is divided by the FUA’s population to be comparable across units of different sizes in the sample and to avoid spurious correlations between the total sizes of cities and their number of governments. The spatial scope of functional relations in European cities, especially large cities, likely extended over time to neighbouring municipalities due to lowering transport costs and agglomeration effects, causing today’s FUAs to be large in terms of built-up area and administratively fragmented. As control variables, we include the (log of) distance from the nearest capital (lnDIST) and the degree of polycentrism (POLY). The distance from the capital city is computed as the Haversine distance between the geographical centroid of the FUA and the centroid of the nearest capital and captures the geographical and hierarchical relations between cities. The degree of polycentrism is computed as the share of people living outside the urban core. The original monocentric theoretical construct of the AMM model does not trustfully reproduce the reality of modern cities, especially large urban agglomerations, where secondary poles have emerged alongside the spatial expansion of the main urban cores. The share of people living outside the main core is expected to account for the dispersion that likely generates a larger urban development ceteris paribus. In addition to the share of people living outside the core, we considered an alternative measure of polycentrism proposed by Meijers et al. (2018), which takes into account the spatial distribution and geographical concentration of people in a more rigorous manner compared to the dualism between core and non-core. However, considering the nature of the spatial unit of observation in this work, the FUA, the measure shows two drawbacks that eventually led us not to include the variable in the model. The first is that being a Helfindhal-type index measure, an artificially high number of FUAs with little municipalities show the highest values of spatial concentration and are, hence, considered strongly monocentric, even though the core is made by a single municipality. The second is that the measure is strongly correlated with the governance fragmentation, causing collinearity issues in the regression model.
The dependent variable is constructed using CLC data from the year 2012, and all the explanatory variables that vary with time refer to the same year or the preceding year when 2012 is not available. Only the POP variable makes an exception, which is measured in years 2009–2010, to avoid any simultaneity bias that may occur because the availability of houses that follows an urban spatial expansion may attract new households through the mechanism of lower average prices, reverting the hypothesised causality between population and the spatial extent of cities. Simultaneity bias is not expected to affect the other variables of the model, which are either measured as regional averages, as with income and land price variables and, hence, not affected by the single city land-use dynamics, or, as in the case of transport costs related to culture, lifestyles and network infrastructures, which may be considered strictly exogenous.

In the empirical section of the paper, we test the robustness of our result using different specifications and, among others, alternative dependent variables (summary tables with results are in Appendix C). The choice of the total urbanised area is dictated by the use of the Brueckner and Fansler (1983) empirical model, but administrative fragmentation may also impact other aspects of the urban development, in particular, those related to the spatial distribution of built-up areas. We define three measures of urban sprawl, both derived from land cover data and commonly used to describe the morphological characteristics of urbanisation (Schwarz, 2010). The first is the number of patches of artificial area (1 in the CLC classification) over the total area or the inverse of the average patch size (IAPS). The indicator represents the morphological fragmentation of the city caused by alternating urban areas and agricultural and natural spaces. The second (DISC) is the percentage of discontinuous urban fabric (112 in the CLC classification) on the total residential area (11 in the CLC classification). Urban fabric is considered discontinuous when less than 80% of the space of its cells is occupied by impermeable features (buildings, roads and artificially surfaced areas). Accordingly, the variable measures the intensity of sprawling patterns. The third (DEV) is the percentage of artificial land (1 in the CLC classification) over the total developed and developable land (1 and 2 in the CLC classification). All

<table>
<thead>
<tr>
<th>Variable</th>
<th>Variable description</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>UA</td>
<td>Total artificial area (CLC 1), km² - Urban Atlas 2012</td>
<td>239.513</td>
<td>315.280</td>
</tr>
<tr>
<td>POP</td>
<td>Total population (1,000 inhabitants) - Urban Audit</td>
<td>549.556</td>
<td>1100.558</td>
</tr>
<tr>
<td>INC</td>
<td>Gross domestic product per inhabitant, 100 euro - Urban Audit</td>
<td>266.332</td>
<td>122.336</td>
</tr>
<tr>
<td>FRAG</td>
<td>Number of FUA municipalities per 1,000 inhabitants – authors’ calculation based on GISCO official boundary files</td>
<td>0.340</td>
<td>0.359</td>
</tr>
<tr>
<td>TC</td>
<td>Inverse of the number of private cars per inhabitant – Urban Audit</td>
<td>0.274</td>
<td>0.115</td>
</tr>
<tr>
<td>ALV</td>
<td>Agriculture value-added per employee, 100 euro - Eurostat (NUTS III)</td>
<td>62.872</td>
<td>142.638</td>
</tr>
<tr>
<td>POLY</td>
<td>Percentage of people living outside the core city</td>
<td>0.434</td>
<td>0.181</td>
</tr>
<tr>
<td>InDIST</td>
<td>log of the Haversine distance from the FUA centroid to the nearest capital centroid, km – authors’ calculation based on GISCO official boundary files</td>
<td>7.047</td>
<td>1.569</td>
</tr>
<tr>
<td>IAPS</td>
<td>Number of artificial patches over total artificial area in the FUA</td>
<td>73.125</td>
<td>22.568</td>
</tr>
<tr>
<td>DISC</td>
<td>Percentage of discontinuous urban fabric</td>
<td>0.792</td>
<td>0.111</td>
</tr>
<tr>
<td>DEV</td>
<td>Percentage of developed land on total developed and developable land</td>
<td>0.738</td>
<td>0.157</td>
</tr>
</tbody>
</table>
the agricultural land, including arable land, land for permanent crops, land for pastures, and other land is considered developable, although we know that the development of these different types of land also has different costs. The measure expresses the degree of urbanisation in relation to its possibility frontier. By construction, a high value of each measure indicates a wider urban sprawl, and we expect to find a positive relationship with the administrative fragmentation accordingly. The three-level CLC classification is described in detail in Appendix B.
We borrow the empirical specification from Brueckner and Fansler (1983), who linked the actual spatial size of cities to its determinants described in the AMM framework. The model takes the form in equation 1:

\[
UA = \beta_0 + \beta_1 POP + \beta_2 INC + \beta_3 TC + \beta_4 AGRI + \varepsilon
\]

where \(UA\) is the vector of the spatial extent of urbanised areas, while the regressors \(POP\), \(INC\), \(TC\) and \(AGRI\) stand for the total population, average household income, transport costs and agricultural land values, respectively. The assumption of the monocentric model being valid, we expect \(\beta_1 > 0; \beta_2 > 0; \beta_3 < 0; \beta_4 < 0\).

We augment the model in equation 1 including the \(FRAG\) variable to capture the effect of administrative fragmentation on land use. The augmented model, in equation 2, also controls for the (log of the) distance from the capital (\(lnDist\)) and the degree of polycentrism (\(POLY\)).

\[
UA = \beta_0 + \beta_1 POP + \beta_2 INC + \beta_3 TC + \beta_4 ALV + \beta_5 FRAG + \beta_6 POLY + \beta_7 ln DIST + \varepsilon
\]

We look at the estimated \(\beta_5\) to test the hypothesis that more fragmented metropolitan areas consume more land on average and other things being equal. We expect this coefficient to be positive and statistically significant. About the relationship with control variables, we expect the coefficients of \(lnDIST\) and \(POLY\) to be lower and greater than zero, respectively. In the first case, we expect a negative relationship because capital cities are on average larger than other cities and because their growth certainly influences neighbouring cities but closer cities more. In the second case, we expect a positive relationship because the average density of the population tends to be lower in cities where more people live outside the core. Based on the discussion of existing literature in the second section it is reasonable to believe that all the effects that shape the relationship between administrative fragmentation and land use are relevant only for large metropolitan areas. Firstly, in small cities, people concentrate mainly on the largest municipality and only to a lower extent in adjacent ones. Thus, the competition to attract residents in the suburbs is lower. Secondly, the self-selection of residents in specific areas operates at the neighbourhood rather than at the municipality level in smaller cities, and this neutralises the effects of zoning on land use. Finally, spillover effects are expected to be less relevant. In small urban areas, land is not as scarce a resource as in large metropolitan areas,
and each municipality has little incentive to implement strict land-use regulations. It follows that while the prescription of the monocentric model is valid for both small and large urban areas, it may be valid in a different manner. The effect of administrative fragmentation may differ among cities of different sizes. In accounting for these different effects, there are no prior indications to suggest what the critical dimension of the FUA should be to see changes in the impact. Thus, the standard approach is to classify cities in different size groups, defining arbitrary size thresholds based on common sense and consolidated and commonly accepted values.

The threshold regression approach we propose departs from this logic and considers the threshold parameters endogenous, which means it has to be estimated alongside the other parameters from the data. The gain of using such an approach is twofold. The first and most obvious is to waive the arbitrariness of the threshold, defined using common sense and accepted standards. For the city size hierarchy, for instance, 5 million, 1 million, 500 thousand, and 100 thousand are quite standard thresholds, but may not give meaningful insights when applied to the specific object of governance fragmentation. The second is to pick only relevant thresholds. The approach, in fact, is used to estimate the threshold value from the data, but applied recursively and using inference can inform about the number of thresholds to be used, avoiding an excessive, and perhaps useless, sub-sampling of the dataset.

To illustrate the general approach, let X be the matrix of variables whose effects are expected to vary across cities of different sizes and Z the matrix of variables whose effects are expected to be constant across the sample of cities. Having defined a threshold variable (TV) for the size of cities (the total population in our case), the specification of the coefficient-varying model can be formalised as in equation 3.

\[
UA = \theta_1 X_1 \cdot I(TV \geq \gamma) + \theta_2 X_2 \cdot I(TV < \gamma)
\]

The equation allows different parameter estimates, \( \theta_1 \) and \( \theta_2 \), for the two groups of observations, respectively. One shows a value of TV larger than or at least equal to the threshold value, and the other does not. The groups are identified using the indicator function \( I(.) \).

To estimate the parameter \( \gamma \) jointly with the parameters set \( \theta_1 \) and \( \theta_2 \), we rely on the threshold regression approach (Hansen, 1996, 2000), whose procedure we briefly describe here.

1. For each value of \( \gamma \in TV \), estimate the model

\[
UA = \theta_1 X_1 \cdot I(TV \geq \gamma) + \delta' Z + \epsilon
\]

that is equivalent to estimating the model in equation 3, being \( \theta_1 \) and \( \theta + \lambda \) and \( \theta_2 = \theta \); to operationalise this step, we trim the distribution, cutting the 5\% of the values lying in the two tails to avoid generating too small sub-samples;

After estimating each model, we collect the
concentrated sum of squared errors $S(\theta, \gamma, \lambda)$ and the estimator $\hat{\gamma}_1$ of the first threshold $\gamma_1$ is the value of $\gamma$ that minimises $S(\gamma) = S(\theta, \gamma, \lambda)$.

3. To obtain a confidence interval for $\gamma$, we construct the LR statistic

$$LR = \frac{n S(\gamma) - S(\gamma_1)}{S(\gamma_1)}$$

$n$ being the sample size, and compare the statistic with the 95% confidence interval value of 7.35 (Hansen, 2000) by line-plotting the values of $LR(\gamma)$ against the values of $\gamma$ and adding a horizontal line at 7.35: the confidence interval is identified as the interval in which $LR(\gamma) < 7.35$.

In the specification of the threshold model, we consider that only three variables with an impact on urbanisation conditioned on size, FRAG, POP, and INC. For population and income, there is previous empirical evidence showing size-related effects (Paulsen, 2012). Concerning administrative fragmentation, the review of existing theoretical and empirical literature in the second section supports this hypothesis. The remaining variables are not expected to show an impact conditioned on size. However, since this a discreptional choice, its validity is tested empirically (results in the appendix) also to assess the robustness of the results.
Table 3 shows the results of the regression models with and without the threshold effect. The first set of results concerns the model without the threshold, estimated on the full sample of 359 European FUAs for which all the data were available. Consistent with all the previous literature, the estimated coefficients for population and income were both positive and statistically significant. An increase of 1,000 inhabitants, other things being equal, results in a rise in the artificial area by 0.24 km². This amount is lower by a magnitude of one-third compared to the estimate in Paulsen (2012), and we attribute this to the structural differences between the EU and the US, with EU cities traditionally more compact (Schneider and Woodcock, 2008). The estimate is larger compared to those in Oueslati et al. (2015), who provided an elasticity estimate of 0.17. Using the sample averages reported in their table, POP=939.8 (in thousands of inhabitants) and UA=211.41 (in km²), the elasticity translates into a marginal consumption of land per 1,000 inhabitants six times smaller compared to ours. We exclude that the difference in the estimated coefficient is due to the inclusion of additional variables in our model, as we estimate the linear regression also using the Brueckner and Fansler (1983) variables only, and the estimated marginal land consumption increases from just 0.24 to 0.25. Instead, this difference is a result of the sampling, because we include in our sample all the FUAs with 50,000 to 100,000 inhabitants and existing empirical evidence (Guastella et al., 2017) demonstrates that small cities have, on average, higher marginal land consumptions.

An increase in the average GDP by 100 euros produces approximately half the effect on the total artificial area of a population increase by 1,000 inhabitants. Although the income effect may appear surprisingly high, it is worth noting that it is coherent with the estimates provided in previous studies, including Paulsen (2012). In addition, the high value of this coefficient may well represent the higher land consumption for non-residential purposes induced by a high GDP level. The estimated transport costs coefficient is negative, consistent with previous theoretical and empirical literature and statistically different from zero. The coefficient related to agricultural land values is correctly sloped but not statistically significant. The POLY coefficient is not statistically different from 0 and, contrary to our intuition that cities where people live outside the core are larger because the average density is lower, negatively sloped. This suggests that in larger cities, the share of people living in the core is higher. We obtain the same result (not shown) in terms of slope and significance when using an alternative measure of polycentricity that overcomes the core-periphery dualism and looks at the spatial distribution of populations, such as that used in Meijers at al. (2018). A negative and significant effect is found for the distance from the capital city, consistent with our expectations. Results about governance fragmentation support our
research hypothesis: an increase in the number of municipalities per thousand inhabitants leads to a spatial expansion of the urbanised area. This effect is estimated at 25.59 km² per additional municipality holding constant the total population and is statistically significant at a 5% level. Overall, including the main variables suggested by the monocentric model framework, an indicator of administrative fragmentation, and our controls, the covariates explain 87% of the total variance of the spatial extent of cities.

Table 3: Linear and threshold regression models for the urbanised area, EU cities, 2012

<table>
<thead>
<tr>
<th>Variable</th>
<th>No-threshold</th>
<th>Threshold</th>
<th>Threshold and country fixed effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALV</td>
<td>-0.026</td>
<td>-0.042 (0.031)</td>
<td>-0.008 (0.025)</td>
</tr>
<tr>
<td>POLY</td>
<td>-41.911 (57.554)</td>
<td>-86.914 (62.693)</td>
<td>2.296 (101.280)</td>
</tr>
<tr>
<td>InDIST</td>
<td>-22.760*** (8.672)</td>
<td>-18.723** (8.319)</td>
<td>-17.714** (8.912)</td>
</tr>
<tr>
<td>Intercept</td>
<td>295.563*** (98.151)</td>
<td>242.099** (101.104)</td>
<td>139.702 (238.464)</td>
</tr>
<tr>
<td>POP</td>
<td>0.241*** (0.016)</td>
<td>0.358*** (0.028)</td>
<td>0.394*** (0.025)</td>
</tr>
<tr>
<td>INC</td>
<td>0.141** (0.063)</td>
<td>0.086** (0.029)</td>
<td>0.076 (0.081)</td>
</tr>
<tr>
<td>FRAG</td>
<td>25.595** (12.829)</td>
<td>52.515* (12.323)</td>
<td>52.515* (12.432)</td>
</tr>
<tr>
<td>conf int</td>
<td>911.8</td>
<td>911.8</td>
<td>911.8</td>
</tr>
<tr>
<td>Adj:R²</td>
<td>0.87</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>Country dummy</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes to table: Robust SE in parenthesis. ***, **, and * indicate statistical significance at 0.1%, 1% and 5% levels

Table 3 also presents the estimation results of two threshold models, the second of which includes country dummies in the specification. In both models, the threshold size is estimated at $\gamma = 911.8$ thousands of inhabitants in the FUA and, based on the 95% confidence interval estimates, the threshold is statistically significant. As for control variables, few differences exist between the model without and the one without the threshold, as all coefficients in the threshold model keep the slope, magnitude and significance of the model without threshold effects. Instead, there are substantial differences between the threshold models when including country dummy variables. In particular, the TC coefficient turns positive and insignificant. The explanation is that country fixed effects capture, among other things, culture and lifestyle-related household preferences for transport and mobility, which also affect the urbanisation patterns of cities and cars as primary transport vehicles.

Looking at the threshold-dependent variables, relatively small differences exist between the models with and without the country dummies, but substantial differences appear when comparing the models with and
without threshold effects. Firstly, the impact of population is higher in FUAs below the threshold. This means that the marginal land consumption per new household is greater in smaller metropolitan areas. Precisely, an increase by 1,000 inhabitants in a large FUA causes an increase in the spatial extent of the urban area of 0.19 km\(^2\). The estimated figure for smaller FUAs is 0.36, 84\% larger. The estimated difference is even larger, almost double, when considering country-specific effects. Secondly, the income effect is higher in FUAs above the threshold. An increase in the average FUA income of 100 euros determines an increase in the spatial extent of the FUA of 0.086 km\(^2\) if the FUA is small and seven times higher if the FUA is large. When considering country dummy, however, both income-related coefficients turn insignificant. Finally, the effect of administrative fragmentation is much higher in large FUAs compared to small ones. In the latter group, the coefficient is also statistically significant only at the 10\% level, even though it becomes more when including country dummy in the threshold regression.

Comparing the results of the models with and without the threshold effects, the change in coefficient magnitude is worth discussing. For both population and income-related coefficients, the full sample estimate lies between the estimates based on the two sub-samples in the threshold regression. This evidence provides a clear indication that the full sample estimate somehow averages the two sub-sample estimates, which are either overestimated or underestimated. In the case of the coefficient related to governance fragmentation, the full sample estimate underestimates the sub-sample ones in both cases. This evidence provides further support for the need to account for size-related structural heterogeneity when considering the relationship between urbanisation and administrative fragmentation.

The significance of the threshold parameter and the evidence of structural heterogeneity invite a search for a second threshold. Considering the already limited number of observations in the \(TV ≥ γ_1\) subsample, we search for the second threshold on the left of \(γ_1\). Applying the same procedure to this sub-sample, we are able to estimate the second threshold \(γ_2\) = 526, but this threshold cannot be considered further: more than 90\% of the values of TV are part of the estimated confidence interval for \(γ_2\) in the range \([111; 837]\). In addition, a test on the hypothesis of structural heterogeneity within this sub-sample does not allow rejecting the null hypothesis of no difference in the coefficient estimates between the two groups.
Concluding remarks

Cities are the engines of growth and the catalysts of Europe’s economic and social development, but urbanisation and the expansion of artificial areas bring adverse environmental effects caused by land-use transformation and, simultaneously, require substantial infrastructure investments to guarantee efficient alternatives to car-based commuting. When these services are not provided, expansion translates into car-dependent, longer and more frequent commuting with even worse environmental effects. However, people demand more space to live in larger houses, and measures aimed at reducing land consumption are at risk of lowering people utility below desirable levels at best and making housing unaffordable in the worst scenario.

This paper frames into the stream of empirical literature aimed at better understanding the determinants of urbanisation and emphasises the role of governance fragmentation in combination with city size. Using EU data on FUAs, we try to respond to the question of whether more fragmented governance leads to more rapid urbanisation, other things being equal. In summary, the answer is yes. However, evidence suggests that this effect is much higher in large metropolitan areas, cities with approximately more than 1 million inhabitants, compared to relatively smaller cities. Metropolitan cities that are less fragmented from the institutional point of view enjoy significant efficiency benefits in terms of land consumption.

From the policy perspective, this result cannot be translated into an indication to reduce the number of municipalities in metropolitan cities. Instead, the problem with excessive institutional or administrative fragmentation is that it increases competition among municipalities, encourages zoning and the spreading of suburbs, and makes land-use restrictions less effective. Policy actions aimed at containing urban development should promote better coordination of the different levels of government at the metropolitan scale to curtail the planning practices that fuel urban sprawl while continuing to encourage traditional instruments for more sustainable urbanisation patterns, such as investing in public infrastructures, vehicle sharing, and electricity-based commuting.

The issue still requires better understanding, and significant aspects of the mechanism through which better governance may ensure adequate land development still need to be explored. For instance, the question remains whether the creation of a central administrative body at the FUA level is the solution to balance the excessive fragmentation, or if a simple coordination of planning policies mediated by a metropolitan agency can be effective, being the latter more feasible to implement.

Another issue unexplored in the paper but deserving attention is the connection between...
land-use planning and other policies, such as those on taxation and mobility. Coordination of fiscal policies limiting the use of land development charges to finance current expenditures may alone reduce competition for land. Likewise, coordination of transport policies aimed at lowering car-dependency may undoubtedly affect the creation of suburbs and the spread of leapfrog urban development. The increasing availability of urban-level data certainly has the potential to boost research in this direction.
Endnotes

1 In Musgrave (1939) and Samuelson (1954), consumer-voter’s preferences are given, and the central government adjusts the revenue and expenditure pattern to them. At the local level, the municipalities’ revenue and expenditure patterns are given, and the consumer-voter chooses the municipality that best fits his set of preferences.

2 In contrast, strong evidence exists of tax mimicking among neighbouring municipalities (Bocci et al., 2017).

3 CEC (1997) defines spatial planning systems as ‘the various institutional arrangements for expressing spatial planning objectives and the mechanisms employed for realising them’.

4 According to the CLC classification, the aggregate ‘1-Urban Area’ comprises all artificial surfaces, including urban fabric, industrial, commercial and transport units, mine, dump and construction sites, and artificial non-agricultural vegetated areas.

Acknowledgements

We thank the editor and two anonymous referees for their valuable criticisms and suggestions to earlier drafts of the paper.
References


## Table A1: Summary review of existing studies reporting about the positive (Benefits) and negative (Costs) effects of dispersed urban form, by outcome type

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Potential costs and benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Productivity</td>
<td><strong>COSTS:</strong> Impossibility to create agglomeration economies, firms’ networks and interactions decrease productivity and innovation levels [Neuman, 2005; Maskell &amp; Malmberg, 2007; Jones et al., 2010; OECD, 2012b] <strong>BENEFITS:</strong> Less road congestion [Burton, 2000]</td>
</tr>
<tr>
<td>Value of space</td>
<td><strong>COSTS:</strong> Both demand-side (e.g., lower productivity and services availability) and supply-side (e.g., abundance of land) factors decrease the value of usable space and land [Alexander, 1993; Churchman, 1999; Glaeser et al., 2003; Cheshire &amp; Hilber, 2008; Epple et al., 2010; Knox, 2011] <strong>BENEFITS:</strong> Increased open space and biodiversity within city limits [Burton, 2001; Neuman, 2005; Ikin et al., 2013; Wolsink, 2016]</td>
</tr>
<tr>
<td>Job accessibility</td>
<td><strong>COSTS:</strong> Longer trip length, less public transport connectivity [Burton, 2001; Boussauw, Neutens &amp; Witlox, 2012] <strong>BENEFITS:</strong> Lower number of crimes [Burchell et al., 1998; Burton, 2000; Chhetri et al., 2013]</td>
</tr>
<tr>
<td>Public services</td>
<td><strong>COSTS:</strong> Impossibility of implementing scale economies [Carruthers &amp; Ulfarsson, 2003; Matsumoto, 2011] <strong>BENEFITS:</strong> Increased open space and biodiversity within city limits [Burton, 2001; Neuman, 2005; Ikin et al., 2013; Wolsink, 2016]</td>
</tr>
<tr>
<td>Social equity</td>
<td><strong>COSTS:</strong> Less surveillance and street lighting [Farrington &amp; Welsh, 2008; Braga &amp; Weisburd, 2010; Tang, 2015] <strong>BENEFITS:</strong> Lower concentration of air pollutants in densely populated areas [Burton, 2000]</td>
</tr>
<tr>
<td>Safety</td>
<td><strong>COSTS:</strong> Higher fragmentation of habitats outside city limits; reduced soil capacity to provide ecosystem services [Cervero, 2001; Burton et al., 2003; Dieleman &amp; Wegener, 2004; Alberti, 2005; European Environment Agency and European Commission, 2006, Mcdonald et al., 2008, Turbe et al., 2010; Seto et al., 2012, Siedentop and Fina, 2012] <strong>BENEFITS:</strong> Increased open space and biodiversity within city limits [Burton, 2001; Neuman, 2005; ikin et al., 2013; Wolsink, 2016]</td>
</tr>
<tr>
<td>Air pollution</td>
<td><strong>COSTS:</strong> Higher pollution emissions [Stone, 2008; Schweitzer &amp; Zou, 2010; Echenique et al., 2012; Berrigan et al., 2014, Creutzig, 2014] <strong>BENEFITS:</strong> Lower concentration of air pollutants in densely populated areas [Burton, 2000]</td>
</tr>
<tr>
<td>Energy efficiency</td>
<td><strong>COSTS:</strong> Lower buildings and single-family houses are usually less energy efficient [Gordon &amp; Richardson, 1997; Verlinden &amp; Rooijers, 1998; Neuman, 2005; Ewing &amp; Rong, 2008; OECD, 2012] <strong>BENEFITS:</strong> Lower road congestion; lower number of traffic fatalities [Burchell et al., 1998; Ewing, Schieber &amp; Zegeer, 2003; Burton et al., 2003; Ewing, Pendall, and Chen, 2003; Dieleman &amp; Wegener, 2004; Zolnik, 2011; Holcombe &amp; Williams, 2012]</td>
</tr>
<tr>
<td>Traffic flow</td>
<td><strong>COSTS:</strong> Longer trips; increased car-dependency; reduced walkability and cyclability [Newman &amp; Kenworthy, 1996; Cervero, 2001; Glaeser and Kahn, 2004; Dieleman &amp; Wegener, 2004; Ewing, Pendall &amp; Chen, 2003; Bart, 2010; Zolnik, 2011; Holcombe &amp; Williams, 2012] <strong>BENEFITS:</strong> Increased open space and biodiversity within city limits [Burton, 2001; Neuman, 2005; ikin et al., 2013; Wolsink, 2016]</td>
</tr>
<tr>
<td>Sustainable mode choice</td>
<td><strong>COSTS:</strong> Lower share of walking and cycling because of longer average trip length; lower share of public transport because discontinuous areas are more difficult to serve [Thomas &amp; Cousins, 1996; Churchman, 1999; Burton, 2000; Echenique &amp; Saint, 2001; Neuman, 2005] <strong>BENEFITS:</strong> Positive health effects due to lower density of road traffic and number of accidents [Troy, 1996; Burton, 2000; WHO, 2011]</td>
</tr>
<tr>
<td>Health</td>
<td><strong>COSTS:</strong> Physical inactivity, obesity, heart disease, cancer prevalence [Ewing et al., 2003; Kelly-Schwartz et al., 2004; Kim et al., 2006; Cho et al., 2006; Doyle et al., 2006; Ewing, Brownson &amp; Berrigan, 2006; Plantinga &amp; Bernal, 2007; Joshu et al., 2008; Fan &amp; Song, 2009; Lee, Ewing &amp; Sesso, 2009; Kostova, 2011; Griffin et al., 2012] <strong>BENEFITS:</strong> Physical inactivity, obesity, heart disease, cancer prevalence [Ewing et al., 2003; Kelly-Schwartz et al., 2004; Kim et al., 2006; Cho et al., 2006; Doyle et al., 2006; Ewing, Brownson &amp; Berrigan, 2006; Plantinga &amp; Bernal, 2007; Joshu et al., 2008; Fan &amp; Song, 2009; Lee, Ewing &amp; Sesso, 2009; Kostova, 2011; Griffin et al., 2012]</td>
</tr>
<tr>
<td>Wellbeing</td>
<td><strong>COSTS:</strong> Limited access to services and amenities, lower consumption variety [Churchman, 1999; Burton, 2000; 2002; Bontanantini, 2013; Schiff, 2015] <strong>BENEFITS:</strong> Reflection of consumer preferences for low-density living; more domestic space (due to lower rent), higher sense of community [Wiersinga, 1997; Burchell et al., 1998; Neuman, 2005; Bruegmann, 2005]</td>
</tr>
</tbody>
</table>
## Table B1: 3 level Corine Land Cover classification

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1 Artificial surfaces</strong></td>
<td>11 Urban fabric</td>
<td>111 Continuous urban fabric</td>
</tr>
<tr>
<td></td>
<td></td>
<td>112 Discontinuous urban fabric</td>
</tr>
<tr>
<td>12 Industrial, commercial</td>
<td>121 Industrial or commercial and transport units</td>
<td></td>
</tr>
<tr>
<td></td>
<td>122 Road and rail networks and associated land</td>
<td></td>
</tr>
<tr>
<td></td>
<td>123 Port areas</td>
<td></td>
</tr>
<tr>
<td></td>
<td>124 Airports</td>
<td></td>
</tr>
<tr>
<td>13 Mine, dump and construction sites</td>
<td>131 Mineral extraction sites</td>
<td></td>
</tr>
<tr>
<td></td>
<td>132 Dump sites</td>
<td></td>
</tr>
<tr>
<td></td>
<td>133 Construction sites</td>
<td></td>
</tr>
<tr>
<td>14 Artificial, non-agricultural vegetated areas</td>
<td>141 Green urban areas</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>142 Sport and leisure facilities</td>
</tr>
<tr>
<td><strong>2 Agricultural areas</strong></td>
<td>21 Arable land</td>
<td>211 Non-irrigated arable land</td>
</tr>
<tr>
<td></td>
<td>212 Permanently irrigated land</td>
<td></td>
</tr>
<tr>
<td></td>
<td>213 Rice fields</td>
<td></td>
</tr>
<tr>
<td>22 Permanent crops</td>
<td>221 Vineyards</td>
<td></td>
</tr>
<tr>
<td></td>
<td>222 Fruit trees and berry plantations</td>
<td></td>
</tr>
<tr>
<td></td>
<td>223 Olive groves</td>
<td></td>
</tr>
<tr>
<td>23 Pastures</td>
<td>231 Pastures</td>
<td></td>
</tr>
<tr>
<td>24 Heterogeneous agricultural areas</td>
<td>241 Annual crops associated with permanent crops</td>
<td></td>
</tr>
<tr>
<td></td>
<td>242 Complex cultivation patterns</td>
<td></td>
</tr>
<tr>
<td></td>
<td>243 Land principally occupied by agriculture, with significant areas of natural vegetation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>244 Agro-forestry areas</td>
<td></td>
</tr>
<tr>
<td><strong>3 Forest and semi-natural areas</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>4 Wetlands</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>5 Water bodies</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix C - Robustness checks

We check the robustness of the result considering alternative specifications of the main empirical equation used to detect the threshold effect as well as alternative measures of urbanisation.

Table C1 summarises the result of an econometric model in which all the variables are affected by structural heterogeneity. The equivalent to the standard specification used to detect the threshold, but now all the covariates are grouped into the X matrix, and the threshold parameter is replaced with the value estimated previously:

(C1)  \[ UA = \theta_1'X_1' \cdot I(TV \geq \gamma) + \theta_2'X_2' I(TV < \gamma) + \phi M + \varepsilon \]

This model also includes country dummy variables M and thus the estimates are comparable with the threshold model with country fixed effects in the third column of table 3.

Table C1: Alternative specifications of the regression model for urbanised area, EU cities, 2012

<table>
<thead>
<tr>
<th>Main terms</th>
<th>Interaction terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>865.3089*</td>
</tr>
<tr>
<td></td>
<td>(458.5387)</td>
</tr>
<tr>
<td>POP</td>
<td>0.2002***</td>
</tr>
<tr>
<td></td>
<td>(0.0123)</td>
</tr>
<tr>
<td>INC</td>
<td>0.3534</td>
</tr>
<tr>
<td></td>
<td>(0.2953)</td>
</tr>
<tr>
<td>FRAG</td>
<td>1004.153***</td>
</tr>
<tr>
<td></td>
<td>(278.6097)</td>
</tr>
<tr>
<td>TC</td>
<td>-726.4336</td>
</tr>
<tr>
<td></td>
<td>(590.6851)</td>
</tr>
<tr>
<td>ALV</td>
<td>-0.0239</td>
</tr>
<tr>
<td></td>
<td>(0.1492)</td>
</tr>
<tr>
<td>POLY</td>
<td>-715.0769*</td>
</tr>
<tr>
<td></td>
<td>(418.3775)</td>
</tr>
<tr>
<td>lnDIST</td>
<td>-25.8334*</td>
</tr>
<tr>
<td></td>
<td>(13.3176)</td>
</tr>
<tr>
<td>F (4; 349)</td>
<td>1.82</td>
</tr>
<tr>
<td>p-value</td>
<td>[0.1237]</td>
</tr>
<tr>
<td>Country dummy</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes to table: Robust SE in parenthesis. ***, **, and * indicate statistical significance at 0.1%, 1% and 5% levels.

The first and second columns of coefficients in table C1 report the values of coefficients in the \( \theta \) and \( \lambda \) vectors of the equation A1, respectively. The coefficients of the interaction terms express the difference in the impact of the related variable on urbanisation between the two groups. Consistently with the evidence in table 3, the marginal land consumption per new household is equal to 0.20
for small FUAs and 0.41 for large FUAs. Similarly, the income coefficients are estimated at 0.35 for large FUAs and at -0.10 for small ones, and the coefficients for administrative fragmentation at 1004 and 31 respectively. In summary, the inclusion of interaction terms for the variables that were previously excluded does not change the main results. Thus, we look at the additional coefficients estimated and note that only for the POLY variable there is a statistically significant difference in the coefficients between the two groups. Note that the coefficient for the interaction term of the TC variable is significant but only at a 5% level and the coefficient on the main term is altogether insignificant.

In the lower part of the table, we report the results of an ANOVA test on the hypothesis that the terms referred to the variable TC, ALV, POLY, and lnDIST. The test is used to compare this model in table C1 with the model with country fixed effects in table 3. The value of the F-statistic does not allow rejecting the null hypothesis, and we conclude that the difference between the two models is not statistically significant and that allowing structural heterogeneity between size groups in all the variables of the model does not result in substantial changes.

Next, we perform robustness checks on the choice of the dependent variable. The use of the total built-up area in the city, the dependent variable in our models, is a standard choice in the literature and follows the reduced form specification (Brueckner and Fansler, 1983) originally presented. However, the review of the literature on the impacts of administrative fragmentation on urban development in section 2 highlighted multiple effects that may best be captured by other indicators of urban development. The three indicators we consider here, namely the inverse of the average size of urban patches, the percentage of discontinuous urban development, and the share of developed land.

For each indicator, we run simple regression and threshold regression, and in no case, a threshold is detected. The procedure ends up indicating the value of a potential threshold, but the estimated confidence interval is too large, covering almost entirely the sample of observations. Hence, we present in table C2 the linear model estimates.
Table C2: The effect of administrative fragmentation on different indicators of urban sprawl, EU cities, 2012

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>IAPS</th>
<th>DISC</th>
<th>DEV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>71.9877***</td>
<td>61.4994***</td>
<td>26.7044**</td>
</tr>
<tr>
<td></td>
<td>(18.6052)</td>
<td>(10.7799)</td>
<td>(11.8101)</td>
</tr>
<tr>
<td>POP</td>
<td>-0.0015*</td>
<td>-0.0005</td>
<td>-0.0017**</td>
</tr>
<tr>
<td></td>
<td>(0.0009)</td>
<td>(0.0004)</td>
<td>(0.0008)</td>
</tr>
<tr>
<td>INC</td>
<td>-0.0042</td>
<td>0.0014</td>
<td>-0.0034</td>
</tr>
<tr>
<td></td>
<td>(0.0097)</td>
<td>(0.0077)</td>
<td>(0.0081)</td>
</tr>
<tr>
<td>FRAG</td>
<td>7.1160**</td>
<td>6.8127***</td>
<td>19.2942***</td>
</tr>
<tr>
<td></td>
<td>(3.5557)</td>
<td>(1.6812)</td>
<td>(2.6917)</td>
</tr>
<tr>
<td>TC</td>
<td>22.0108</td>
<td>-3.1683</td>
<td>0.2072</td>
</tr>
<tr>
<td></td>
<td>(30.0834)</td>
<td>(8.5775)</td>
<td>(19.8046)</td>
</tr>
<tr>
<td>ALV</td>
<td>0.0106</td>
<td>-0.0004</td>
<td>-0.0126</td>
</tr>
<tr>
<td></td>
<td>(0.0181)</td>
<td>(0.0030)</td>
<td>(0.0106)</td>
</tr>
<tr>
<td>POLY</td>
<td>22.3968</td>
<td>0.1202</td>
<td>15.7066*</td>
</tr>
<tr>
<td></td>
<td>(14.9314)</td>
<td>(4.2736)</td>
<td>(9.5366)</td>
</tr>
<tr>
<td>lnDIST</td>
<td>0.7050</td>
<td>-0.8782***</td>
<td>-0.7505</td>
</tr>
<tr>
<td></td>
<td>(0.6526)</td>
<td>(0.3305)</td>
<td>(0.5153)</td>
</tr>
<tr>
<td>R2</td>
<td>0.36</td>
<td>0.46</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Notes to table: Robust SE in parenthesis. ***, **, and * indicate statistical significance at 0.1%, 1% and 5% levels.

In all the regressions the population coefficient is negatively sloped, insignificant when DISC is the dependent variable, significant but only at a 10% level when IAPS is the dependent variable, and significant when DEV is the dependent variable. This evidence suggests that the population is negatively (and weakly) related to urban sprawl or, put differently, sprawl occurs more likely in less populated metropolitan areas. There are several justifications for this evidence, but the most prominent is that in large metropolitan areas developable land is a scarce resource and it is used more efficiently. There is no evidence relating urban sprawl to income, but this lack of evidence is likely the result of the inclusion of country dummy variables, consistently with the evidence already discussed in table 3.

The most interesting result concerns the administrative governance, whose coefficient is always positive and statistically significant. This evidence confirms that more fragmented governance impacts not only the amount of total built-up area in FUAs but also its geographical distribution within the FUA, as measured by different sprawl indicators. All the other estimated coefficients are not significant except the distance decay coefficient when DISC is the dependent variable.

The insignificance of the coefficient of many variables points to misspecification problems in the results of the models summarised in table C2. Not surprisingly, the goodness of fit of these models is remarkably low compared to the previous ones. We are aware of this problem with omitted variables, but we are also confident that the inclusion of country-level fixed effects can effectively mitigate the bias. Also, we have included population and income information that, according to existing evidence on the determinants of urban sprawl, are the main explanatory variables which omission may cause severe estimation bias. Having included the most relevant information and having controlled for missing information with higher level effects the result about the positive relationship between administrative fragmentation and urban sprawls is deemed reliable.
References


Boussauw, K., Neutens, T., & Witlox, F. (2012). Relationship between spatial proximity and travel-to-work distance: the effect of the compact city. Regional studies, 46(6), 687-706.


1. 2020, FACTS Series, Alessandra Celani de Macedo, Nicola Cantore, Laura Barbier, Marco Matteini, Giorgia Pasqualetto, The Impact of Industrial Energy Efficiency on Economic and Social Indicators
3. 2020, 2030 Agenda Series, Marco Buso, Cesare Dosi, Michele Moretto, Do Exit Options Increase the Value-For-Money of Public-Private Partnerships?
4. 2020, FACTS Series, Ricardo Nieva, A Tragic Solution to the Collective Action Problem: Implications for Corruption, Conflict and Inequality
5. 2020, FACTS Series, Charles Fang Chin Cheng, Nicola Cantore, The Inclusive and Sustainable Development Index: a Data Envelopment Analysis Approach
6. 2020, 2030 Agenda Series, FACTS Series, Coker Eric, Cavalli Laura, Fabrizi Enrico, Guastella Gianni, Lippo Enrico, Parisi Maria Laura, Pontarollo Nicola, Rizzati Massimiliano, Varacca Alessandro, Vergalli Sergio, The Effects of Air Pollution on COVID-19 Related Mortality in Northern Italy
7. 2020, 2030 Agenda Series, Laura Cavalli, Giulia Lizzi, Port of the future: Addressing Efficiency and Sustainability at the Port of Livorno with 5G
8. 2020, FACTS Series, Federica Cappelli, Gianni Guastella, Stefano Pareglio, Institutional Fragmentation and Urbanisation in the EU Cities