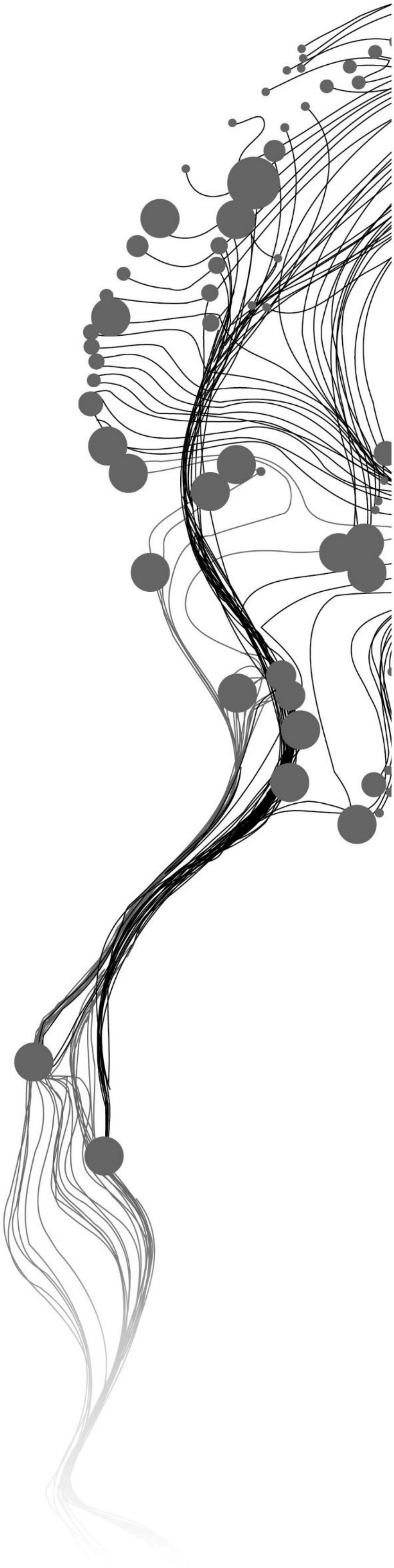
Economic diversity and complexity in the urban context

Exploring the links between urban morphology and economic performance in Belo Horizonte, Brazil

LUCAS VIEIRA MAGALHÃES
July, 2020

SUPERVISORS:
Dr. M. Kuffer
Dr. N. Schwarz
Dr. M. Haddad (External Advisor, Iowa State University)



Economic diversity and complexity in the urban context

Exploring the links between urban morphology and economic performance in Belo Horizonte, Brazil

LUCAS VIEIRA MAGALHÃES

Enschede, The Netherlands, July, 2020

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Urban Planning and Management

SUPERVISORS:
Dr. M. Kuffer
Dr. N. Schwarz
Dr. M. Haddad (External Advisor, Iowa State University)

THESIS ASSESSMENT BOARD:
Dr. J. Martinez (Chair)
Dr. H. Pearsall (External Examiner, Temple University)

ABSTRACT

The diversity of economic activities is one of the key characteristics that make cities attractive in our society. Not only a natural consequence of how cities are built, different aspects of the urban form influence the enhancement or decline of such economic diversity. Urban morphology is the study of the set of physical features a city offers to its inhabitants, such as buildings and road layout. This research explores how urban morphology influences levels of economic diversity within a city, considering it as a complex system, and how this in turn influences the economic performance achieved in different areas. Economic performance is the expected collective development that thriving economic activities bring to a society. The aspects of urban morphology analysed in this research are related to the built environment, to land-use patterns and to road layout connectivity. This is statistically compared to multiple economic diversity and complexity indices, such as Shannon's entropy, Simpson's diversity and the Economic Complexity Index. A place's economic diversity is, then, statistically analysed in relation to different proxies for economic performance, such as rates of innovation, entrepreneurship, and resilience. That is done in an intra-urban approach, comparing different scales within the city of Belo Horizonte, Brazil, and in different timeframes. This research finds statistically relevant enhancers for economic diversity and complexity in multiple urban morphology indicators, as well as proxies of economic performance significantly influenced by both economic diversity and complexity. Factors such as built-up density and a high land-use mix have a strong positive effect on both the diversity and complexity of urban economic activities. Proxies for economic performance, such as the emergence of new firms and the rate of innovation for new firms are also found to be positively influenced by higher economic diversities, while a proxy for economic resilience is found to be highly influenced by a place's economic complexity. The research concludes that a quantitative analysis for previously qualitative urban theories can confirm some assumptions in urban planning, such as that built-up densities can enhance economic development and could, thus, be stimulated by urban policies. Moreover, further research taking different urban contexts into consideration is encouraged, so findings can be further expanded.

Keywords: economic diversity; economic complexity; urban morphology; economic performance

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Monika and Nina, for guiding me through this journey of personal, professional, and academic enrichment. To my external advisor Mônica, for having encouraged me, believed in me and having worked together since UFMG. To ITC and the University of Twente, for granting me with this opportunity, especially by means of the Excellence Scholarship. Without this support, my journey in this Master's and in the Netherlands would not have been possible.

To my parents, Ana and Márcio, for having always believed in me, given me the best opportunities in life, and being so supportive. This would not be possible without your love and support. To my brother, André, for having shared with me the funniest moments in literally our whole lives, even from a distance. To my little sister Maria Paula, for making me laugh and reminding me how simple enjoying life can be. To Carol, needless to say the best companion life has given me. My path has been moulded by hers and vice-versa, we have grown to become inseparable since childhood, pursuing our futures together, in both the most hilarious and challenging steps of our lives.

To my friend Camila, for being by my side in these challenging times, the best addition this master's abroad has brought me. To the friends I've met or grown closer to in the Netherlands, especially João, Gabriel and Bruno. Thank you for being with me and making this journey more pleasant. Although recent, it feels like we've known each other since long before. To my godmother Didi, and my cousin Gabi, definitely a second mother and a sister to me. To all my uncles, aunts, cousins, and grandmothers Sônia and Vera, grandfathers Brandão and Guido (*in memoriam*). It was very challenging to be far from you in these times, but I wouldn't make it without knowing about your love and support.

To my dearest friend Marcelo, with whom I share the strongest of bonds for over a decade. To Débora, on whom I've found the most sincere and caring friendship. To Fernanda and Tatá, for building together the best stories we will tell our grandchildren. To Marina and Henrique, for being with me for the first opportunity I had to study in the Netherlands, that allowed for this second opportunity to be even more enjoyable. To all my dearest friends from Comitiva, Full Pistola, and Galerinha, I would definitely cite you one by one if I had more space. Thank you for making me feel part of something greater and worth living for.

A special acknowledgement goes to Webmapper, where I have conducted my internship. Thank you Edward, Steven, Frederieke and Niene for all the technical support and your patience and joy on sharing time and knowledge with me, especially in the initial phase of this research. My Mondays and Thursdays in Utrecht have certainly been more pleasant by your side.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Economic diversity can be considered a consequence of urban life as well as a factor for why cities thrive. There are multiple aspects of the urban morphology that might influence the emergence of highly diverse economic environments. On the other side, the diversity of economic activities is seen as one of the factors that contribute to stronger economic development. This research proposes to measure economic diversity in an intra-city scale, to explore the aspects of urban morphology that contribute to a higher economic diversity and to test the consequences of such diversity for economic performance.

1.1. Background and justification

Cities are seen as the spatial materialisation of the economic system. The very idea of urban is often mixed with the availability of the necessary conditions for economic activities to prosper (Monte-Mór, 2005). A market-based economy was characterised, according to Smith (1776), by the specialisation of labour, the process in which production is broken into more specific tasks carried on by workers. Jacobs (1970, 1985) went further on this idea, claiming that this specialisation is only possible to happen within the environment of cities. For Jacobs, the process of innovation under a market-based economy happens when new labour is added on top of old labour, where new businesses arise as breakaways from existing businesses. This is true both for the emergence of new businesses within the same industry and for when new industries arise from others, leading to completely unrelated products, which is often called innovation.

These factors lead to the idea of economic development being fairly linked to the diversification of an economy. If new businesses are emerging, it is a sign that the whole economy is performing well in a determined area. If a certain region is presented with different types of economic activities functioning well together, it indicates that a certain level of development has taken place. Further literature in urban and regional sciences confirms the link between regional economic performance and diversity of economic activities – more specifically, how economically diverse activities enhance a region's resilience (Frenken, Van Oort, & Verburg, 2007; Sprague, 2018; Xiao & Drucker, 2013). This means that these regions will better withstand external shocks: either by minimising economic losses, easing the transition towards other markets or by recovering more quickly from significant declines in economic performance (Pant, Barker, & Zobel, 2014). The idea behind it is that economically diverse places have their economies functioning more like a network, better capable of reshaping itself when affected by external factors; whereas less diverse places tend to have most activities fluctuating around central nodes, causing greater losses if these nodes are disrupted (Martin, Sunley, Gardiner, & Tyler, 2016; Modica & Reggiani, 2015).

Economic diversity, when observed at the city level, is both related to the predominance of small firms within a specific industry and to the diversity of industries in relation to the whole economy (Jacobs, 1970). There are a few examples in history of how an effervescence of economic activities led cities into times of prosperity and how the dominance of single large industries led to stagnation or even decay. One very prominent example that incorporated both of them in less than a century is the city of Detroit, in the United States, with the rise and fall of the automobile industry (Glaeser, 2011). A closer example to the context where this research is conducted would be Enschede's textile industry, which represented 64% of total jobs in the 1950s and fell to just 3.4% in the 1990s, an indication of a dependence that led to industrial and economic decline (Vissers & Dankbaar, 2013). It is therefore essential for addressing economic development in cities to understand the patterns that emerge from the location of economic activities.

The focus of economics scholars in relation to the geography of innovation is, traditionally, firms, the entrepreneurs themselves or national economies. However, Florida et al. (2017) stressed the need to focus on cities. The authors highlight the works of Jacobs and Schumpeter to defend that novelty – or the emergence of new activities on top of older – should be prioritised, rather than lowering costs of production through specialisation. Christaller (1966) relates specialisation to the demand-side: he claims that specialised activities are more specific for the consumer, and so require a larger demand-catchment area to become economically viable. Therefore, these activities would tend to appear less often in space and they are more likely to locate in central areas. This "Central Place Theory" has been extensively studied in the literature (see Chen, 2014; Hsu, 2012; Mulligan, Partridge, & Carruthers, 2012). Batty (2017) refers to specialisation not as an antonym for economic diversity, but rather as a complement. If specialisation is considered as the emergence of less common activities – thus, specialised – it can go hand-in-hand with the idea of economic diversity when economic complexity comes into play.

Hidalgo (2015) shows how specialised, complex industries depend on an amalgam of less specialised production to emerge and contribute to the overall economic performance of countries. Bustos, Gomez, Hausmann, & Hidalgo (2012) and Gomez-Lievano & Patterson-Lomba (2018) also demonstrate that in order for complex economic environments to emerge, a cluster of diverse, less complex, more prevalent activities is a requirement. Specialisation can be interpreted, then, as the process through which economic diversity feeds itself. Specialised activities depend on economically diverse, centrally located places to emerge. They also tend to increase the complexity of economic activity clusters, which in turn increases the potential for economic diversity to grow and foster further economic performance. Therefore, economic development should be considered this endogenous emergent process that leads to a more complex self-organised set of functions within the cities, instead of the mere economic expansion via simply specialising into existing industries.

Jacobs (1961) and, more recently, Cozzolino (2019) point to different factors of urban morphology as being responsible to why some areas emerge as highly adaptive, and therefore potentially economically successful, and others do not. Jacobs (1961) mentions the mixture of other primary uses (such as residences); the lower average size of blocks; diverse ages, types, and sizes of buildings; and higher population density as factors essential for successful consumer-serving enterprises to flourish. Cozzolino (2019) also lists factors responsible for neighbourhoods to be more dynamic: the small scale of design; incremental construction time of buildings – as opposed to all at the same time; proscriptive planning rules; low percentages of public open space; independent ownership systems, amongst others. However, these are taken as assumptions in most urban theories, and eventually base urban plans and regulations, although not often tested quantitatively.

Being frequently an object of regulation or direct intervention by local governments, certain morphological configurations should be fostered if they are found to have a strong influence on the emergence of diversity. Urban plans and regulations should focus on enhancing local diversities if these are expected to enhance economic performance and, by extension, quality of life for the citizens. It is, therefore, of vital importance to understand how these dimensions of urban morphology, economic diversity, and economic performance are linked.

1.2. Research problem

Cities are the engines for the diversification and complexification of the economy and, therefore, for development. Distribution of knowledge and production networks are limited spatially and their spatial distributions are unequal between countries, cities and different areas within the same city. Therefore, it is essential to identify what are the factors influencing these unequal patterns of economic diversity in space. Urban morphology is frequently mentioned in the literature as an influencing factor but is not commonly quantitatively assessed. Economic performance has different indicators, but literature tends to focus on technological innovation at firm-level or aggregated regional levels, not at an intra-urban scale. This research will explore the most appropriate methods of measuring economic diversity and complexity within cities, testing the relevance of cause-and-effect links between urban morphology aspects, economic diversity and complexity measurements and economic performance indicators.

Firstly, it is essential to build a coherent analysis of economic diversity. Many of the measurements of economic diversity mentioned in the literature are aggregated to some degree by pre-delimited boundaries, which may cause a loss of information. The aggregation is in itself a challenge, since an analysis of the unequal distribution of economic activities within cities needs intra-urban micro-clustering to be detected in diversity measurements.

Secondly, Florida et al. (2017) emphasize the necessity to redirect the focus of regional science towards cities. More specifically, the authors also praise for a refocus on the visions of authors such as Jacobs and Schumpeter in geographies of development, prioritising diversity rather than scale economies and specialisation. But other authors (Batty, 2017; Hidalgo, 2015) highlight specialisation as the emergence of specific economic activities that are not prevalent everywhere and that contribute for the complexification of the local economy. A clearer connection needs to be done between these concepts, exploring how diverse economic clusters might act as self-feeding mechanisms to foster complexification and enhance diversity itself.

Thirdly, Jacobs (1961) and Cozzolino (2019) point to aspects of the urban morphology that make urban places dynamic. This dynamism is very related to the conditions expected from the urban form to allow economically diverse environments to emerge. However, the authors themselves do not test quantitatively the validity of the defined factors on fostering a cauldron of a complex set of commercial functions. This gap is proposed to be addressed by this research, quantifying the factors related to building configuration, connectivity and land-use patterns and testing their influence on the presence of economically diverse environments.

Moreover, Florida et al. (2017) suggest further research on innovation and entrepreneurship to expand away from technological innovation per se, in order to include factors such as business processes, the service industry and occupations. The authors' call for a refocus on Jacobs's ideas can be addressed in this research by measuring economic performance using Jacobs's own definitions of innovation: the rate at which new business categories emerge on top of existing businesses. An analysis of entrepreneurship can be conducted by looking at the rate of predominance of small businesses and the rate of emergence of new firms. Together with an indicator for economic resilience, economic performance of different urban areas can be related to the level of diversity where these phenomena took place.

Overall, this research has the potential of adding to existing literature an analysis of what factors of the urban morphology contribute to a stronger economic diversity and to what extent this diversity actually translates into stronger economic performance for cities.

1.3. Research objective and questions

This research aims at measuring spatial patterns of economic diversity and complexity in the intra-urban environment of Belo Horizonte, Brazil, exploring their causes and consequences. It has as specific objectives and research questions the following:

1) Apply different measurements of economic diversity in a case study

   1. a. What are the main concepts and measurements of economic diversity available in the literature and what are their limitations?

   1. b. What spatial clustering patterns emerge from mapping economic activities in the urban fabric?

   1. c. How can diversity measurements be applied in a case study and how can their expansion towards complexity allow for understanding the economic composition of urban areas?

2) Analyse what aspects of urban morphology might cause the emergence of high economic diversity and complexity

   2. a. Which quantifiable aspects of urban morphology are most commonly used in literature?

   2. b. To what degree can these aspects of urban morphology explain the presence or absence of high economic diversity and complexity?

3) Assess whether the effects of economic diversity on economic performance can be detected and measured

   3. a. What indicators of economic performance are most commonly measured quantitatively in space?

   3. b. How related are economic diversity measurements to the indicators of economic performance?

1.4. Conceptual framework

The conceptual framework (Figure 1) shows an overview of concepts used in this research. The concepts are grouped by sub-objective. *Number 1 – Diversity analysis* shows the relations around economic diversity itself: economic diversity is detected by analysing the clustering patterns of the economic activities. This diversity will be assessed using diversity measurements, as well as experimenting in different aggregation levels. These clustering patterns are expected to be analysed in the light of the proximity analysis of activities: how likely some types of activities are to appear near others. The specialisation of activities is depicted as the process through which these clusters increase their complexity, enhancing back the diversity itself.

To explore the factors that might cause the emergence of economic diversity, *Number 2 – Causes for diversity* highlights the aspects of urban morphology to be reviewed in the literature. Three main dimensions were preliminary detected: the built environment, connectivity of public and private spaces, and land use patterns. For each one of them, a set of aspects of urban morphology were highlighted: built-up density and compactness under built environment; street network and public-private interfaces under connectivity; and mixture of primary uses under land-use patterns.

For the effects of economic diversity to economic performance, three dimensions were selected in light of preliminary reviewed literature (see *Number 3 – Effects of diversity*). Innovation can be indicated by the emergence of new business categories throughout time; entrepreneurship in terms of the emergence of new firms and the predominance of small firms; economic resilience in terms of the number of closed businesses throughout time for each area.

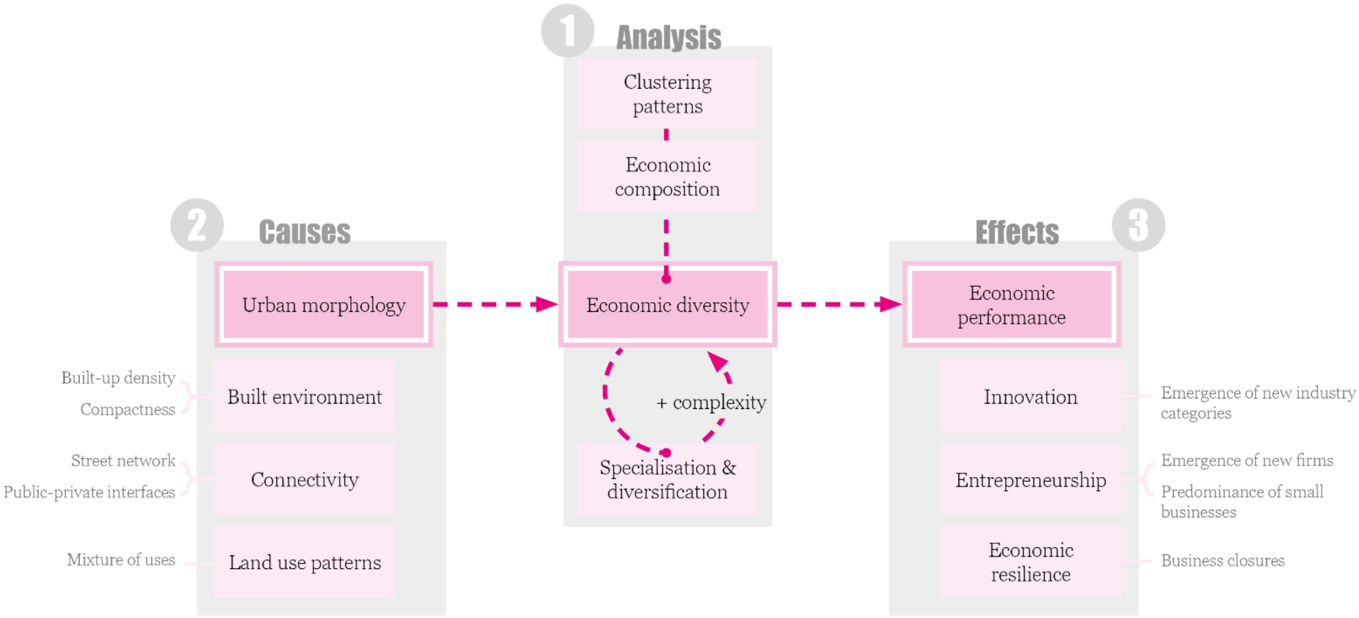


Figure 1. Conceptual framework for this research.

2. LITERATURE REVIEW

This section will systematically review previous literature on relevant topics for this research. Important contributions to be used in this research's methods are highlighted and applied in the Methodology section. Topics here described are economic clustering detection, diversity measurements, urban morphology indicators, and economic performance measurement.

2.1. Cluster detection

Several literature sources have studied the patterns emerging from the location of economic activities and how these are clustered within cities (Araldi & Fusco, 2019; Hidalgo & Castañer, 2015). Specifically to retail activities, Araldi & Fusco (2019) apply an assessment of the retail fabric in the region of the French Riviera metropolitan area. The authors identify three levels of retail cluster agglomerations, which they called Street level (150m), Neighbourhood level (300m) and Community level (600m). They argue, accordingly, that current regional science literature tends to deal with aggregated data at some sort of administrative boundaries, incurring in statistical bias due to the Modified Areal Unit Problem (MAUP) (Araldi & Fusco, 2019). Instead, they propose a bottom-up, theory-driven approach to detect clusters of retail activities based on Bayesian inference methods, that assign activity points to clusters based on their probability to be associated to a few stablished categories.

Araldi & Fusco's (2019) analysed cluster types and categories of the retail landscape; their approach considers all possibilities of combination between activity-type predominance, the morphology of clusters, and regional function. However, the authors suggested that future research in Latin America, for instance, should consider the presence and importance of shopping centres. Since this research proposes to analyse not only retail activities, but all economic activities of a city, the presence of office buildings needs to be considered as well, which was highlighted by Araldi & Fusco (2019) as the "vertical dimension". Another two aspects highlighted by them for future development are temporal comparisons (with time-series data) and urban morphology parameters, which they mention to be vital for linking retail policies to urban planning and urban design strategies.

Hidalgo & Castañer (2015) propose to create what was called "amenity space" of neighbourhoods in American cities. This involves identifying what activities tend to cluster with one another in different neighbourhoods of American cities. Their first step is to identify the neighbourhoods themselves, by first applying an Accessibility Index (1) to all economic activities in a city. The index has a distance-decaying function, being summed up to the Accessibility (A_i) of a single activity (i) all the other activities (N) accessible by it (inversely weighted by distance d in km between activity i and each other activity j). The formula is described as follows:

$$A_i = \sum_{j=1}^{N} e^{-\gamma d_{ij}} \quad (1)$$

The constant γ is set by the authors at $\gamma = 16$, which means the influence of an activity j in activity i's accessibility decreases by half roughly every 42 meters. For these settings, an activity *j* located in $d = 0$ (the same location as activity *i*) contributes to +1 in relation to A_i, whereas when $d = 0.5\ km$ the contribution is practically null. They claim that this constant setting creates meaningful neighbourhoods for their dataset. The result of applying the formula (1) to all activities in the dataset would be a landscape of Accessibility peaks and valleys, being the peaks considered first for the assignment of neighbourhoods. All activities within 500 m from a peak are automatically assigned to it as a neighbourhood. The remaining activities in the landscape are assigned to the neighbourhoods with the highest number of nearby activities, using a Nearest Neighbour analysis, in an iterative process until all activities are assigned to a specific neighbourhood.

The type of landscape derived from applying the Accessibility index resembles a topographic surface, with its peaks and valleys representing respectively concentration and absence of economic activities. A Watershed Segmentation algorithm (Bandara, 2016; Rahman & Islam, 2013) could be applied to topographic surfaces, assigning the activities organically to spatial units of analysis, in a bottom-up, theory and data-driven approach, with no pre-defined threshold. Since the effects of aggregating economic activities vary by spatial scale, the usage and testing of multiple spatial scales is encouraged (Andersson, Klaesson, & Larsson, 2016). Previous literature had already discussed the use of hierarchical clustering of economic activities in the study of economic diversity (Carol, 1960; Cutrini, 2009).

2.2. Diversity measurements

Diversity is a concept extensively explored in urban and regional economics. Urban areas are credited for promoting economic advantages for firms and workers, by generating spillover effects from the concentration of people, companies and wages (Cottineau, Finance, Hatna, Arcaute, & Batty, 2018). Kajtazi (2007) measures multifunctionality of economic activities in urban areas. The author uses the concepts of richness – total amount of economic activities; evenness – distribution of economic functions between different activities; Simpson's and Shannon-Wiener's Index (explored also by Baeza, Cerrone, & Männigo, 2017); density-diversity (explored also by Batty, Besussi, Maat, & Harts, 2004) that tries to include the total number of activities in a given area.

Jost (2006) describes that a concept of diversity can be categorised as α or β diversities. The former considers only the composition of categories within a unit of analysis, while the latter takes into account the distribution of a certain category among the units of analysis. As such, the concept of density-diversity, as worked by Batty et al. (2004), can be considered a type of β-diversity, as takes into account the distribution of economic activities among different areas. Baeza et al., (2017) calculate Shannon's entropy index using two methods: aggregating by a grid base or by street segments.

These classic diversity measurements have been further expanded towards complementing concepts, such as specialisation and complexification. As Batty (2017) discusses, the concepts of specialisation and diversification are often seen as contradictory in the study of urban economics. However, both phenomena are better understood if seen as complementary to one another (Batty, 2017; Burlina & Antonietti, 2020). Cutrini (2009) also explores economic diversity as an emergent manifestation of two parallel phenomena: specialisation of geographical units and concentration of economic activities in space. The rise in economic diversity, in a process defined by Jacobs (1970) as "import replacement", can create new comparative advantages for products an area did not produce before. This is a reason for a certain region to abandon previous products or production processes, generating specialisation. This can lead to the emergence of related services and products, that can again increase economic diversity and lead to further specialisation. This recurring diversification and specialisation shows that these two concepts do not oppose each other, but a successful economy combines them in a harmonic way (Hong & Xiao, 2016).

This process of diversification and specialisation is the bridge between urban economics and complexity science. Urban economies can be characterised as complex systems since they are composed of heterogeneous agents in a variety of groups, acting in different times and spaces, incurring in non-linear patterns and generating unexpected outcomes (Burlina & Antonietti, 2020). The recurring process of diversification and specialisation is the mean by which cities enhance the complexities of their economies and, by definition, manage to develop products and processes ever more complex. In an attempt to measure both the complexity of places and that of the products these places produce, Hidalgo & Hausmann (2009) have developed an index that was called the Economic Complexity Index.

The Economic Complexity Index is an attempt to qualify products exported by countries as more or less complex. And, by extension, to qualify the countries exporting them also as having more or less complex economies. In an iterative process, Hidalgo & Hausmann (2009) describe how to assign complexity to both products and countries, representing respectively the amount of crystallised knowledge involved in a product, and the presence of the right set of capabilities necessary for producing such products (Hidalgo, 2015). Economic Complexity levels of countries are extensively linked with levels of economic development (Hidalgo, 2015) and has been similarly applied for regional economics (Bishop & Mateos-Garcia, 2019; Burlina & Antonietti, 2020), comparing rates of growth in economic activity to the index. However, no study has been found that applies the same index for intra-urban areas, and considering urban services and retail, in extension to industries, as products to be classified. In addition to the Economic Complexity Index, this study will explore the following diversity measurements: Shannon's entropy, Simpson's diversity, the density-diversity, and richness.

2.3. Urban morphology

The urban form is believed to influence a place's economic performance in several ways. Firstly, it is worth mentioning the knowledge spillover effects consequence of urban concentrations (Cottineau et al., 2018; Forman, Goldfarb, & Greenstein, 2016). Neighbourhood variables are said to influence the degree of impact in economic innovation in urban areas (Smit, Abreu, & de Groot, 2015). Florida et al. (2017) mention how urban morphology can play a role in determining the emergence of diverse economic environments, even at smaller scales. They link "micro-clusters" to the reproduction of large-scale clustering patterns at neighbourhood scale, such as the technological clusters in San Francisco, that have emerged in areas of the city previously run-down or decadent even. Aspects of the urban fabric and the morphology in these successful areas are expected to play a role in allowing for diversity to emerge (Cozzolino, 2019).

Some of the studies found have proposed explicitly quantitative measurements of urban form, the focus for this study. Berghauser Pont & Haupt (2009) point to two essential quantitative characteristics of urban form as a basis for the characterisation of built-environment: the built-up intensity (called Floor Space Index) and the degree of compactness (called Ground Space Index). From these two indices, one can derive average floor size and Open Space Ratio, for instance (Mashhoodi & Berghauser Pont, 2011). The characteristics of plots are also emphasised as an important factor that influences the intensity of economic activities (Bobkova, Marcus, Berghauser Pont, Stavroulaki, & Bolin, 2019).

Another important dimension influencing economic intensity in cities is the mixture of primary uses. Jacobs (1961) describes that as being essential for multiple movements of people within an urban area, being important in fostering encounters and, by extension, economic development. An attempt to quantify the degree of multifunctionality of an urban area is done via the Mixed-Use Index (MXI) (Mashhoodi & Berghauser Pont, 2011; Nes, Pont, & Mashhoodi, 2012). The authors integrate built environment indicators with land-use division, indicating that lower built-up densities tend to generate mono-functionality.

Another aspect often taken into account is the public spaces, or, more specifically, the street network. Berghauser Pont & Haupt (2009) pointed to the street network as an important factor for analysing urban morphology. Other authors have quantified the interaction between public and private spaces, this being called Frontage Index (Bobkova et al., 2019; Feliciotti, Romice, & Porta, 2016, 2017). The fragmentation of the street network is also mentioned theoretically by Jacobs (1961), highlighting the importance of corners as places of encounter, and small sizes for blocks. Small block sizes also generate redundancy of street networks, or their capacity of performing repetitive roles, an important proxy for the resilience of an urban form and was quantified in the study of Feliciotti et al. (2016).

This research proposes to encompass these detected three greater dimensions of quantifiable urban form characteristics. They are the built environment itself, with its built-up densities and ground coverage indicators; the mixture of land-uses; and the connectivity of the public spaces, as the fragmentation of the street network and the permeability between private and public spaces.

2.4. Economic performance

Economic performance within cities is extensively linked to the levels of innovation and entrepreneurship (Florida, Adler, & Mellander 2017). Innovation in cities has been defined and measured in terms of the observed concentration of scientific publications and patents (Forman, Goldfarb, & Greenstein, 2016 in Florida et al., 2017), the concentration of innovative commercial products (Acs & Audretsch, 1988 in Florida et al., 2017) and network-based flows of venture capital between so-called financial centres and innovative centres (Martin, Sunley, & Turner, 2002 in Florida et al., 2017). They all conclude that innovation clustered in cities is one of the factors that contribute to higher economic performance.

As well as innovation, entrepreneurship can also be considered a phenomenon that is fostered by cities. According to Florida et al. (2017), the concept is often measured in terms of the rate of emergence of new firms, the predominance of successful small businesses, or in terms of the technology applied in new businesses that emerge. They refer to extensive literature that shows how both successful and failed entrepreneurial experiences follow clustering patterns (Folta, Cooper, & Baik, 2006; Gilbert, McDougall, & Audretsch, 2006), a further indication of the importance of locational economics to the subject. It is clear that the patterns of both success and failure of entrepreneurship are unequally distributed in space. It is vital to understand the spatial factors behind both success and failure of entrepreneurial and innovative activities.

Jacobs (1970) mentions that the most important development process that happens within a city is the act of adding new work on top of old work, which can be considered as a concept of innovation per se. While most authors tend to turn towards patent data (Balland et al., 2020), venture capital (Adler, Florida, King, & Mellander, 2019), research networks structures (Quatraro & Usai, 2017; Smit et al., 2015), or productivity (Hamidi & Zandiatashbar, 2019) as signs of economic performance, this research returns to simpler concepts, following Jacobs (1970), such as growth in number of firms (Glaeser, Kerr, & Ponzetto, 2010), and in the business categories these firms operate on (Smit et al., 2015).

In addition to that, the concept of economic resilience comes as complementary. Shutters, Muneepeerakul, & Lobo (2015) define as resilience of a complex system the capacity of its parts to perform together as a network, with nodes and links. In economic geography, the concept of resilience is defined with a multitude of meanings. Pant et al. (2014) define resilience as two things: the capacity of an economic system to not get affected by external shocks (robustness), and the speed through which an economic system is capable of recovering from a shock (recoverability). The concept of robustness within resilience completes the three major quantifiable dimensions of economic performance addressed in this research, the other two being entrepreneurship (or emergence of firms) and innovation (or emergence of business categories).

3. METHODOLOGY

This section describes the study area for the research, the dataset availability and how they were used. After that, it describes the methods for detecting clusters of economic activities, analysing aspects of the clusters' forms, performing measurements on economic diversity and complexity, and analysing economic composition derived from complexity. Thereafter, it describes which indicators were selected and how they were calculated for urban morphology and economic performance. These are summarised in Figure 2. Finally, it explains how all these were included in statistical analyses to test relevant cause-and-effect patterns between them. Figure 2 shows an overview of methods and their respective results as a flowchart.

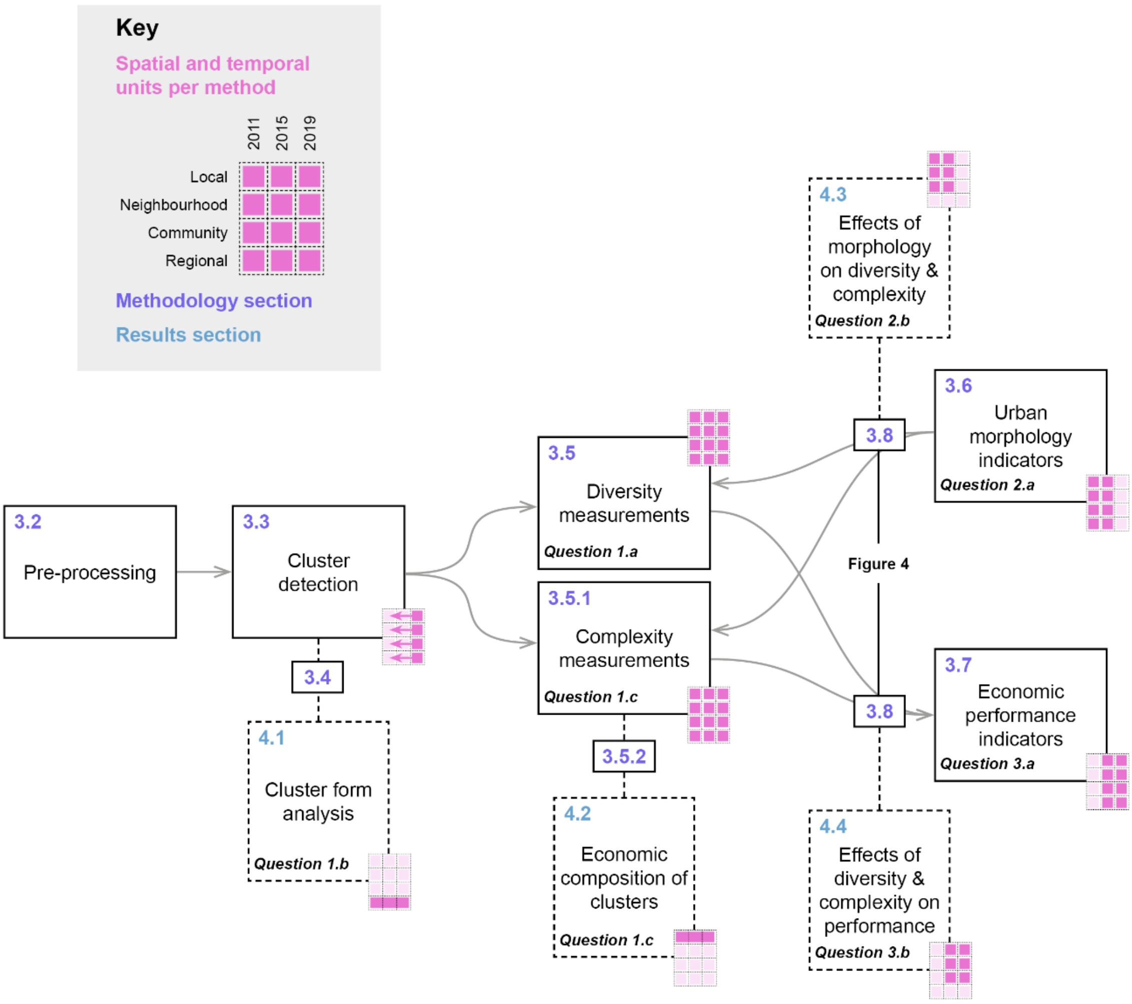


Figure 2. Flowchart showing an overview of methods and generated results.

3.1. Study area

In order to test the applicability of all the measurements of economic diversity, the city of Belo Horizonte, in Brazil was selected as a case study. The city is located in the third largest Metropolitan Area of the country, accounting to more than 5 million inhabitants. The city is in the heart of the country's iron ore mining region, called the Iron Quadrangle. Although the city's dynamics have expanded beyond the dependence from mining wealth, the state's economy has been severely affected by the recent Brazilian economic crisis, as well as by the decade-long worldwide drop in commodity prices.

Brazil has a generous policy in relation to open data, being the Access to Information Act (Government of Brazil, 2011) a milestone in relation to providing free information to its citizens. Besides that, the Municipal Government of Belo Horizonte (*PBH* in the Portuguese acronym) offers a database server with multiple spatial datasets. One essential for this research which is a Vector – Point dataset with all economic activities registered in the municipality. Each point refers to a single economic activity, characterised by its official names, addresses, area used within a building and the type of activity. The categorisation of the type of activities follows a nationally standardised coding system which makes it possible to refer back to the categories and subcategories of each activity.

The fact that the city is the author's home city also plays a role in the decision. Since different areas of the city are to be compared in this research, empirical knowledge of fundamental characteristics, historical backgrounds and functional roles of these areas to the general functioning of the city is quintessential. The author has previously worked with urban analysis in a specific area of this city, having some background research been done already in relation to the urban functional roles that take place. For this previous work, it was necessary to solicit access to some extra datasets from the Municipal Government (*PBH*), being the channels of communication for such requests already known in case they were needed.

3.2. Data definition

3.2.1. Datasets

The dataset for the location of economic activities is available for the years of 2011, 2015 and 2019. Therefore, a temporal comparison is possible throughout these 8 years. Other datasets were included in this research in order to address all questions proposed, specifically related to the built environment and possible demographic figures. Table 1 below shows an overview of all datasets selected.

Table 1. Overview of datasets used in this research.

Object	Type	Spatial level	Acquired date	Information included	Source
Economic activities	Vector – point data	Firm (328,494 entries)	2011 2015 2019	Category of main activity (6-digit level based on CNAE 2.0) Date of opening Ownership system (limited society, open society, public, private) Size of company (individual owner, small, medium, large) Type of facility (branch, franchise, headquarters) Type of practise (fixed establishment, ambulatory, door-to-door, online) Area used (m²) Address	*PBH* WFS

Object	Type	Spatial level	Acquired date	Information included	Source
Plots	Vector – polygon data	Plot (358,209 entries)	2011 2019	Boundaries Area ID	*PBH* WFS
Roads	Vector – line data	Streets (51,306 entries)	2019	Axial polylines Street name Width	*PBH* WFS
Estates	Vector – polygon data	Plot (679,085 entries)	2011	Address Plot ID Area of individual unit (m²) % of unit area from total lot built-up area Use of unit (residential, commercial – retail, commercial – offices, industrial, storage, garage)	*PBH Department of Planning*
Buildings	Vector – polygon data	Buildings (730,185 entries)	2015	Summary of statistics (mean, median, mode, standard deviation) for Digital Elevation Model and Digital Terrain Model within boundaries of polygons.	*PBH* WFS
Approved projects	Vector – Polygon data	Plot (42,151 entries)	2015	Built-up area Year of approval Dedicated floor space usage	*PBH Department of Planning*
Demography	Vector – polygon data	Census tracts (3,936 entries)	2010	Population and households, divided by age groups, income brackets, access to sanitation, education levels, and others.	*IBGE (Brazilian Institute for Geography and Statistics)*

3.2.2. Data pre-processing

A thorough inspection of the mapped points, did not detect any mistakenly duplicate entries, mispositioned points or other similar inconsistencies. This indicates that the process of maintaining and updating the dataset (which is kept up to date in the Municipality's Geodata server) is a meticulous process, most likely linked to Brazilian Federal Revenue Office's registries since the legal entities' registration numbers (*CNPJ*, in the Portuguese acronym) are present in the dataset.

However, some degree of filtering is necessary for the purpose of the research. The first criterion used was to filter out the legal entity of the "Residential Building Condominium" (*Condomínio Edilício* in Portuguese). These are special legal entities created upon completion of apartment buildings' constructions and later maintained by the dwellers themselves. Since they act as a residents' association and do not offer any economic function per se – acting as an economic entity only while hiring external services – it was decided that they are not relevant for this research. In fact, their presence in the analysis would distort the results since most residential buildings in the city are registered as such an association, generating a faux presence of economic activities even in strictly residential areas.

Two other fields were considered for the purpose of filtering the dataset. These were the Type of Practise performed by the registered economic activity and the Type of Facility within which the activity is registered.

These are split into multiple categories, explained in Appendix 01 (Table 11 and Table 12). Three conditional criteria were chosen in order to define whether an economic activity point would be considered in the analysis, according to these two fields: (1) whether economic activity's work is conducted within the registered and depicted address; (2) whether transaction or provider-customer contact is conducted at the depicted location; and (3) whether a location is characterised mainly by the presence of people (employers, employees, clients or others) instead of machinery or goods. A rationale of the chosen criteria is described in Table 2.

Table 2. Criteria for including Type of Practise or Type of Facility's categories in the analyses.

Criteria	**Rationale**
(1) whether economic activity's work is conducted within the registered and depicted address	Spatial effects derived from the proximity between economic activities can only be captured if the activity is performed in a certain place.
(2) whether production, transaction or provider-customer contact is conducted at the depicted location	The very definition of economic activity depends on production or transaction being conducted where it is depicted.
(3) whether a location is characterised mainly by the presence of people (employers, employees, clients or others) instead of machinery or goods	Knowledge spill-overs for externalities of firms' agglomeration are characterised mainly by the exchange of information between people, fostered by interpersonal encounters.

An overview of Type of Practise categories and Type of Facility categories and how they were judged by these three criteria is described in Appendix 01. Finally, under the field of "Type of Practise", only the category "At a fixed facility" was considered into the research. For the field "Type of Facility" the following categories were included: "Productive unit", "Headquarters" and "Collection post". All the other field categories were filtered out, which resulted in 199407 entries for the year of 2019, 167274 entries for the year of 2015 and 100738 entries for the year of 2011.

3.3. Cluster detection

The first step to detect clusters for the chosen dataset is to apply the Accessibility Index (Hidalgo & Castañer, 2015) for the filtered dataset of the most recent year (2019). The choice of detecting clusters solely for the most recent year of availability of data is related to the purpose of defining the clusters in the first place. They are conducted to become the basic spatial units for this research, defining data-driven boundaries within which measurements of economic diversity, performance and urban morphology will be made and compared with one another. In order to be able to compare different years, the chosen approach was to define the clusters for the most recent year available and apply their boundaries to the datasets for the previous two years.

Equation 2 (also described in section 2.1), is used to detect the four levels of clustering (Local, Neighbourhood, Community and Regional levels). In order to do so, the constant γ is set to vary, respectively, between the values of 32, 16, 8 and 4. For $\gamma = 32$, the influence of an activity (j) on another's (i) accessibility (A_i) decreases by half roughly every 21m, being neglectable at around 150m. For $\gamma = 16$, the influence of an activity on another's accessibility decreases by half roughly every 42m, being neglectable at around 300m. For $\gamma = 8$, the influence of an activity on another's accessibility decreases by half roughly every 84m, being neglectable at around 600m. For $\gamma = 4$, the influence of an activity on another's accessibility decreases by half every 168m, being neglectable at around 1200m.

$$A_i = \sum_{j=1}^{N} e^{-\gamma d_{ij}} \tag{2}$$

For each level of detection, the values achieved per activity are interpolated into raster files. The method chosen for the interpolation of the data points is the Inverse Distance Weighted (IDW) interpolation, more specifically using an exponential function to weigh values by distance. The IDW tends to preserve the values achieved for each point, avoiding the displacement of peaks and artificial values found in polynomial interpolation methods. Moreover, the voids between point concentrations, for this analysis, need to decay towards zero (0), since the absence of economic activity must result necessarily in a valley, which is achieved by the use of the exponential variable. This way, a vicious cycles of geographical relatedness is avoided, since the input point-values were already defined according to the proximity to other points, detecting yet another spatial correlation to extrapolate these values to the whole surface would be redundant.

The segmentation of the interpolated surfaces into regions of analysis is conducted using SAGA's region-growing algorithm of Watershed Segmentation. In the field of hydrology, watersheds are defined as the drainage basins within which water runs towards a single point. Within the algorithm's functionalities, local maxima are selected as seeds, sectioning the area using the valley lines as edges, assigning for each cell a unique peak to which it belongs. The result is a segmentation of the area into spatial units. In order to avoid oversegmentation, a peak-to-valley threshold was defined. In literature (e.g., Liu et al., 2018), commonly, the trial-and-error approach is used to define optimum threshold values. The authors used the elbow-method for thresholding, however, this approach in the given dataset did not result in the expected insights. In fact, all attempted statistics resulted in no elbow whatsoever. Therefore, a more empirical visual inspection was seen as a better suited approach.

By testing subsequent thresholds for the lowest level (150m) and comparing the resulted division of areas to an empirical functional knowledge of the different areas of the city, it was noticed that after some point the algorithm would join together areas very different from one another. This was the case, for instance, for the joining of the northern and southern banks of the Pampulha lake. They are known to be areas very different in function (the southern being more residential and the northern with higher presence of commerce), in levels of wealth (the southern being one of the richest areas of the city, while the northern has more humble housing), in urban morphology (the southern having large lots with more green areas between houses, the northern more compact). They should not – as they are not actually – be considered in the same neighbourhood, let alone in the same Local-level cluster.

A satisfying level of segmentation, leaving apart functionally different areas while still avoiding over segmentation, was being observed with thresholds close in value to the Standard Deviation value for the input raster. Despite not being able to systematically test the statistical relevance of using the Standard Deviation as threshold, it resulted in meaningful clustering for the 4 levels observed. It is recommended that further research, specially in the profound field of image segmentation algorithms, tests the statistical relevance for using this value as threshold.

The outputs for this process are other raster files, with the segments themselves, that will later on define the clusters' boundaries; and point-data for the seeds that generate these segments. These seeds are the peak-values detected in the input interpolated Accessibility Index (Hidalgo & Castañer, 2015) raster files. The seed points for Community segments are then used for extracting the Regional segment where they belong. The seed points for Neighbourhood segments are used for extracting the Community segment where they belong. Finally, the seed points for Local-level segments are used for extracting the Neighbourhood segments where they belong. These Street-level segments are then joined to the Neighbourhood seeds' table, and so on, to have an entangled cluster relationship between the four levels. These are then assigned to the economic activity points. The expected result from this entanglement is a hierarchical clustering in which a Local-level cluster will not be contained by a Community-level cluster that is different than the one that contains its Neighbourhood-level parent, and so forth.

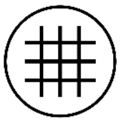
3.4. Cluster form analyses

Two steps were defined to understand the resulting spatial units of the segmentation process. Firstly, simple statistics for the composition of these areas were plotted, so that their dimensions in containing one another are visualised. Second, these spatial units were mapped within the city's boundaries. Another aspect to be detected is the location and characteristics of the peaks for these spatial units. Since they are the points where economic activities concentrate, it provides meaningful insight of how economic activities are agglomerated. For such, a classification of the peaks of the highest level of aggregation (Regional level) was conducted. Since this was a visual classification, based on the forms of the peaks detected, it would be unfeasible to be conducted for all aggregation levels.

This classification followed Lynch's (1960) characterisation of the urban landscape's elements, as perceived by citizens into five categories. These are: paths, landmarks, nodes, districts and edges. Apart from edges, which did not find any correspondence within this research, the other four categories were reinterpreted into the following: Path, Building, Node and Grid. These are described in Table 3. Two conceptual differentiations were made for these categories. Landmarks were renamed into Buildings, mostly because on Lynch's definition landmarks are impenetrable totems in the landscape, whilst in this research, Buildings are usually shopping centres where customers enter the buildings to access the economic activities. Lynch's definition of districts was renamed into Grid, since for this research the most noticed feature of this category was the street layout, following the form of an orthogonal grid.

Table 3. Morphological categorisation of cluster peaks.

Peak form	*Description*
Path 	The peak is characterised by activities clustered in a linear pattern, following the road layout of a single street. These are observed as being the main commerce streets of their own regions, similar to the idea of a high street configuration
Building 	The peak's seed falls within a single large building (or a conglomerate of small buildings), being characterised by a large number of activities within the same footprints. These are usually shopping centres, or open malls, very present in the Latin American urban landscape. They are seen as islands of commerce in the landscape and it was observed that they polarise much smaller regions than the other types.
Node 	The peak's seed falls exactly at, or very near to a large crossing of streets or avenues, usually configuring a square or roundabout. This means that, around this node, activities are concentrated in such a way that the crossing itself emerges as their natural centre. Three observations of this form were detected at the Regional level, two of them being the main centres of the city.

Peak form	*Description*
Grid 	Deviating from Lynch's (1960) classification of landscape elements, grids are characterised by a regular orthogonal road layout, within activities are clustered. Rather than being detected within a single building or a single crossing, the activities are spread within this grid, generating a two-dimensional area similar to Lynch's District definition.

3.5. Classic diversity measurements

The following diversity measurements were chosen for this research: richness index (SRI), Shannon's entropy index (SEI), Simpson's diversity index (SDI), Batty et al.'s (2004) Density-Diversity index (DDI). These were applied to each spatial unit of analysis, for all levels selected, and terms were unified according to Kajtazi (2007), Jost (2006), Batty et al. (2004).

A category iterator i ranges from 1 to the total number of categories S. An area iterator k ranges from 1 to the total number of areas K. As such, the simplest measurement, base for all calculations, is the number of individuals a belonging to category i present in area k, being represented here as $a(k,i)$. Some derived measurements are the total number of individuals for one specific category in all areas (a_i), the total number of individuals of all categories in one specific area (a_k), and the maximum number of individuals, considering all areas, for a specific category ($a_{\max(i)}$). One important derived measurement is, within an area k, the proportion between the number of individuals for a single category and the total number of individuals in that area. This was called $p_{k,i}$ and is defined in Equation 3. From this basis, all other measurements were determined (Table 4). All calculations were done using R packages *vegan* (Oksanen et al., 2019) and *EconGeo* (Balland, 2020).

$$p_{k,i} = \frac{a(k,i)}{a_k} \tag{3}$$

Table 4. Definitions for diversity measurement indexes.

Index	*Abbreviation*	*Definition*	*Equation*
Richness index	SRI	The simplest of measurements. A simple count of how many different categories exist in a given area.	$\sum_{i=1}^{S} {p_{k,i}}^0$
Shannon's entropy	SEI	Most commonly used index as a proxy for diversity. Represents, from the information theory field of study, the amount of information needed to communicate a system's state.	$-\sum_{i=1}^{S} p_{k,i} \ln p_{k,i}$
Simpson's diversity	SDI	Commonly used in measurements of biodiversity. Ranges from 0 to 1, being 0 a perfect concentration in one class and 1 an infinite diversity.	$1-\sum_{i=1}^{S} {p_{k,i}}^2$

Index	*Abbreviation*	*Definition*	*Equation*
Density-Diversity	DDI	Considers not only the mix of categories within an area but also distribution of categories between areas. Considered as an introduction of density in the studies of diversity.	$\sum_{i=1}^{S} a(k,i) / a_{\max(i)}$

3.5.1. Going beyond classic diversity measurements: the Economic Complexity Index

Recent developments in economic geography point to the complexity economics, complementary to diversity, as a more robust indicator for resilience in an area's economic activities (Bishop & Mateos-Garcia, 2019; Burlina & Antonietti, 2020). This was applied to this research having as spatial units the clusters detected and as industries all the business categories recorded within a spatial unit.

The preliminary measurement to be conducted, necessary for the calculation of an Economic Complexity Index, is the area's Balassa Index (Hausmann & Hidalgo, 2011; Hidalgo & Hausmann, 2009). This is calculated by detecting whether a region has a higher share of a certain service than average. This would mean that a region has a Revealed Comparative Advantage (RCA) in producing that product or service. Applying it to the urban economics, an area of a city that has a higher number of a certain business category than the average of all areas indicates that said area has a Revealed Comparative Advantage in that service or retail category. The calculations for the Balassa Index (BI – Equation 4) and the binary RCA (Equation 5) value (whether an area has or does not have a comparative advantage in a certain service) is defined as follows:

$$BI_{k,i} = \frac{a(k,i)}{a_i / K} \tag{4}$$

$$RCA_{k,i} = \begin{cases} 0, & BI_{k,i} < 1 \\ 1, & BI_{k,i} \geq 1 \end{cases} \tag{5}$$

By positioning the binary values for RCA in a proximity matrix with rows as spatial units and columns as business categories, it derives the diversity of an area as the number of categories in which it has a comparative advantage, and the ubiquity of a category as the number of areas with a comparative advantage in it. Hidalgo & Hausmann (2009) defined a method of reflections: the diversity of an area is updated according to its categories' ubiquities; the ubiquity of a category is updated according to their areas' diversities; the diversity is again updated according to the average diversity of the other areas with the same categories; the ubiquity of a category is again updated according to the average ubiquity of the other categories within the same areas; and so forth, in an iterative process. A satisfying level of iteration is also met by taking the eigenvector ($\vec{C}$) with the second largest eigenvalue of the proximity matrix. In the equation below, $\langle \vec{C} \rangle$ represents an average and *std*($\vec{C}$) represents a standard deviation. Economic Complexity Index (ECI – Equation 6) for an area k is then defined as follows:

$$ECI_k = \frac{\vec{C}_k - \langle \vec{C}_k \rangle}{std(\vec{C}_k)} \tag{6}$$

All these calculations are done using R Package *EconGeo* (Balland, 2017) and *EconomicComplexity* package. Other derived products from these calculations are a complexity index for business categories, being able to asses which sort of categories were classified as complex; proximity matrices for business categories, being able to relate which categories tend to appear near one another, among others.

3.5.2. Economic composition of clusters

By using proximity matrices for business categories derived from the previous item, it is possible to develop a network of categories based on the likelihood of them appearing near one another. This is defined as the Product Space (Hidalgo, Klinger, Barabasi, & Hausmann, 2007). The product space can give meaningful insights on which business categories the Economic Complexity calculations are assigning to more or less complex. And its network-like form can highlight aspects also relevant for this research, such as how central are nodes of economic activities in relation to a whole network, indicating specific economic activities that enhance local complexity or fosters diversification.

For this purpose, a network of Product Space is built for the lowest level of aggregation (Local level), this being the level with the highest number of spatial units. The resulting graph is analysed using *igraph* and characterised in relation to the business categories themselves, the betweenness of nodes in the network, the values for complexity assigned for the highest-classified business categories, and for the lowest-classified business categories. This is done by filtering business categories with more than 100 appearances in the study area.

3.6. Urban morphology

Urban morphology indicators are divided in three dimensions. Built environment dimension analyses the relations between built-up structures (buildings) and the space. Land-use analyses the composition of floor space dedicated to each major floor usage. Connectivity analyses the permeability between public and private spaces, as well as the fragmentation of the street network. All the indicators within each dimension were calculated as aggregated for all spatial units, for all spatial levels, for years 2011 and 2015.

3.6.1. Built environment

The first indicators of urban morphology are related to the built environment of each analysed area. They were defined by Berghauser Pont & Haupt (2009) as part of the Spacematrix: a set of measurements that, together, help to understand spatial patterns, general built landscape characteristics and even propose changes to the built environment. The intensity of the built landscape is called the Floor Space Index (FSI) in this method. Also known as Floor Area Ratio, it is a mathematical relation between the total floorspace area of a single plot and the plot's area, indicating then the built-up density. The second is called Ground Space Index (GSI), or coverage, and consists of the proportion of a plot's area that is covered by a building footprint. These two indices are derived into two gradient measurements: the Open Space Ratio (OSR), which consists on the difference between plot area and the building's footprint area (i.e. open areas of a plot), divided by its total floorspace area; and average number of storeys (L). Since all these indicators are interconnected, and the inputs available by the dataset are the building footprints' areas (A_f), the plots' areas (A_p) and the number of storeys per building (L), all derived calculations are explained in Table 5.

Table 5. Summary of built-environment indicators for urban morphology.

Indicator	*Abbreviation*	*Definition*	*Formula*
Floor Space Index	FSI	A measurement of occupation intensity, sometimes described as a proxy for built-up density. It is defined by the total floor area of a building divided by a plot's area.	$\frac{L * A_f}{A_p}$

Indicator	*Abbreviation*	*Definition*	*Formula*
		A higher value indicates a higher built-up intensity.	
Ground Space Index	GSI	Also called coverage ratio, is sometimes used as a proxy for compactness. It is defined by dividing a building's footprint area by the plot's area. A higher value indicates a higher compactness of the urban fabric, in terms of open spaces.	$\frac{A_f}{A_p}$
Open Space Ratio	OSR	A relation between the availability of non-built up area (open spaces) and total floor area. Higher values indicate higher availability of open spaces per built-up area.	$\frac{A_p - A_f}{A_f * L}$
Building Size Diversity	BSD	Calculated as the standard deviation for the aggregated FSI. A higher value indicates that a spatial unit has a more heterogeneous built landscape.	$std_k\left(\frac{L * A_f}{A_p}\right)$

3.6.2. Land-use

Detailed data for the amount of floorspace area dedicated for each use at plot level is available for the year of 2011. From 2011 to 2015 another dataset is used to complement the previous one, related to the approved projects for the city. The latter shows to which plots new buildings have been approved for construction, also detailing floorspace usage for those buildings. Therefore, it is possible to overwrite previous plot information with new projects' information in order to generate a floorspace use dataset for the year of 2015.

The composition of land uses for an area of analysis is defined by the proportion between Amenities (retail), Residential and Office (including factories) floorspace calculated. These compose a Mixed-Use Index (Mashhoodi & Berghauser Pont, 2011; Nes et al., 2012), represented as a ternary visualisation (Figure 3), in which each corner of the triangle indicates a monofunctional area (either predominantly Amenities, Residential or Office). The sides of the triangle indicate a bifunctionality, combining two of the three categories. The centre of the triangle indicates a mix between all three land-uses, called multifunctional.

In order to classify land-use shares into these classes, the thresholds defined by Mashhoodi & Berghauser Pont (2011) and Nes, Pont, & Mashhoodi (2012) is used. The authors use a 10% share for each land-use as a threshold. This means that, if only one land-use category has more than 10% share of floorspace area, an urban patch is classified into monofunctional. If two classes have more than 10% as bifunctional, and if all of them are presented with more than 10% as multifunctional. Table 6 shows this classification, while Figure 3 shows the ternary representation of this class division, following also other similar studies that used similar representation for other contexts (Ridd, 1995).

Table 6. Classification thresholds for land-use analysis.

Class	*% Amenities*	*% Residential*	*% Office*
Monofunctional amenities (MF-A)	More than 10%	Less than 10%	Less than 10%
Monofunctional residential (MF-R)	Less than 10%	More than 10%	Less than 10%
Monofunctional offices (MF-O)	Less than 10%	Less than 10%	More than 10%
Bifunctional amenities and residential (B-AR)	More than 10%	More than 10%	Less than 10%
Bifunctional amenities and offices (B-AO)	More than 10%	Less than 10%	More than 10%
Bifunctional residential and offices (B-RO)	Less than 10%	More than 10%	More than 10%
Multifunctional (MUL)	More than 10%	More than 10%	More than 10%

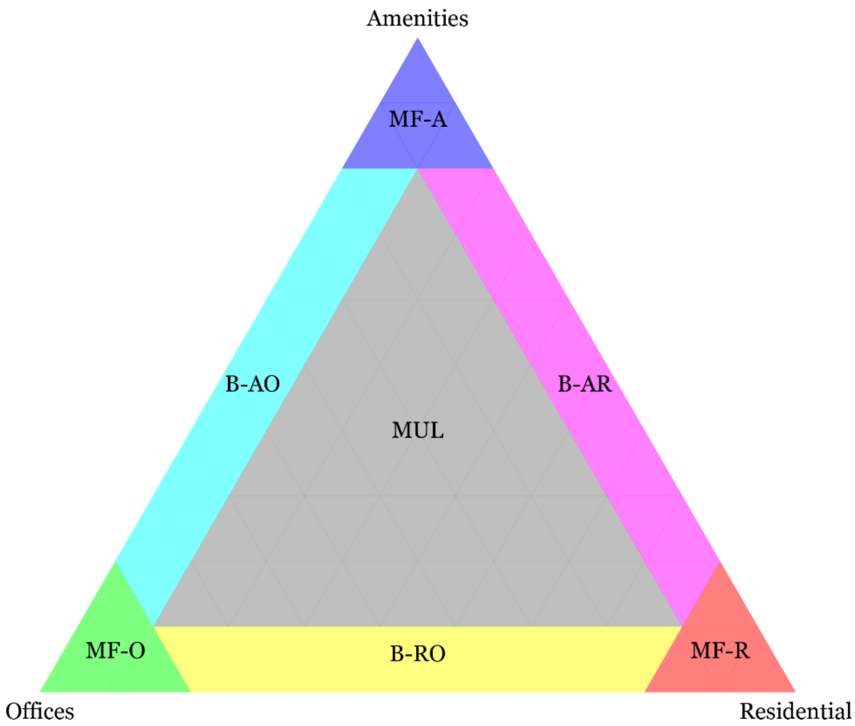


Figure 3. Ternary representation for land-use classification.

The ternary plot will be used for visually testing the influence of different land-use classifications on the values achieved for economic diversity and complexity. A **Mixed-Usedness (MIX)** indicator was also calculated by applying the Shannon entropy calculation for the composition of primary land-uses for each area and normalised. This means that a perfect division between land-uses, 1/3 for each, results in a MIX equal to 1, while a perfect concentration of one land-use results in a MIX equal to 0.

3.6.3. Connectivity

Another dimension for indicators in urban morphology is related to the public spaces and the interface between these and private spaces. Public spaces are where people in a city move around; they promote random encounters and allow potential customers to access privately managed economic spaces. These are not only determined by the fragmentation of the public space itself, but the degree of permeability between public and private areas. In order to test these effects, four indicators are chosen.

As input, the street network of the city is transformed into axial lines, with intersections being defined as a single point. This transformation is done by using the Place Syntax Tool QGIS plugin (KTH School of

Architecture & Chalmers Architecture and Spacescape AB, 2019), which is also used for other calculations in this dimension. **Intersection Density (ISD)** is a simple relation between the number of intersections per unit area.

Average Block Size (ABS) is also an indicator for the performance of the urban form. According to Jacobs (1961), very long blocks diminish the possibility of citizens to choose between different paths, harming the street network's capillarity and the possibility of different shops to be easily accessed by a multitude of people. As such, it is defined for this research as the average length of the stretches of road between two intersections within an area of analysis.

Address Fragmentation (ADF) is defined as the number of addresses per plot. It is considered a proxy for the intensity of the interface between private plots and public streets. In literature, it is common to see the Frontage Index as representing this characteristic (Bobkova et al., 2019; Feliciotti et al., 2016, 2017). However, it is common for the study area in question to have multiple shops within a plot, or multiple street frontages for a single plot, each represented by a distinct address. Therefore, a count of how many addresses fall within a plot is presented as a more accurate proxy for the intensity of this interface. For this analysis, an average is taken from all address counts in plots within each spatial unit.

Street Network Betweenness (SNB) is an indicator for redundancy of the street network (Feliciotti et al., 2016). Redundancy is characterised by the ability of street stretches to perform repetitive roles (being passed by different paths for instance). It is one of the proxies described by the authors as important for the resilience of urban form (Bobkova et al., 2019; Feliciotti et al., 2016). It is calculated using the PST QGIS plugin, using a threshold of 1200m, and the average for an area of analysis is taken for all street stretches that fall within.

3.7. Economic performance

Economic performance indicators are divided into three dimensions. The entrepreneurship dimension analyses proxies for the emergence of new economic activities. The innovation dimension includes proxies for the creation of new processes and categories of economic activities. The resilience dimension refers to how well the spatial units could withstand external shocks in terms of disappearing economic activity. All the indicators within each dimension were calculated as aggregated for all spatial units, for all spatial levels, for years 2015 and 2019 – or changes between 2011 and 2015, and between 2015 and 2019 for temporal comparisons indicators.

3.7.1. Entrepreneurship

Three indicators were chosen to represent economic performance in terms of levels of entrepreneurship in an area. The growth in the number of firms is seen as a simple indication of economic performance for an area (Maraschin & Krafta, 2013). Therefore, it was chosen as an indicator for this dimension, called **Emergence of New Firms (ENF)**. As it is a comparative indicator, it is available for years 2015 and 2019.

Employment growth and wage availability to the number of small firms, and the average diversity of firm sizes in an area are important measures of economic performance (Folta et al., 2006; Glaeser et al., 2010). Thus two indicators are used, i.e., **Company Size Diversity (CSD)** and **Predominance of Small Firms (PSF).** The dataset for firm locations has a government-defined category related to the size of the company: whether Micro-company, Small, Medium-sized or Large. These are defined according to the company's yearly earnings and the number of employees, which are not directly available at the dataset. Company Size Diversity (CSD) will be conducted by taking the standard deviation for aggregating the firm's floor space area, whereas Predominance of Small Firms (PSF) will count the percentage of companies falling under the Micro- or Small-sized company.

3.7.2. Innovation

Innovation in regional economics is usually referred to as new products or processes emerging from existing ones (Smit et al., 2015). It is often associated with new patent solicitations within an area (Folta et al., 2006) or creation of jobs (Glaeser et al., 2010). Jacobs (1970) refers to innovation as the process of adding new work on top of older work. In the context of this research, this means the emergence of new business categories from one year to another. As such, it is a comparative dimension.

A first indicator chosen for innovation is the increase in the richness of activities from one year to another. This was called **Increase in Richness (INR)** and, for every region in the analysis, it is measured as the increase in the value achieved for richness between two years. This indicator, as a comparative measurement, is valid for the year of 2015 and 2019 only.

The second indicator chosen seeks to analyse whether new activities that emerge belong to economic categories already existing in a place or whether they breakaway as new categories not previously existing there. It is seen as an intersection between the emergence of new firms (ENF) and an increase in richness (INR): it is calculated as the percentage of new firms that fall within previously non-existing categories. It was called **Companies in New Categories (CNC)** and, being also a comparative indicator, it is available for the years 2015 and 2019 only.

3.7.3. Resilience

A simple proxy for measuring the resilience is the rate at which businesses close in a certain area. If few businesses closures happen, it is an indication that this system has a high robustness and is less affected by external shocks. For the context of this research, this is calculated using the businesses' unique IDs, comparing their disappearance in consecutive years. The number of closed businesses is taken as a percentage of the total number of businesses in every area of analysis. This indicator was called **Rate of Business Closures (RBC)** and, as a comparative indicator, is available for years 2015 and 2019.

3.8. Statistical analyses

Several steps are necessary to test the influence of urban morphology indicators on economic diversity indices (research question 2. b), and economic diversity indices on economic performance indicators (research question 3. b). The first step is to calculate the descriptive statistics for all indicators and indices, for all years and spatial scales, and decide on transformations of data to be used for correlations and regressions. Secondly, correlation tables between morphology and diversity will test which indicators are individually correlated, and the same will be done between diversity and performance. Thirdly, it is tested whether indicators within the same scope are highly correlated. Highly correlated ones are excluded from regression analyses, to avoid multicollinearity problems. The fourth step is to rank the importance of indicators to be selected for regression analyses, testing whether they are significant to be included or not. And finally, the regressions themselves will show how these factors act together to influence the desired relations. Figure 4 shows an overview of the statistical analyses.

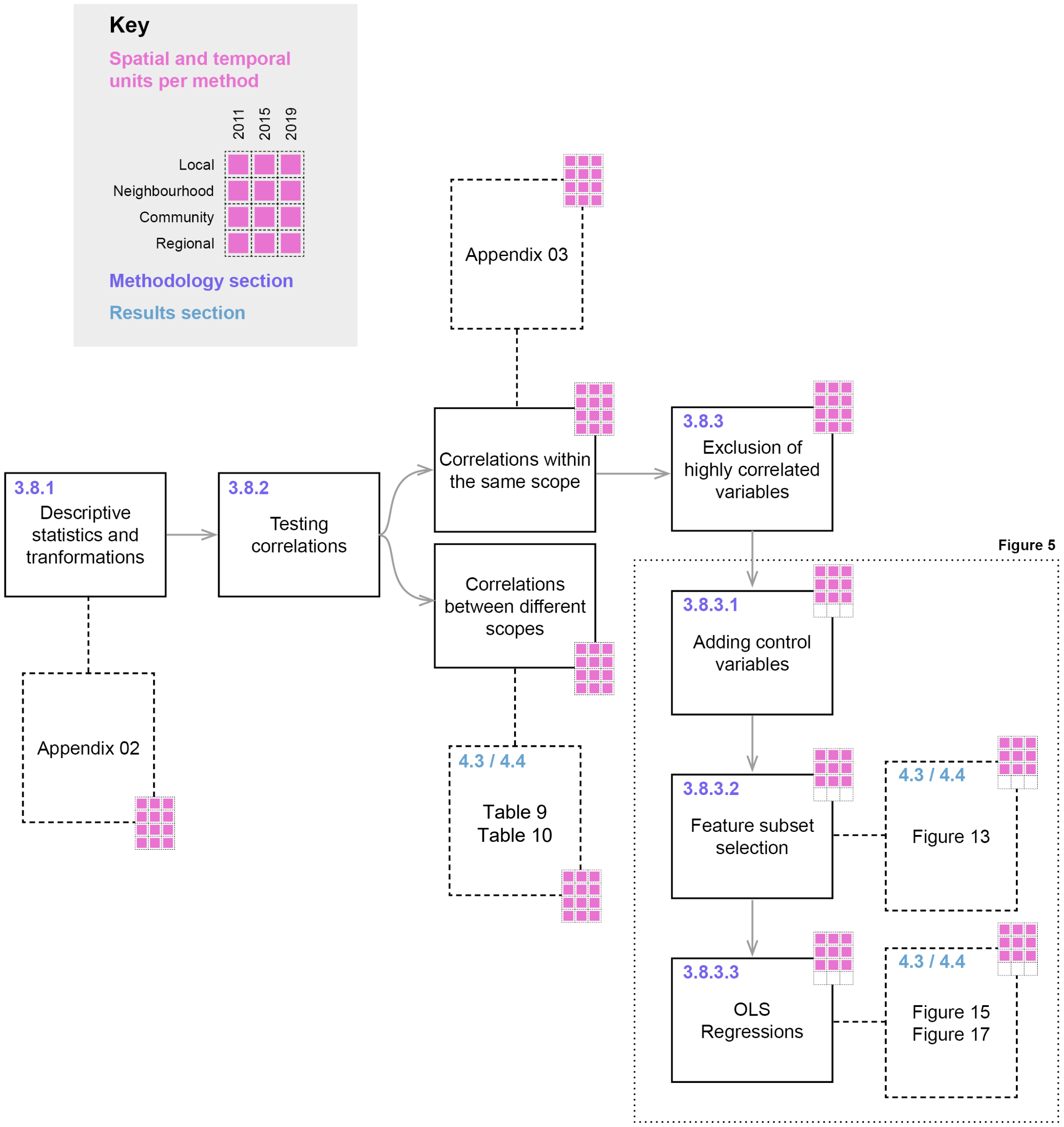


Figure 4. Flowchart with an overview of statistical analyses.

3.8.1. Descriptive statistics and transformations

The descriptive statistics for all indicators can be found in Appendix 01. Before testing their correlation, their histograms were used to transform the values to reach a normalised distribution. The transformations chose for each variable can be found in Table 7. Transformations to adequate α-diversity indices were based on variety normalisation by Jost (2006). Table 7 also includes the expected direction of influence (positive or negative) between morphology and diversity, and between diversity and performance, based on previously discussed literature.

Table 7. Summary of selected indicators and indices, transformation performed on them (f(x)) and indicators' expected relation to diversity and complexity.

Scope	*Dimension*	*Indicator*	*Abbreviation*	$f(x)$	*Expected relation to Diversity/Complexity*
Diversity	α-diversity	Richness Index	SRI	x	-
		Shannon's Entropy	SEI	exp x	-
		Simpson's Diversity	SDI	$1/(1-x)$	-
	β-diversity	Density-Diversity	DDI	$\ln(x)$	-
	Complexity	Economic Complexity Index	ECI	x	-
Urban morphology	Built environment	Floor Space Index	FSI	$\ln(x)$	Positive
		Ground Space Index	GSI	x	Positive
		Open Space Ratio	OSR	$\ln(x)$	Negative
		Building Size Diversity	BSD	$\ln(x)$	Positive
	Land-use	Mixed-Usedness	MIX	x	Positive
	Connectivity	Intersection Density	ISD	x	Positive
		Average Block Size	ABS	$\ln(x)$	Negative
		Address Fragmentation	ADF	$\ln(x)$	Positive
		Street Network Betweenness	SNB	x	Positive
Economic performance	Entrepreneurship	Emergence of New Firms	ENF	$\ln(x)$	Positive
		Company Size Diversity	CSD	$\ln(x)$	Positive
		Predominance of Small Firms	PSF	x^2	Positive
	Innovation	Increase in Richness	INR	$\ln(x)$	Positive
		Companies in New Categories	CNC	x	Positive
	Resilience	Rate of Business Closures	RBC	x	Negative

After transformations, all indicators or indices were normalised using a Z-score normalisation, and outliers over or below 3 standard deviations were removed from the next steps of analysis.

3.8.2. Testing correlations

To test the influence of urban morphology indicators on economic diversity and complexity indices, a correlation table is done including all morphology indicators with all diversity indices, for all years and all scales. Similarly, a correlation table comparing all economic diversity and complexity indices to the economic performance indicators. These together are the first steps to answer respectively research questions 2. b and 3. b. The results of these correlations can be seen in sections 4.3 and 4.4, respectively.

In order to avoid multicollinearity problems for regressions, a set of correlation tables were produced comparing indicators within each scope. All diversity and complexity indices were correlated with one another, as well as all urban morphology indicators and economic performance indicators. A preliminary result of this step is seen in Appendix 03 and it is used to choose which indicators are used for regression analyses that follows.

3.8.3. Regression definitions

Based on the previous intra-correlations, a set of preliminary results needs to be brought up so it is clear which indicators were included in statistical regressions. All α-diversity and β-diversity indices were found to be highly positively correlated with one another (above 0.7 Pearson correlation coefficients). In order to avoid redundancy and multicollinearity, only Shannon's entropy (SEI) was chosen for further regression analyses. Shannon's entropy is seen in literature as the most common proxy for diversity (Jost, 2006). Economic Complexity Index (ECI) was also included, since it was shown as being just mildly correlated with any other diversity indices. Moreover, ECI represents the expansion of diversity measurements deemed innovative for the scope of analysis this research proposes.

For urban morphology indicators, Floor Space Index (FSI) was highly correlated with Building Size Diversity (BSD). Ground Space Index (GSI) was also found to be strongly negatively correlated with Open Space Ratio (OSR). Since FSI and GSI are the bases for further analysis in literature (Berghauser Pont & Haupt, 2009), and BSD and OSR are derived or aggregated from the former, BSD and OSR were the ones left out for further regression analyses. Average Block Size (ABS) and Intersection Density (ISD) were also found to be strongly negatively correlated, being chosen for the regressions only ISD. The other indicators for urban morphology not previously mentioned were also included for regressions.

For economic performance, Increase in Richness (INR) was found to be strongly positively correlated to the Emergence of New Firms (ENF). Since there was a risk for Increase in Richness (INR) also correlate with the diversity indices, given it is derived from one of them, it was the one chosen to be removed for the regression analyses. It is also important to mention that for these regression analyses only the Local, Neighbourhood and Community levels were included, since the Regional-level aggregated 14 spatial units, that was considered too little for robust statistical results.

3.8.3.1. Control variables

A set of control variables was chosen for the regressions. These encompass other possible influences on economic diversity and complexity, beyond the urban morphology indicators. Most of the control variables chosen were demographic variables, aggregated by census tract for the Census of 2010. The disaggregation of these demographic variables into the spatial units of this research followed the instructions by Gotway & Young (2002) using probabilistic potential mapping with total built-up area by plot functioning as weights. These demographic variables chosen were average income in minimum salaries, percentage of non-white population, population density, and percentage of households with access to sewerage. Percentage of non-whites and access to sewerage are seen as proxies for both informality and poverty since racial issues in Brazil are very much related to socio-economic conditions and household conditions. Besides these, the areas of spatial units and the density of firms were also chosen as control variables. All these were also Z-score normalised.

As described in Figure 5, the regressions conducted for this section are done in a two-step process. First, all morphology and control variables chosen are put in a system of feature selection as the best subset and used as independent variables for economic diversity and complexity indices as the dependent ones. Then, diversity and complexity are used as independent variables for each of the performance variables as the dependent ones. Since the two processes are entangled, done in a subsequent process, control variables were chosen to be used for the first step alone. Their influence on the second step is considered presumed by the degree of influence of each of the independent variables individually.

3.8.3.2. Narrowing down independent variables

All of the morphology and control variables chosen are put into the feature selection that combines multiple linear regression and machine-learning algorithms. This system performs Linear Regression, Ridge, Lasso, Elastic Net, Lasso-Lars-IC and Random Forest regressions with all variables and rank their overall average importance from their resulting coefficients. Methods such as Ridge, Lasso and Elastic Net are characterised by reducing the coefficients of variables seen as colinear to the dependent variable or between themselves, resulting in smaller overall importance for that variable. This process was repeated for each of the dependent variables, for each of the scales (Local, Neighbourhood, Community) and for each of the years (2011, 2015) generating an intermediate result that is the ranking of importance of variables.

Resulting variables with average ranking higher than 0.2 (ranging from 0 to 1) were selected for the subsequent OLS regressions. For Shannon's entropy (SEI) as a dependent, this was the case for Floor Space Index (FSI), Ground Space Index (GSI), Mixed-Usedness (MIX), percentage of non-whites, population density, and area. For Economic Complexity (ECI) as a dependent, Floor Space Index (FSI), Ground Space Index (GSI), Mixed-Usedness (MIX), average income, and percentage of non-whites composed the resulting subset.

3.8.3.3. Ordinary Least Squares regressions

The final process of statistical analyses was to run multiple Ordinary Least Squares regressions. Economic diversity (SEI) and complexity (ECI) were selected separately as dependent variables for the previously described subsets of independent variables in urban morphology, for years 2011 and 2015, for scales Local, Neighbourhood and Community. Economic performance indicators were selected separately as dependent variables for SEI and ECI as independent, for years 2015 and 2019, for scales Local, Neighbourhood and Community. The β-coefficients, p-values and Adjusted R^2 values resulting from these regressions finalise the answering of research questions 2. b and 3. b. A summary for the whole process can be seen in Figure 5.

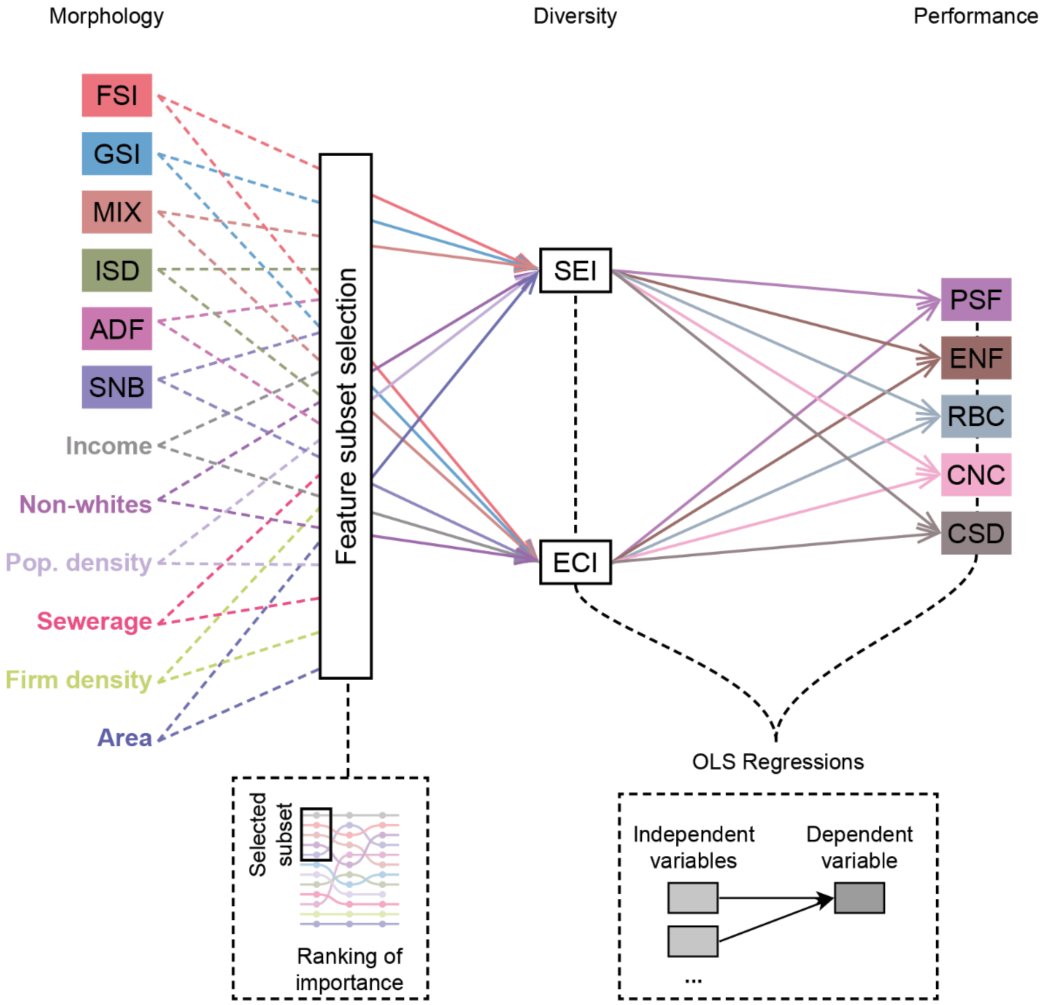


Figure 5. Summary framework for regression analyses and generated intermediate (ranking of importance) and final (OLS coefficients) results.

4. RESULTS

This section summarises the results related to the following research questions: emergent patterns of economic activities' locations (research question 1. b in sub-section 4.1); the economic compositions of these clusters (research question 1. c in sub-section 4.2); the influence of urban morphology indicators on economic diversity and complexity indices (research question 2. b in sub-section 4.3); and the influence of economic diversity and complexity indices on economic performance indicators (research question 3. b in sub-section 4.4). The research questions 1. a, 2. a, and 3. a are related to finding relevant indicators and are already described in the Literature review and Methodology sections.

4.1. Clustering patterns in the urban fabric

For generating the spatial units of this study, the city was split into areas of analysis for Local Neighbourhood, Community and Regional clustering levels. The resulting spatial units were used to analyse the economic activities of 2011 and 2015. To understand how the activities cluster in space, an overall analysis of these units' compositions were conducted. Table 8 depicts the average number of activities per spatial unit, per year and clustering level. It shows that during the period 2011-2015, there was a greater change in the total number of activities than the period 2015-2019. On average, each of the 14 Regional-level clusters contains 5.3 Community-level clusters, each containing 3.6 Neighbourhood-level clusters, each in turn containing 3.1 Local-level clusters.

Table 8. The average number of economic activities clustered for each level, for each year.

		Average number of activities per unit		
Level of clustering	**Resulting spatial units**	**2011**	**2015**	**2019**
Local-level	854	118	196	233
Neighbourhood-level	273	369	613	730
Community-level	75	1 343	2 230	2 659
Regional-level	14	7 196	11 948	14 243
	Total number of activities	100 738	167 274	199 407

In order to understand the scale of these spatial units, Figure 6 shows how these are distributed in relation to the whole city, and a localised window towards the east of the city shows how this happens in a zoomed-in level.

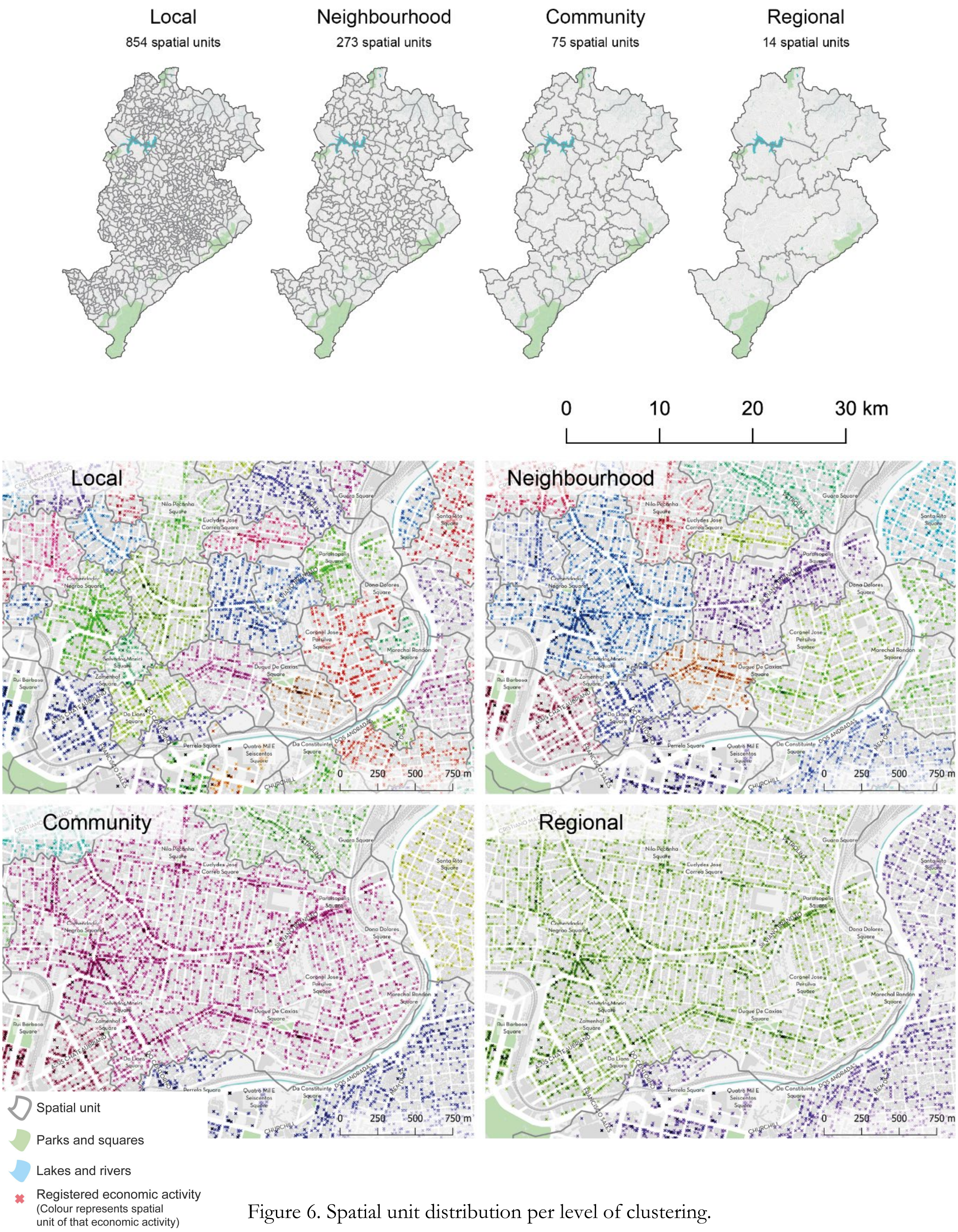


Figure 6. Spatial unit distribution per level of clustering.

By mapping the Regional-level clustering and their respective seeds (peaks of activity concentration that generated the unit) some emerging patterns were observed. Firstly, the total number of Regional-level units (14) is very close to the city's regional administrative areas (11), subdivisions of the Municipality that function as local governance areas. Some boundaries of the clusters even coincide with these administrative areas (for the Barreiro region, for instance, in the southern-most corner of the city). Moreover, it was noticed that the 14 different peaks for these areas had similarities between one another that allowed for their classification into four categories (see section 3.4), which are Building, Linear, Grid, and Node (Figure 7).

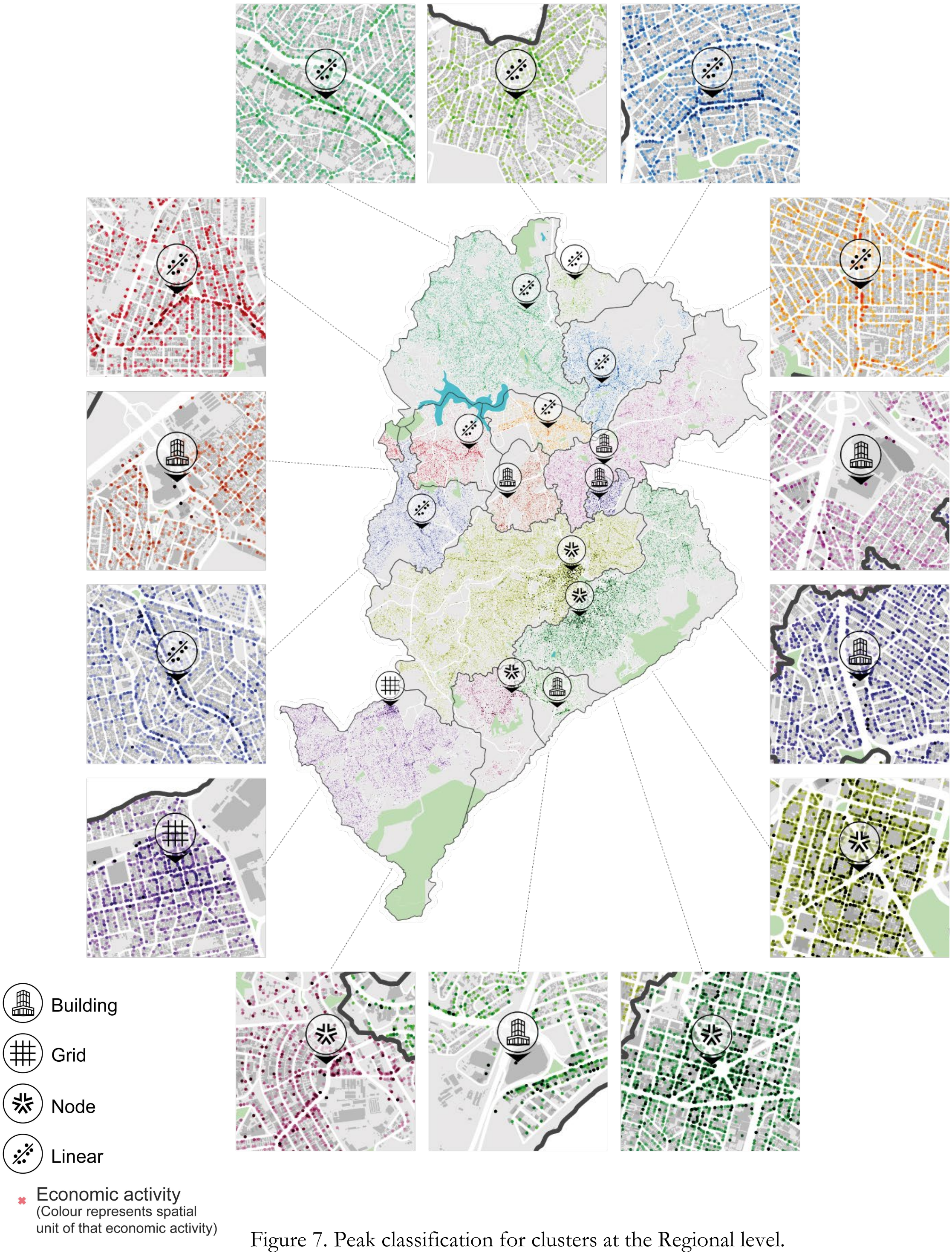


Figure 7. Peak classification for clusters at the Regional level.

Figure 7 shows that areas polarised by the Building category as a peak are usually smaller or encompasses less dense areas. In general, the northern part of the city is characterised by Linear peaks of economic concentration. The two most important economic centres of the city are polarised by Nodes (respectively, in yellow and dark green).

4.2. Economic composition of clusters

After calculating economic diversity and complexity of the spatial units, it is important to understand what the Economic Complexity Index means in terms of the categorical composition of economic activities within these areas and how it expands from classic diversity measurements. The Economic Complexity Index calculation assigns complexity values to business categories, besides the expected complexity assigned to the spatial units. This generates a proximity matrix that was used for creating a network of business categories called the Product Space. By plotting all business categories, Figure 8 shows a remarkable central core concentrating most of the economic categories observed in the network. A few branches are observed as well, being the branches located in the bottom corner of the chart, for instance, more dedicated to retail activities.

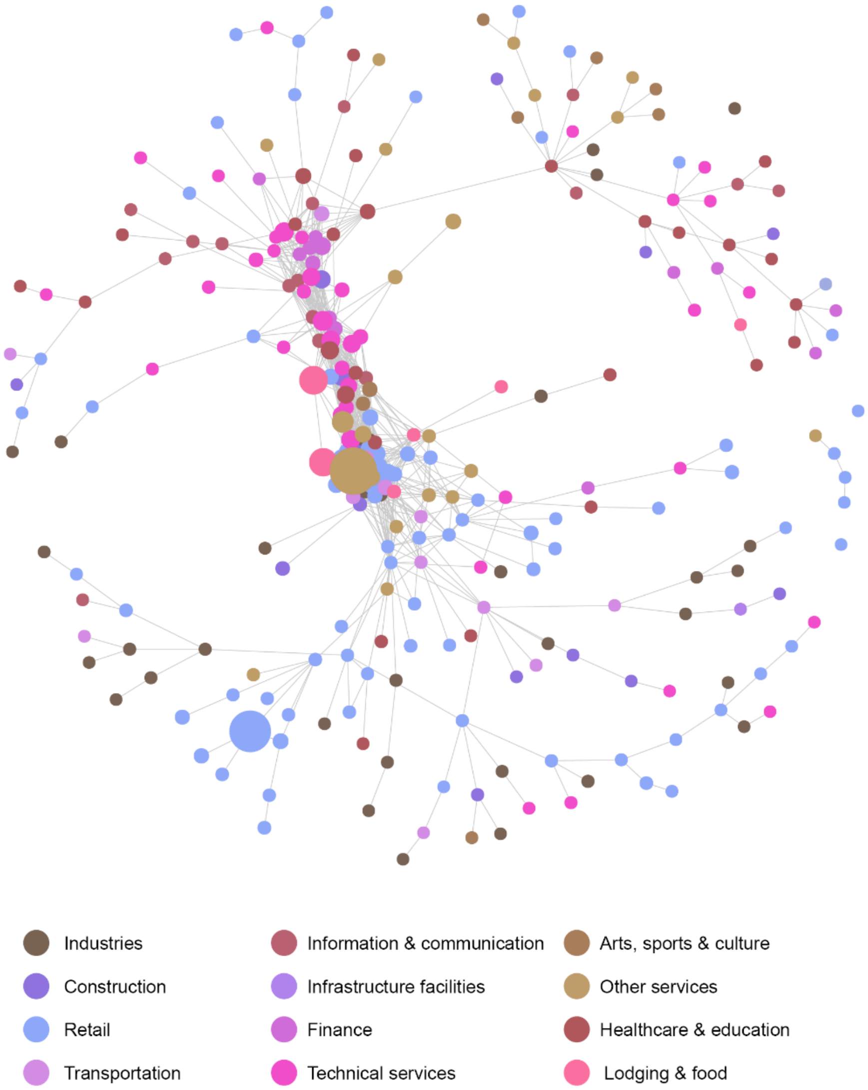


Figure 8. Product-Space network for business categories in Belo Horizonte. Size represents number of occurrences

The strength of the connection between these nodes is determined by the relevance of co-occurrence between two business categories in all spatial units. This indicates that there are categories that tend to co-occur more often than others, generating nodes being crossed by most lines. To identify this, a test for betweenness was conducted for these nodes (Figure 9).

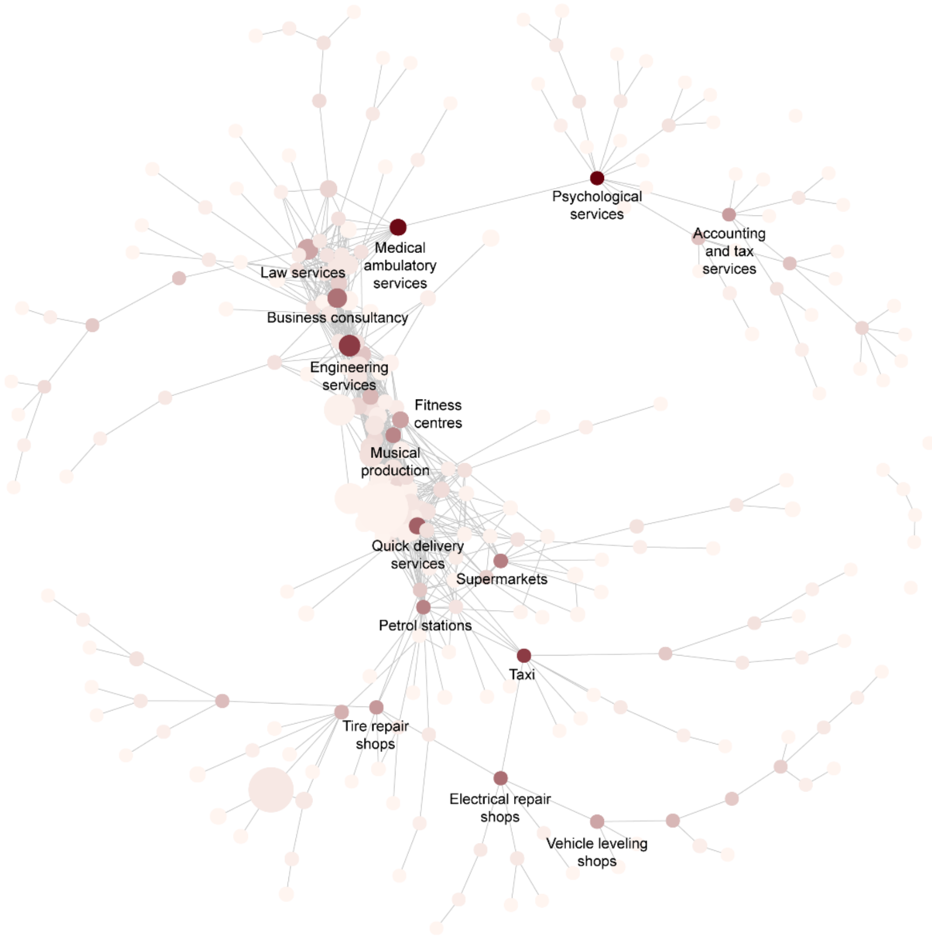


Figure 9. Business categories with highest betweenness values in the network, being more highly present together with others.

Since Economic Complexity calculations assign complexity values for both areas of analysis and products, it is deemed important to check which products are being considered complex, so that the areas they are in are also being considered as economically complex. These are expected to be business categories that require higher capabilities, higher technological requirements, or more knowledge-intensive firms. For such, it was highlighted the business categories with highest and lowest complexity values associated with them, seen in Figure 10 and Figure 11, which confirm these expectations.

Figure 10. Business categories with highest Product Complexity associated to them.

Figure 11. Business categories with lowest Product Complexity index associated to them.

4.3. The influence of urban morphology on economic diversity and complexity

The first analysis involving the influence of different urban morphology indicators on the economic diversity and complexity indices is to test how the different compositions of land-use affect the indices for the most recent year. This is done by visual analysis, plotting in the ternary plot (Figure 12) the three major land-uses, and graduating by colour the normalised diversity and complexity indexes. Figure 12 shows that values more centrally located in the charts, indicating a higher mix of primary uses, tend to have a stronger red colour, that indicates a higher index value. This is true for both indices, although Shannon's entropy (SEI) has a stronger red incidence towards the purely residential land-use category, whilst Economic Complexity (ECI) does not depict this.

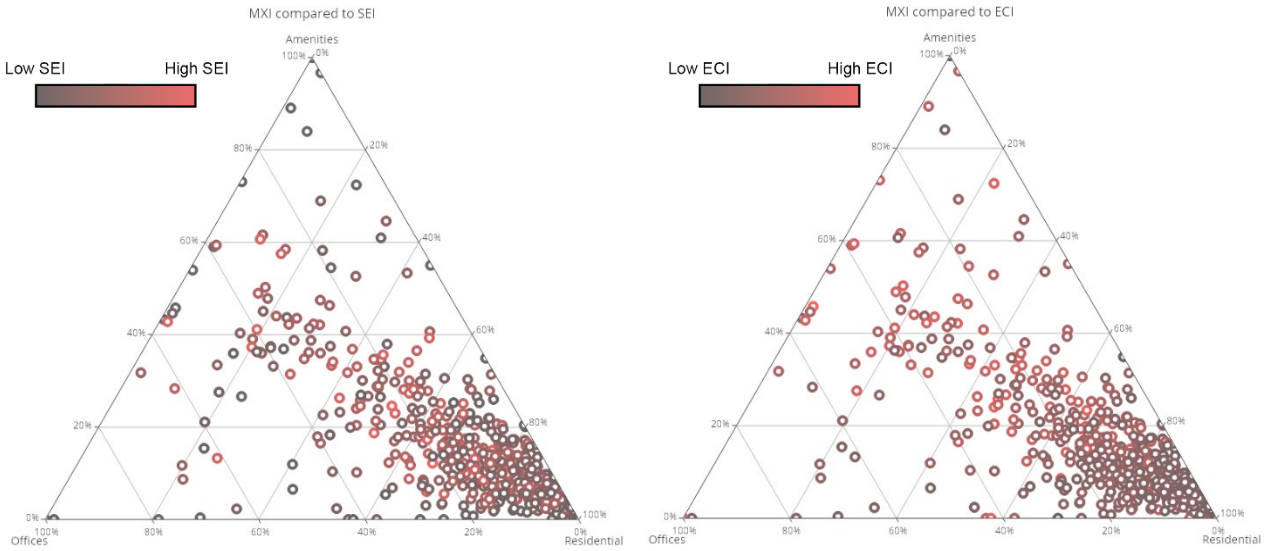


Figure 12. MXI plot compared to Shannon's entropy (SEI) - left, and Economic Complexity (ECI) - right.

The next analysis involved a correlation table between morphology indicators of a year and the economic diversity of that same year. Table 9 shows that the Floor Space Index (FSI) is positively strongly correlated to the Economic Complexity Index (ECI). This correlation increases with the increase of scale and is similar for both years of analysis. Building Size Diversity (BSD) follows a similar trend since both are correlated with one another. Another morphology indicator with consistent, positive correlation with all economic diversity indices is the Mixed-Usedness indicator (MIX), regardless of scale or year of analysis.

		2011					2015				
		SEI	ECI	DDI	SDI	SRI	SEI	ECI	DDI	SDI	SRI
Local	FSI	0.34	0.67	0.35	0.19	0.40	0.22	0.71	0.30	0.12	0.29
	GSI	0.32	0.17	0.31	0.19	0.35	0.29	0.25	0.35	0.14	0.34
	BSD	0.34	0.68	0.35	0.18	0.40	0.22	0.71	0.31	0.11	0.29
	OSR	-0.31	-0.34	-0.33	-0.28	-0.35	-0.25	-0.31	-0.30	-0.22	-0.31
	ABS	0.05	0.37	0.13	0.03	0.08	-0.01	0.41	0.06	-0.03	0.02
	ISD	0.07	-0.23	0.04	0.15	0.07	0.13	-0.25	0.10	0.20	0.12
	ADF	0.08	-0.03	0.10	0.04	0.07	0.05	-0.06	0.11	-0.00	0.07
	SNB	0.19	0.44	0.23	0.09	0.20	0.09	0.46	0.15	0.02	0.12
	MIX	0.39	0.41	0.44	0.20	0.42	0.24	0.45	0.31	0.09	0.29
Neighbourhood	FSI	0.21	0.74	0.31	-0.02	0.34	0.09	0.79	0.21	-0.02	0.22
	GSI	0.32	0.18	0.36	0.17	0.38	0.24	0.26	0.34	0.03	0.35
	BSD	0.18	0.77	0.31	0.01	0.32	0.07	0.77	0.22	-0.06	0.22
	OSR	-0.27	-0.33	-0.25	-0.07	-0.35	-0.19	-0.28	-0.22	0.04	-0.28
	ABS	-0.03	0.41	0.06	-0.11	0.04	-0.09	0.43	-0.00	-0.10	-0.02
	ISD	0.16	-0.20	0.05	0.17	0.12	0.20	-0.23	0.13	0.09	0.16
	ADF	0.13	-0.03	0.21	0.09	0.17	0.07	-0.09	0.19	-0.07	0.16
	SNB	0.10	0.48	0.19	-0.05	0.20	0.01	0.50	0.12	-0.08	0.11
	MIX	0.30	0.47	0.39	0.14	0.39	0.16	0.47	0.26	0.01	0.25
Community	FSI	0.09	0.77	0.28	-0.05	0.21	0.12	0.78	0.19	0.05	0.19
	GSI	0.21	0.07	0.35	0.17	0.26	0.21	0.12	0.30	-0.08	0.30
	BSD	0.03	0.87	0.22	-0.12	0.16	0.00	0.85	0.14	-0.02	0.10
	OSR	-0.26	-0.23	-0.32	-0.22	-0.30	-0.23	-0.16	-0.23	0.03	-0.27
	ABS	-0.16	0.30	-0.10	-0.15	-0.13	-0.18	0.34	-0.13	0.03	-0.18
	ISD	0.23	-0.12	0.22	0.26	0.22	0.25	-0.14	0.20	0.03	0.25
	ADF	0.05	-0.08	0.11	0.08	0.05	-0.05	-0.14	0.11	-0.10	0.04
	SNB	-0.00	0.46	-0.01	-0.13	0.06	-0.02	0.48	0.02	0.16	-0.00
	MIX	0.26	0.53	0.34	0.06	0.31	0.15	0.52	0.22	0.17	0.17
Regional	FSI	0.07	0.83	0.23	0.04	0.15	0.23	0.88	0.14	0.45	0.14
	GSI	0.52	0.05	0.59	0.41	0.58	0.58	0.07	0.62	0.36	0.63
	BSD	0.11	0.92	0.33	-0.03	0.25	0.25	0.93	0.31	0.35	0.29
	OSR	-0.40	0.12	-0.27	-0.40	-0.36	-0.46	0.17	-0.24	-0.33	-0.37
	ABS	-0.25	0.12	-0.47	-0.08	-0.39	-0.18	0.22	-0.51	0.14	-0.43
	ISD	0.29	0.07	0.53	0.19	0.49	0.32	-0.01	0.56	0.13	0.53
	ADF	0.39	-0.05	0.43	0.35	0.40	0.32	-0.11	0.47	0.03	0.44
	SNB	0.37	0.32	0.24	0.29	0.30	0.43	0.38	0.18	0.64	0.24
	MIX	0.54	0.48	0.50	0.21	0.51	0.40	0.41	0.31	0.22	0.37

Table 9. Correlation table between urban morphology indicators and economic diversity and complexity indexes.

A negative, moderate correlation is observed for both years between the Open Space Ratio (OSR) and all economic diversity indices, especially for the Local level. Table 9 depicts that the direction and strength of all urban morphology indicators' correlations follow the same trend for all economic diversity indexes (SEI, DDI, SDI, SRI), and differs mostly for the Economic Complexity Index (ECI). This is expected as these four diversity indices are shown before as being strongly correlated with one another, while ECI is mathematically orthogonal to them. To avoid multicollinearity problems, the most commonly used in the literature from these four indexes (Shannon's Entropy – SEI) was chosen for further regression analysis, together with ECI.

The second step to analyse the influence of urban morphology on economic diversity and complexity was to rank the average importance of factors in a series of statistical models. As described in the Methodology (sub-section 3.8.3), from this step on some of the urban morphology indicators were removed and control variables were added. The highest level of clustering, Regional, was not considered, since the number of observations (14 areas) was too little for robust regression models. The results of this ranking encompass a Linear Regression, Ridge, Lasso, Elastic Net, Lasso-Lars-IC and Random Forest estimators and can be seen in Figure 13.

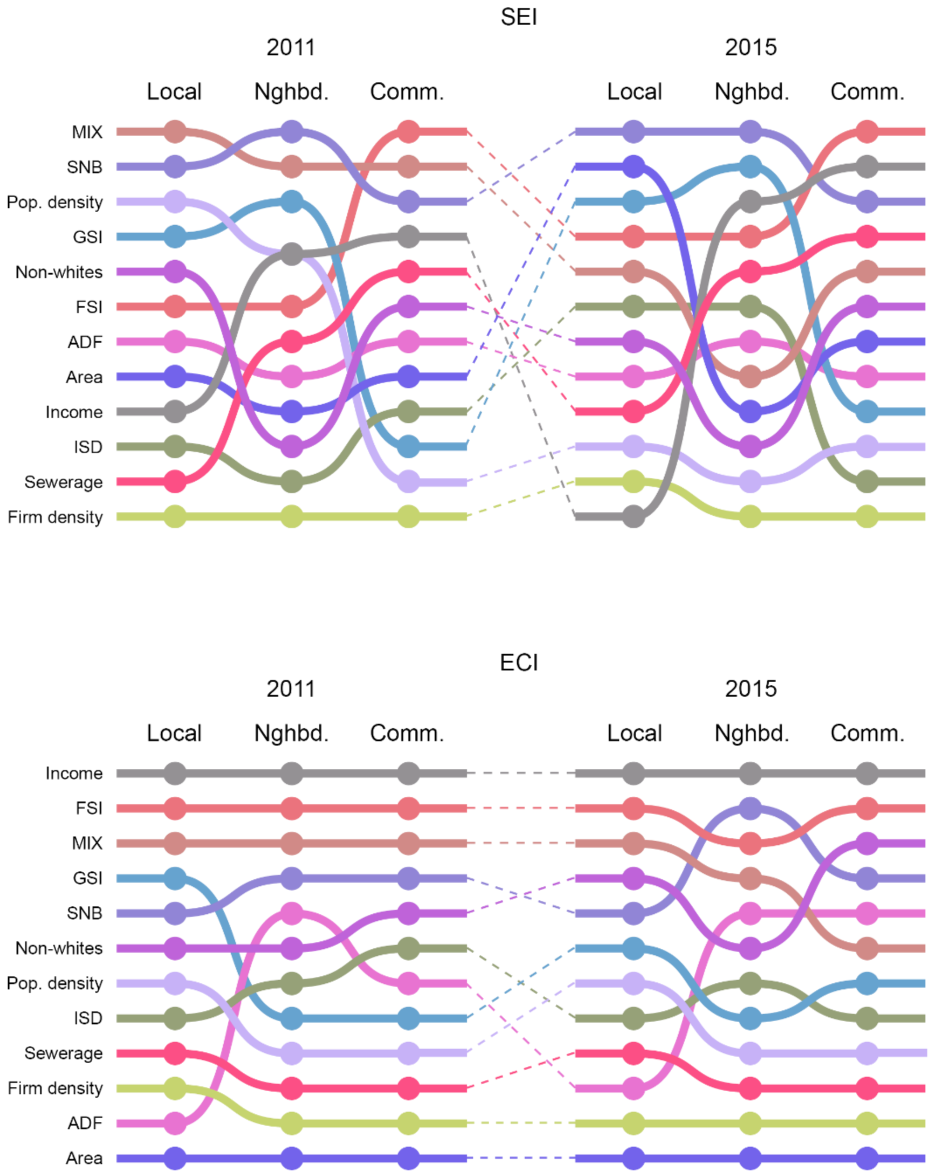


Figure 13. Ranking of importance of urban morphology factors on influencing economic diversity in a multi-model analysis. Order of the variables is their position in the ranking, the top being highest ranked in importance.

The ranking of importance of urban morphology indicators for both Shannon's Entropy (SEI) and Economic Complexity (ECI) show different patterns. At first, Figure 13 shows that the influence of urban morphology indicators on ECI is considerably steadier over different scales and different years of analysis, as compared to SEI. For both dependent variables, the Floor Space Index (FSI), related to built-up density, is consistently ranked in the top half of importance for all years and all scales. Other morphology variables, such as the degree of Mixed-Usedness (MIX) and Street Network Betweenness (SNB) also appear in the top tier of importance. For the Ground Space Index (GSI), related to the compactness of urban form, Figure 13 depicts that for both dependent variables, for both years, its importance decreases with the increase of scale. This is an indication that this factor is more localized, influencing more the composition of economic activities at a smaller scale than at larger ones.

Relevant control variables, such as access to sewerage and average income, are presented with an increasing degree of importance with the increase in scale when SEI is the dependent variable. Average income as a control variable is consistently the most important factor when ECI is the dependent variable. These rankings (importance) were used to formulate Ordinary Least Squares linear regressions' independent variables that can be seen in Figure 15. For a consistent comparison, these were selected for the smallest scale of analysis and applied to all scales for the same dependent variable. Figure 14 is a key for reading the chart depicted by Figure 15.

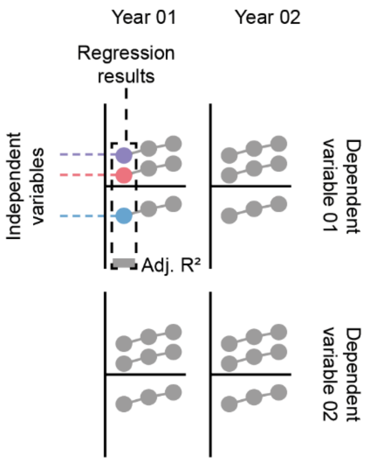


Figure 14. Key on how to read regression results for chart in Figure 14.

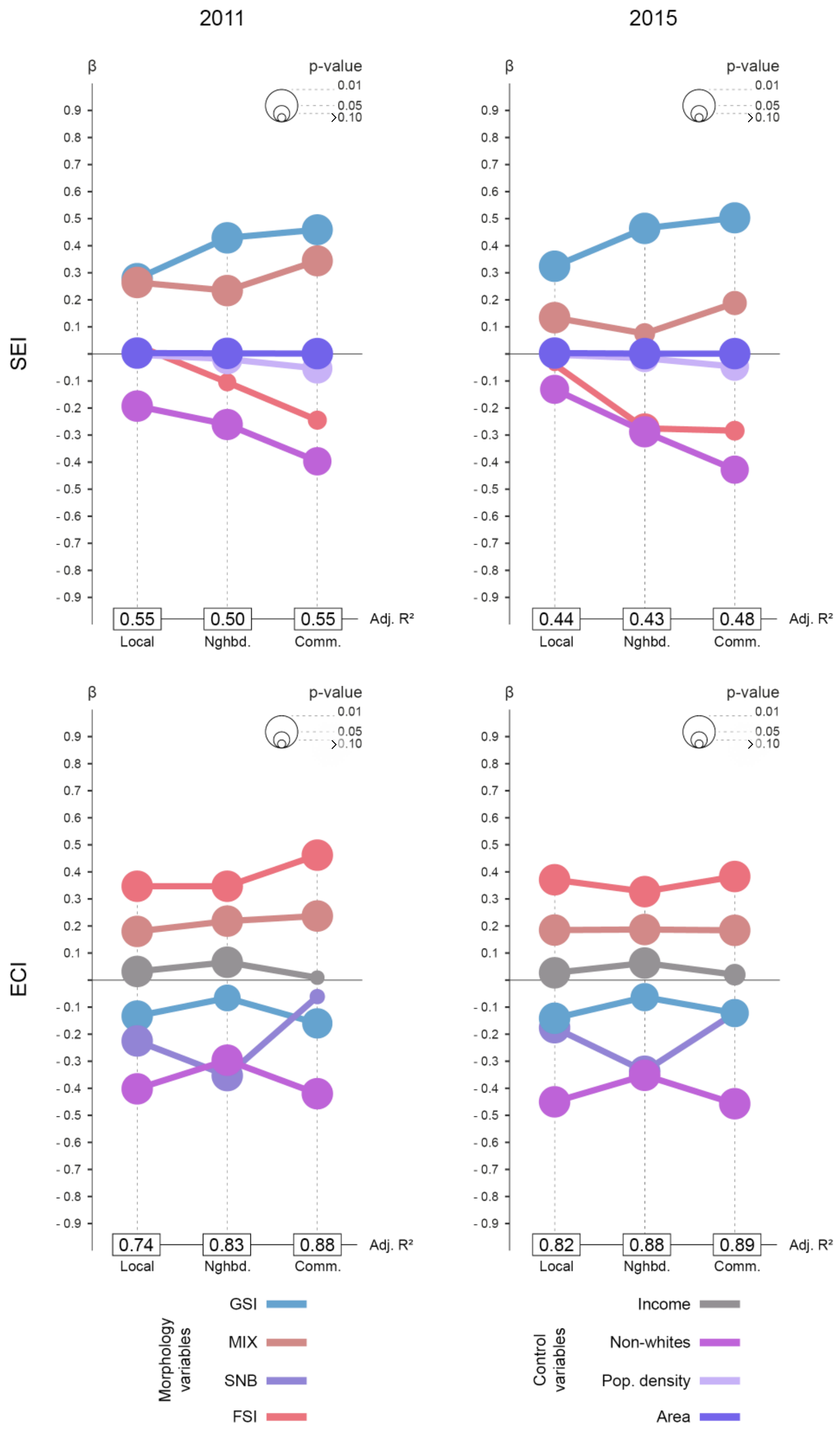


Figure 15. Regression coefficients and adjusted R^2 for both SEI and ECI as dependent variables for all scales and years of analysis.

Figure 15 depicts that Mixed-Usedness (MIX) significantly positively influences both diversity and complexity, in both years, for all scales. Since it is a land-use pattern indicator, being higher when the share of residential, retail and office floor space usage is more equally distributed, it represents an important finding to be discussed concerning land-use policy and land allocation. Street Network Betweenness (SNB), as the only connectivity indicator included in the regressions, appears only as negatively influencing the Economic Complexity Index (ECI).

The variable Ground Space Index (GSI) has a diverging influence when comparing diversity and complexity. While it has a moderate positive influence on diversity, its influence on complexity is negative and milder. This is an indication that a diverse economic environment may not necessarily be encompassed by complex economic categories, a key difference between these two indices. Another variable with diverging influence is the Floor Space Index (FSI), representing built-up density. It influences positively economic complexity and negatively economic diversity. However, its influence on economic diversity is not considered too significant by the regressions, with p-values above 0.1.

The control variable representing the percentage of non-white population follows a similar strength for both dependent variables, with a negative coefficient. It was detected that percentage of non-whites, relating to socio-economic conditions, influences negatively both diversity and complexity for all areas. Also controlling for wealth distribution is the variable related to average income. Figure 15 shows that it positively influences the complexity index (ECI).

4.4. The effects of economic diversity and complexity on economic performance

Table 10 shows the results for the correlation coefficients between economic diversity and complexity indices and economic performance indicators. Since all diversity measurements – i.e. except the complexity index ECI – are strongly correlated with one another, it was expected that their relation to the performance indicators would lead to similar coefficients. Table 10 shows that all diversity and complexity indices are strongly or moderately positively correlated to the Emergence of New Firms (ENF) for both years and all spatial scales. For Company Size Diversity (CSD), it is noticed an overall positive correlation to all indices, but diverging tendencies with the increase in spatial scale: for the complexity index (ECI), CSD decreases its correlation with the increase in scale, until almost insignificance at the highest level, whilst for all diversity indicators the other way around, showing an increase in correlation with an increase in scale.

The Percentage of Small Firms (PSF) indicator depicts largely an insignificant correlation to all diversity indices, while a strong negative correlation is observed between this variable and the complexity index. This is an indication that a business category's economic complexity is not related to the size of firms, measured in number of employees or revenue. For Companies in New Categories (CNC), a general positive correlation was detected for the diversity indices, while being not so significant for the complexity index. For Increase in Richness (INR), no clear trend was detected for different scales, indices, and years.

Rate of Business Closures (RBC) expected a negative correlation (less business closures) for higher economic diversity and complexity indexes. Although for the diversity indexes a correlation is not evident, for the complexity index there is a general positive correlation for the first time interval (2011-2015) and a strong negative one for the next time interval (2015-2019). This is an interesting trend to be discussed since the first time interval is considered one with a strong economic development in the macroeconomic level in Brazil, while the following one a strong economic crisis.

		2015					2019				
		SEI	ECI	DDI	SDI	SRI	SEI	ECI	DDI	SDI	SRI
Local	ENF	0.75	0.25	0.79	0.57	0.81	0.75	0.19	0.83	0.46	0.85
	PSF	-0.24	-0.73	-0.25	-0.06	-0.27	-0.11	-0.79	-0.17	-0.01	-0.15
	INR	0.08	-0.30	0.13	0.20	0.12	-0.18	-0.19	-0.16	-0.17	-0.17
	RBC	0.00	0.24	0.04	-0.10	0.04	0.03	-0.35	0.01	0.09	0.03
	CNC	0.71	0.12	0.71	0.51	0.72	0.59	0.07	0.62	0.35	0.63
	CSD	0.24	0.40	0.27	0.07	0.27	0.14	0.37	0.19	0.01	0.16
Neighbourhood	ENF	0.73	0.23	0.83	0.44	0.83	0.64	0.12	0.87	0.22	0.84
	PSF	-0.17	-0.81	-0.25	-0.03	-0.26	-0.03	-0.84	-0.13	-0.06	-0.12
	INR	0.02	-0.50	-0.05	0.09	-0.09	-0.35	-0.12	-0.20	-0.37	-0.30
	RBC	-0.05	0.29	0.01	-0.24	0.05	0.04	-0.50	0.05	-0.05	0.05
	CNC	0.65	-0.00	0.68	0.49	0.66	0.34	-0.13	0.45	0.13	0.42
	CSD	0.27	0.37	0.33	0.11	0.32	0.13	0.30	0.24	0.04	0.20
Community	ENF	0.77	0.20	0.91	0.44	0.88	0.70	0.09	0.94	0.34	0.88
	PSF	-0.01	-0.89	-0.15	0.12	-0.12	0.13	-0.87	0.01	-0.06	0.06
	INR	-0.07	-0.36	-0.07	-0.02	-0.12	-0.32	0.14	-0.03	-0.42	-0.15
	RBC	0.13	0.55	0.20	-0.20	0.24	-0.03	-0.63	-0.03	-0.18	-0.02
	CNC	0.67	0.02	0.78	0.43	0.75	0.10	-0.12	0.27	0.10	0.16
	CSD	0.25	0.29	0.23	-0.02	0.27	0.32	0.24	0.37	0.23	0.34
Regional	ENF	0.61	0.26	0.94	0.24	0.88	0.47	0.15	0.96	0.11	0.86
	PSF	0.03	-0.80	-0.14	0.26	-0.07	-0.06	-0.82	-0.12	-0.19	-0.08
	INR	-0.07	-0.03	0.21	-0.19	0.12	0.15	0.36	0.36	0.18	0.31
	RBC	0.08	0.62	0.16	-0.23	0.16	-0.14	-0.63	-0.08	-0.27	-0.08
	CNC	0.42	0.61	0.80	0.12	0.71	-0.17	0.01	-0.25	-0.02	-0.31
	CSD	0.54	0.01	0.37	0.30	0.44	0.64	0.10	0.42	0.44	0.52

Table 10. Correlation table between economic performance indicators and economic diversity and complexity indexes.

A second step towards understanding the influence of economic diversity and complexity on economic performance indicators was to create multiple linear regressions having as dependent variables each of the performance indicators and as independent variables one of the diversity indexes (Shannon's entropy – SEI) and the complexity index (ECI). The representation chosen for the regressions' coefficients is similar to the one used in Figure 15, although the meaning of the positions of symbols within the chart is slightly different. Figure 16 is a key on how to read the chart present in Figure 17.

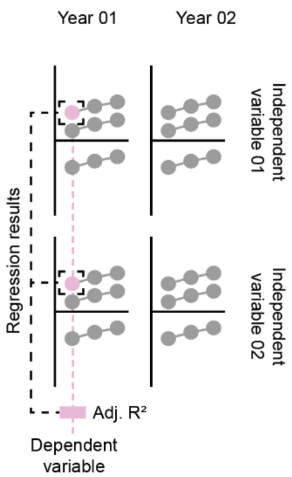
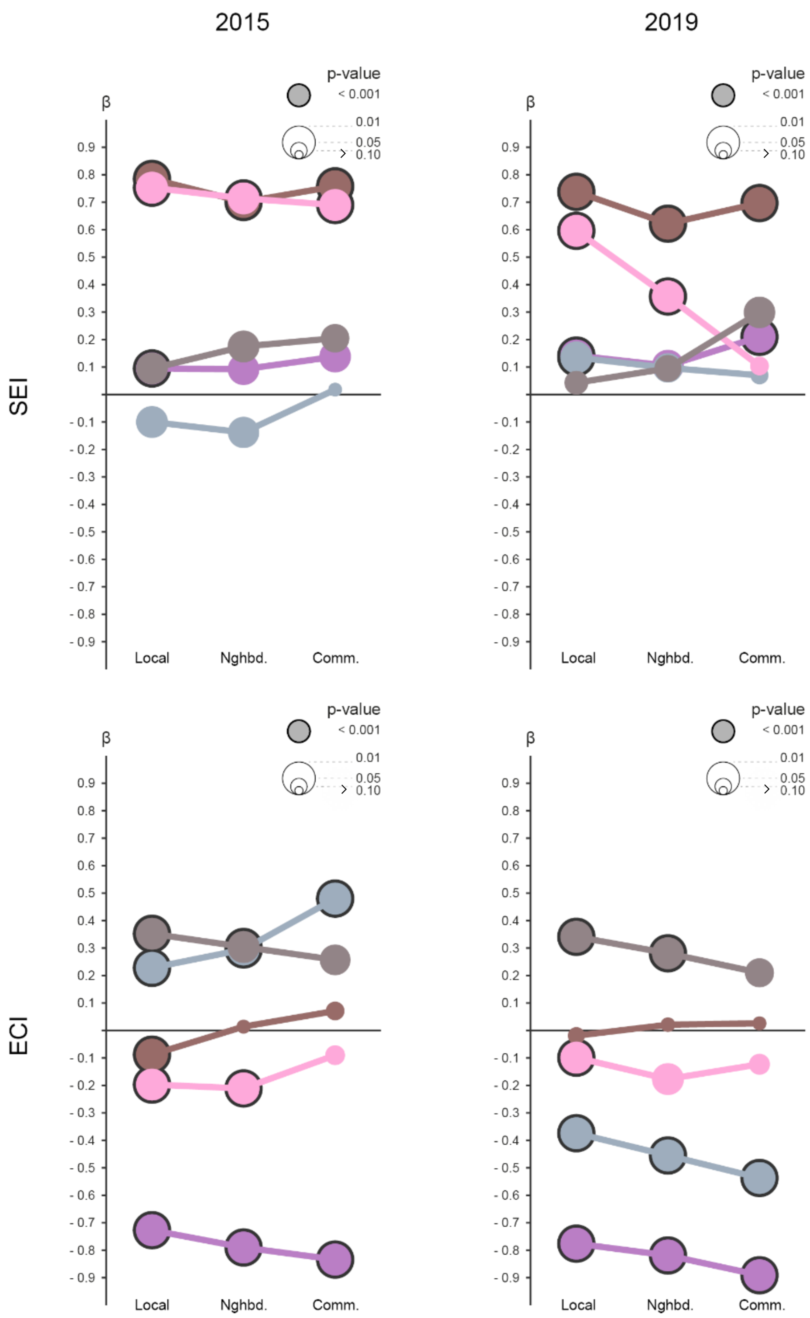


Figure 16. Key on how to read regression results for chart in Figure 16.



Adj. R^2	Local (2015)	Nghbd. (2015)	Comm. (2015)	Local (2019)	Nghbd. (2019)	Comm. (2019)
PSF	0.55	0.67	0.80	0.64	0.72	0.80
ENF	0.57	0.53	0.59	0.56	0.41	0.48
RBC	0.06	0.10	0.28	0.14	0.26	0.39
CNC	0.54	0.46	0.45	0.72	0.15	0.00
CSD	0.16	0.16	0.10	0.41	0.09	0.12

Figure 17. Regression coefficients for both SEI and ECI as independent variables and economic performance indicators as dependent variables for all scales and years of analysis.

Comparing to the trends observed in the correlation table (Table 10), similar trends were observed in Figure 17's regression coefficients. Emergence of New Firms (ENF), a proxy for entrepreneurship, is significantly positively influenced by the diversity index (SEI) and not so significantly influenced by the complexity index (ECI). The Adjusted R^2 values for ENF regressions are also high. Percentage of Small Firms (PSF), also a proxy for entrepreneurship, showed strong negative coefficients with the complexity index, while positive, but not so significant, coefficients for the diversity index. Also, the Adjusted R^2 for these regressions is high and increasing with scale.

Although the Rate of Business Closures (RBC) regressions resulted in a small Adjusted R^2, meaning the variance is not so importantly explained by the dependent variables, its value increases with scale, indicating that aggregation at higher scales increases the importance of both diversity and complexity indexes as explaining factors for this variable. Similarly to the trend observed in the correlations table, positive coefficients for this variable's influence by the complexity index were observed for the first time period (2011-2015) and negative for the second time period (2015-2019). This was a proxy for economic resilience.

Companies in New Categories (CNC), a proxy for innovation, as a dependent variable resulted in positive, significant coefficients for the diversity index, and negative and not-so-significant coefficients for the complexity index. These regressions had a different pattern of decreasing Adjusted R^2 with the increase in scale, indicating that this factor is more influenced by Local-level factors. Lastly, Company Size Diversity (CSD), as a proxy for entrepreneurship, had overall low Adjusted R^2, but positive coefficients for both complexity and diversity indexes.

5. DISCUSSION

This section discusses the results described in the previous section. It gives firstly an overall reflection on the calculation of indicators. Secondly, it discusses findings related to the forms of the spatial units calculated. Thirdly, the meanings of the economic composition of spatial units are brought up. Then, it discusses the main findings related to the influence of urban morphology on economic diversity and complexity. Furthermore, the impacts of economic diversity and complexity on economic performance are addressed. Finally, a discussion of the limitations found in this research and possible ways to address them is presented.

5.1. Reflections on indicators calculations and correlations

The calculation of indicators in economic diversity, complexity, urban morphology and economic performance generated important preliminary discussions on the meaning of these indicators. Economic diversity and complexity were found to be different concepts, measuring different phenomena. Prior to regression analyses, indicators of urban morphology were found to be highly correlated to one another, which based the exclusion of some of the selected indicators from further regressions. The same can be said for the economic performance indicators. These are discussed below.

All economic diversity indices were highly correlated with one another. It was expected that α-diversities would be correlated since they measure essentially the same phenomenon – i.e. the variety of business categories within clusters. However, they were also found to be strongly correlated to the only β-diversity index selected, Density-Diversity (DDI), that compares the concentration of activities between all areas. Economic complexity, on the other hand, is presented with a low correlation towards all other diversities measured, confirming its mathematical orthogonality as previously described in the literature (Burlina & Antonietti, 2020).

For urban morphology indicators, some correlations were also observed. A strong correlation between Building Size Diversity (BSD) and Floor Space Index (FSI) indicates that more densely built-up areas are the same ones that present a highly diverse built-up environment. Areas less intensively built-up on average are concluded to be also more homogeneous. So whenever a high average FSI is mentioned in this research, it can be interpreted as having a high heterogeneity in building size. A negative correlation was found between Ground Space Index (GSI) and Open Space Ration (OSR), indicating that less compact areas tend also to have a higher amount of open space per built-up area. A similarly negative correlation found between Intersection Density (ISD) and Average Block Size (ABS), as expected by the nature of both indicators essentially measuring the same thing.

Internal correlations between economic performance indicators also generated meaningful discussions. Increase in Richness (INR) was removed from regressions for being strongly correlated to both emergence of new firms (ENF) and companies in new categories (CNC). But it is striking to highlight from this that areas where more new business categories emerged were areas where more new firms, regardless of their categories, were emerging already. And it is also striking that areas with the most increase in new categories had the highest percentages of new firms belonging to these new categories. This is a sign that levels of entrepreneurship, as the creation of firms, and innovation, as the emergence of new work, go hand-in-hand. This represents a quantitative confirmation for Jacobs' (1970) qualitative explanation for how urban areas thrive economically.

5.2. Clustering patterns of economic activities in the urban fabric

Mapping the economic activities, detecting their concentration for different clustering levels, and segmenting the economic landscape into spatial units of analysis provide us with interesting observations on how economic activities cluster in space. Firstly, this sub-section discusses the emergent patterns detected by the different clustering levels. Secondly, it discusses the classification of these spatial units regarding the form of their peaks.

The Local level cluster detection for this research generated spatial units very close to the streetscape, being detected that the segmentations followed barriers also found in the landscape, such as larger street crossings, rivers, parks, among other areas with a lower concentration of economic activities. The same pattern is found to be reproduced for higher-level clustering. The Neighbourhood-level cluster detection is close in scope to the city's division of neighbourhoods, while the Regional level follows the city's division into administrative regions that function as local governance bodies.

The peaks of concentration of economic activities in these clusters are also called seeds, since they generated their respective spatial unit. The urban form associated with these peaks was classified into four categories, according to Lynch (1960). This was done for the highest degree of aggregation, the Regional level, since this classification was done manually, based on visual perceptions of where these peaks are located. For further research, it could be an interesting expansion to automatize this peak classification process. As such, the same classification could be done for lower degrees of aggregation as well, and resulting categories could be included in further regression analyses, for instance. This could indicate whether the form of peaks, a derived emergent pattern of agglomeration economies, act together with other studied factors, such as urban morphology, on the intensity of economic diversity or complexity.

Regarding the classification performed, it is seen that the two main economic poles of the city are classified as Nodes, indicating an important crossing of roads around which activities agglomerate. This is the case for the so-called Square 7th of September, traditional city centre landmark and the Savassi Square, a more recent economic centre. Although road crossings are, by definition, public spaces and, thus, devoid of formal economic activities per se, it is possible to conclude that the strong concentration of activities is a result of this type of configuration. This confirms Jacobs's (1961) assumptions that crossings effectively generate places of encounter, favouring knowledge spill-over effects and, therefore, fostering diversity of demands and increasing positive externalities.

It is also noticed by the classification that the Linear category of peaks is more often found towards the North of the city, an area with generally poorer inhabitants and lower densities. Furthermore, Building category of peaks is found to polarise usually smaller spatial units than other categories. This needs to be further expanded in future research, but it is a sign that shopping centres might effectively hinder a smoother distribution of economic activities in their area of polarisation. Based on the size of these spatial units, the concentration of a large number of activities within a single building may indicate that enterprises acting alone in the surroundings find it hard to compete with the convenience of shopping centres.

5.3. Economic composition of spatial units

Analysing the economic composition of spatial units was important to understand how the Economic Complexity Index (ECI) differs in concept with the other diversity indices. Firstly, the Product-Space of business categories, by-product of complexity calculation, shows a remarkable core of activities that co-occur, indicating they are more widespread in the territory, despite their complexity. Secondly, an analysis of betweenness in the network highlights that more generalised business categories act as catalysers for the presence of a higher number of categories and, therefore, increasing diversity. Thirdly, the type of business categories classified as more complex indicates that post-industrial, technology-related, knowledge-intensive categories stand out as being the most complex. Finally, on the other hand, more labour-intensive, less advanced business categories appear classified as less complex. These findings are described below.

An entangled network of firm category colocation does not necessarily translate into an enhanced economic complexity. There are sub-networks of high complexity and sub-networks of low complexity, depending on the nature of economic activities that happen. The remarkable core in the centre of the chart (Figure 8) reflects firm categories that often appear together, suggesting that these are more widespread in the urban tissue. Although it is expected that more widespread categories appear as less complex, this is not the case for this core. Its complexity ranges in mid-values, connecting branches of higher complexity in the top-right corner with branches of lower complexity in the bottom corner (Figure 8).

By classifying the Product Space by the betweenness of its nodes (Figure 9), it is highlighted the business categories that link more intensively other kinds of businesses, regardless of their complexity. It can be concluded that the presence of categories such as law firms, consultancy firms, engineering services, supermarkets, fitness centres and music production, act as a catalyser for a richer composition of firm category within an area. In other words, the presence of these categories of firms fosters diversity itself. This is a curious finding since these categories do not appear to be themselves specialised. In fact, their level of generality – a law firm can act as auxiliary to a multitude of other business categories, for instance – may actually be the reason why they appear as strong catalysers in the first place.

The sort of business categories figured with the highest complexity values (Figure 10) is strongly in line with categories highlighted by Murdoch (2018) as demanding highly-skilled labour in the post-industrial economy and seen as the main sources of the economic prosperity for global cities (Sassen, 2005). These are financial, telecommunications, scientific and technical services, educational services, healthcare and social assistance, among others. In this intra-urban analysis, these categories were also seen to contribute to higher levels of complexity of the economy. It is important to highlight that some of these categories are characterised by having a strong or exclusive governmental presence in the Brazilian context, especially medical services, higher educational facilities, hospitals, and public administration services (government-related by definition). This indicates that local and national governments can have a decisive role in complexifying local economies.

In the other side of the spectrum, Murdoch (2018) describes less advanced industry categories as being economically traditional: construction, manufacturing, retail trade, transportation and warehousing, among others. These coincide with the lowest complexity categories from this research (Figure 11). This indicates a change in the scope when transferring applications of the Economic Complexity Index from the common country or regional level down to city and intra-city levels. Higher levels of aggregation (e.g. country-level) tend to position manufacturing industries within the higher complexity ones (Hidalgo, 2015), as opposed to commodity extraction and agriculture as low complexity, for instance. By transferring it to the city-scale, incorporating the service industry to the analysis, manufacturing figures among the lowest levels of complexity, possibly because agricultural and mining activities are less commonly found in urban areas.

5.4. The influence of urban morphology on economic diversity and complexity

Several factors of the urban morphology were found to influence either economic diversity, economic complexity, or both. Firstly, it is important to notice that some independent variables need to be assessed qualitatively as a group, such as the relations between Floor Space Index, Ground Space Index, the control for income and Economic Complexity Index, since they have underlying urban configurations worth noticing. Secondly, diversity is found to be positively influenced by Ground Space Index, indicating that areas with higher compactness promote diverse economic activities, although these are not translated into complex business categories. Thirdly, most indicators related to connectivity were not relevant enough to be included in the regressions and, the one that was (Street Network Betweenness), had an opposite direction of influence on complexity than expected. Finally, the mixture of uses is found to be the most consistent factor in influencing positively both diversity and complexity.

Some indicators related to the built environment dimension were found to significantly influence economic diversity and complexity. Floor Space Index (FSI) being strongly related to Economic Complexity Index (ECI) in the regression needs to be interpreted in combination with two other variables: the control for income and Ground Space Index (GSI). FSI and the control for income being positively related to complexity, while GSI follows a negative trend is a sign that a specific sort of building form is related to higher complexity, and this building type might be significantly related to the households' income. Buildings for higher-income families in Belo Horizonte tend to be constructed in higher-demand areas, using the maximum built-up potentials allowed by legislation, which also restricts the possible occupation area of a plot (GSI). This generates a specific type of building form, with tall buildings in the centre of plots, combining both high FSI and low GSI. Therefore, this should rather be considered as the type of usual practice concerning urban form production than as separate influencing factors being interpreted individually. A neighbourhood with this sort of building patterns is shown in Figure 18.

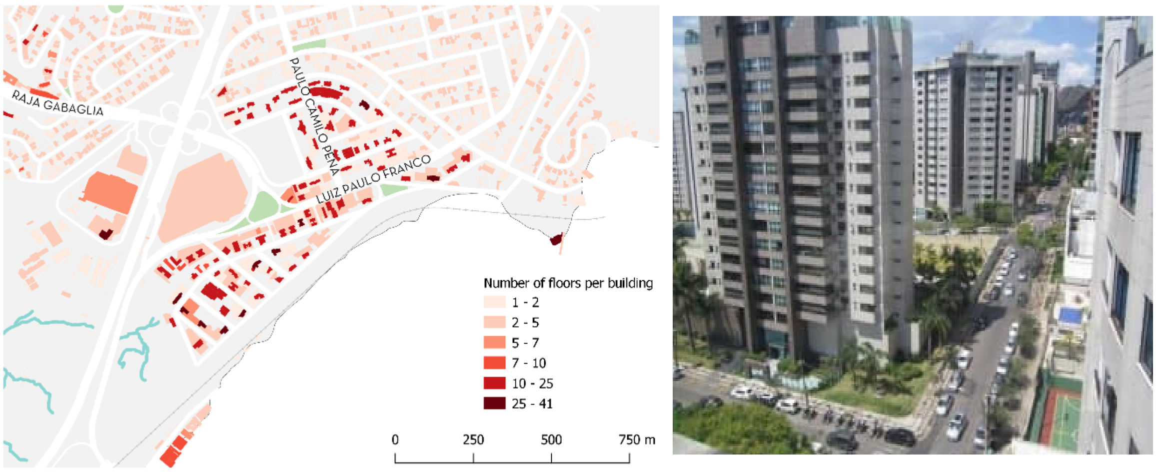


Figure 18. The luxurious neighbourhood of Belvedere (high income), consisting of tall buildings (high FSI), centralised in their plots (low GSI). Source: Trovit Imóveis.

On the other hand, a high Ground Space Index (GSI) was found to positively influence economic diversity. Floor Space Index (FSI) in the same regressions was found as having a negative influence, but highly insignificant. This generates a discussion on the difference in concept between economic diversity and economic complexity since GSI has diverging influence in them. Again, the nature of complexity index is to assign different business categories to different complexity values, so it is safe to say that these businesses are identified to happen where the previously discussed urban form is built. Diversity, on the other hand, does not consider the exact categories of business, but rather just their variety. It is believed, then, that in more traditional urban tissues, where GSI tends to be higher, it might emerge a large variety of low-complex business categories, regardless of the intensity of that built-up area (FSI). This goes along Jacobs' (1961) theory that higher compactness of the urban form incentives inhabitants to use public spaces more intensely, generating a stronger demand for local shops and resulting in a larger variety of categories.

Indicators related to the connectivity of plots were not found to be so relevant in explaining economic diversity or complexity. Street Network Betweenness was the only indicator considered relevant enough to be included in the analysis, and it was detected to have a negative relation to economic complexity, the contrary to expected. This might have to do with the nature of the urban environment in the case study. Informal settlement areas have a more organic settlement that increases the intersection density, reduces block size, increases street betweenness, and increases address fragmentation. So despite all these items having as expectation a positive correlation to both diversity and complexity, a negative one was found for street network betweenness and the other ones were not even considered relevant for the regression analyses by the selection of subsets. This pattern detected for informal settlements is shown in Figure 19.

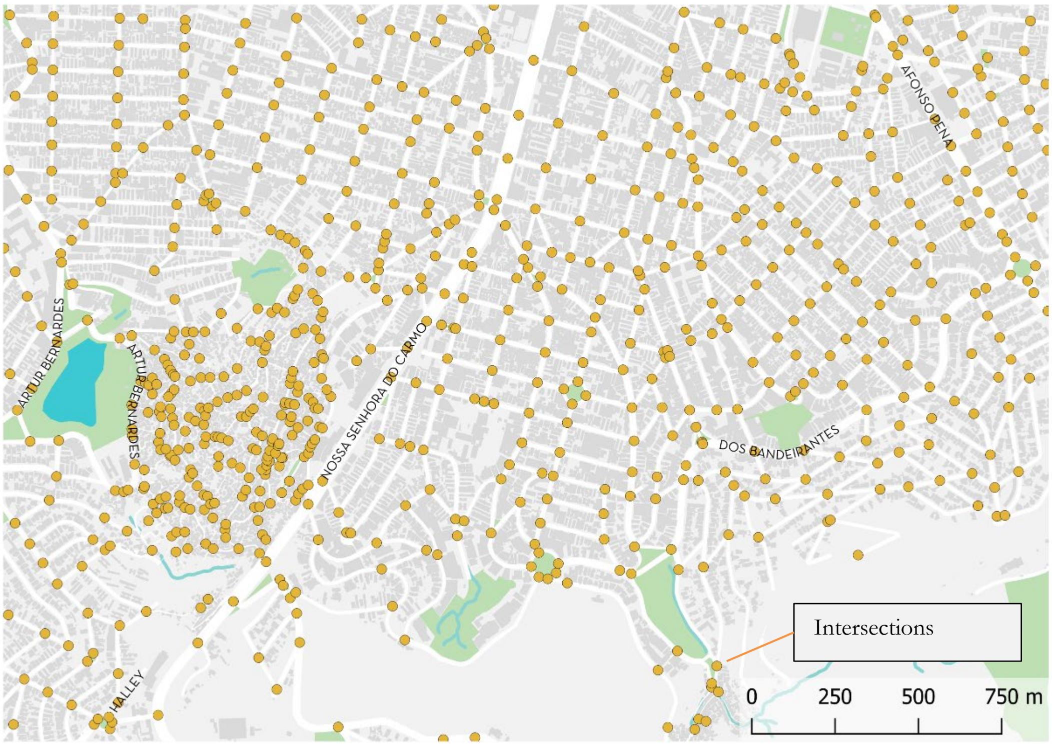


Figure 19. Intersections detected for connectivity dimension indicators. Informal area (left) shows a much higher concentration of intersections than formal areas (others).

Concerning land-use, mixed-usedness was considered positively associated with both diversity and complexity. This can also be confirmed when plotting the MXI ternary plot (Figure 12) against values of economic complexity (ECI) and diversity (SEI), being higher values for ECI and SEI located more centrally in the ternary. This confirms Jacobs' (1961) assumption of the benefits deriving from having a high mix of primary uses, as well as Berghauser Pont & Haupt's (2009), Mashhoodi & Berghauser Pont's (2011) and Nes et al.'s (2012). Despite both diversity and complexity being more intense in multifunctional spatial units, diversity tends to be higher in more residential areas, while complexity tends to have higher values in areas that share offices and amenities. This comparison was done visually, and for further research that intends to search the relevance of specific combinations of land-use could be included as dummy variables in regressions, or even the lone percentages for each land-use could be included as continuous variables.

5.5. The effects of economic diversity and complexity on economic performance

Economic diversity's influence on economic performance follows the expected trends and directions more accurately than complexity's. Diversity's influence on entrepreneurship and innovation indicators were found to be more consistent, whereas complexity negatively influences Rate of Business Closures (RBC) for the second time frame, indicating a stronger resilience. Economic complexity is strongly negatively associated with Percentage of Small Firms (PSF), indicating that complex business categories are expected to be encompassed by larger firms, contrary to expectation. Finally, Company Size Diversity (CSD) is positively associated with both diversity and complexity, although not relevantly.

General observation shows that Shannon's entropy (SEI) regression coefficients follow the expectations for the performance proxies more accurately than the Economic Complexity Index (ECI) coefficients. That is just not the case for the Rate of Business Closures (RBC), which was expected to be negatively connected and appeared as positively influenced by SEI for the year 2019. On the other hand, ECI appeared as influencing positively the RBC for 2015, in an unexpected way, but negatively influencing it for 2019, for which adjusted R^2 is also higher, indicating more relevance for explaining the variance. This is an interesting pattern since the period of 2015-2019 is marked by a strong economic crisis in Brazil. Therefore, a higher ECI is observed to halt business closures in a period where these are more expected to happen.

Another interesting pattern to be discussed is how Economic Complexity (ECI) negatively impacts the Predominance of Small Firms indicator (PSF), opposite to diversity's (SEI) impact on it. This classification for whether a firm is considered small or not depends on a government classification, according to the firm's earnings and the number of employees. Therefore, although the predominance of small firms is often associated with positive economic performance (Glaeser et al., 2010), its negative association to economic complexity might indicate that the type of business categories considered to be complex actually do not fulfil a government requirement to be classified as a small firm. That might be for either having higher revenues than the threshold or a higher number of employees.

Company Size Diversity (CSD) is positively correlated to average firm size, indicating that places with large firms are also the ones with a more heterogeneous composition. This is also related to the previously discussed negative influence of ECI towards PSF, since complex activities are expected to have larger companies in revenue and number of employees. Company Size Diversity has positive relations to both diversity and complexity, despite overall not found to be relevantly influenced by neither, since its regressions' results indicated a low Adjusted R^2.

5.6. Limitations

The main limitations found in this research are regarding the comprehensiveness of the datasets available, the precision of information available in the datasets, and coverage of tools available. These are described in this sub-section, together with preliminary recommendations for addressing these issues in further research.

It was detected that the nature of the dataset used presented an important limitation to be highlighted for the scope of this research. Since the locational data for firms is dependent on a formal registration with municipal authorities, informal economic activities, very present in the context of Brazil, are not well depicted by the dataset. This is especially true for poorer areas, especially slum areas (*favelas*) and the poor peripheries typical of Brazilian cities. Although some control variables were added to account for this absence, such as average income and racial composition of the population, overall results for both urban morphology and economic diversity were impacted by the limitations of the dataset. It is also worth mentioning that this research did not propose to address the effects of informal economic activities on local economies and entrepreneurship. Such a project would need to involve extensive fieldwork and could still be significantly imprecise given the dynamic nature of the informal sector. However, a general inspection of the dataset done in the preliminary phases of this research detected a high accuracy of the economic activities depicted as compared to Google Maps registries or Street View analysis even for informal slum areas.

Luckily, data for urban morphology was not seen as being severely affected by informal settlement areas. Due to the comprehensiveness of the datasets freely provided by the Municipal Government of Belo Horizonte, as well as a perceived extensive attempt to have formalized plot systems within informal areas, data related to both plots and buildings are impressively precise, as compared to satellite imagery for instance. However, direct access to the source surveying files, such as Digital Elevation Models and Digital Surface Models, would significantly improve the accuracy of data for building heights, for instance, that was dependent on previous interpretation summarised in polygon shapefiles.

Throughout the course of this research, some innovative methods were discovered and enhance future applications of similar methodologies. This is the case, for instance, for the morphological tessellation described by Fleischmann, Feliciotti, Romice, & Porta (2020) and implemented as the *python* library *momepy* (Fleischmann, 2019). The authors argue for migrating urban morphology analysis from the scope of plot systems towards such method, that considers immediate surrounding public spaces, such as street space, to be part of a plot as a single unit of analysis.

Another method related to both urban morphology and economic diversity measurements that was partially used in this research and could have had its use further expanded was the Place Syntax Tool (PST). It could allow for a multi-level approach that would prevent a significant Modifiable Areal Unit Problem (MAUP), by considering individual elements such as an economic activity throughout multiple scales by modifying thresholds. Although this research tried to avoid inconsistencies from pre-determined administrative boundaries and developing a bottom-up, data-driven spatial unit definition, it still sliced the area of analysis into observable spatial units, and incurred on the MAUP by transitioning between multiple scales. Further studies related to complex systems and emergent patterns should try to avoid that altogether, by modelling individual elements and their relations to surrounding as close in scale as possible to the actual elements and their characteristics.

6. CONCLUSION

This research tested the relations between economic diversity, urban morphology and economic performance; and between economic complexity, urban morphology and economic performance, in an intra-urban context, comparing different areas of the city of Belo Horizonte, in Brazil, across time and spatial scales. It confirms that economic diversity and economic complexity are different concepts and influenced by urban morphology differently. They also have different impacts on economic performance of urban areas. Using OLS regressions, this research has found meaningful results from these relations, either confirming or contradicting previous theories and assumptions in urban planning and regional science, as described below.

Some theoretical assumptions from previous studies of cities are quantitatively confirmed in this research. One of the findings is that mixed-use areas coincide significantly with higher economic diversity and complexity, indicating that multi-functionality of urban areas, where primary uses such as residences, retail and offices are harmonically present, is confirmed as an important factor for thriving economies. Moreover, built-up intensity is found to be positively associated with economic complexity, and compactness is positively associated with economic diversity. It is interesting to observe that some processes need to be understood as qualitative arrangements between different indicators. Low compactness and high built-up intensity are associated with higher incomes as an indication of a specific type of urban form, common in richer urban areas of Belo Horizonte. This indicates that a qualitative understanding of how cities are built needs to be considered also in quantitative analyses. Thirdly, a higher complexity also indicated a stronger economic resilience in a period of economic crisis (2015-2019), while higher diversity indicated stronger rates for entrepreneurship and innovation.

On the other hand, there are theoretical assumptions contradictory to this research's findings. This is the case for the connectivity dimension of urban morphology, highlighting the importance of considering the context of application when referring to urban theories. In North American and European contexts, where urban theories usually emerge, connectivity indicators, such as small block sizes, are associated with traditional urban cores, commonly areas with thriving economies and local vitality, as opposed to post-war, sprawling modernist neighbourhoods or suburbs. In the Brazilian context, which can also be expanded to other cities from the Global South, the connectivity indicators used in this research are especially higher in organically grown poor peripheries and slum areas (*favelas*). These are often devoid of adequate infrastructure, house the poorer parcels of the population, and lack behind in socioeconomic conditions, for historical reasons. Furthermore, economic activities in these areas are often neglected by official records and statistics, being informal activities the predominant, and economic performance and diversity indicators are lower in comparison to other, richer areas within the same city.

This study has shown that the Economic Complexity Index can be successfully transferred from the traditional application of products exported by a region or a country to services performed within a city. The resulting complexity of business categories confirms provisions by urban theorists that post-industrial business categories guide the development of urban areas. This highlights the importance of development of specific plans and policies considering the scale of application, since the process of complexification of an economy is observed differently from scale to scale. Further research is encouraged to analyse how different industry categories behave in relation to complexity when transitioning between local and national scales.

Further research is also encouraged to bring these quantitative relations between urban form and economic performance to different urban contexts so that findings can be extrapolated into more comprehensive, comparative scopes. Current developments in urban form assessment, such as morphological tessellation (Fleischmann et al., 2020) and place syntax (KTH School of Architecture & Chalmers Architecture and Spacescape AB, 2019), provide tools for a more adequate approach, thinking of cities and their economies as complex systems in constant coevolution, with enablers, catalysts and friction between its parts. Complex systems need to be modelled with special attention to a system's parts since they function together generating emergent patterns that are the commonly observable and studied.

It is important to highlight possible contributions of this research for future public policies to improve economic performance. The urban morphology indicators chosen are usually objects of regulation by local governments. This is the case, in Brazil, of building regulations, masterplans and land-use plans. More specifically, land-use plans consist of the provision of zoning policies, such as single-use zoning strategies, or definition of specific categories of business activities allowed to locate in specific types of road. As shown by this research, a mix of primary uses is one of the main factors influencing a dynamic economic environment in an urban area. Therefore, limiting entire areas of cities to single uses, such as residential, hinders the possibility of a diverse and complex economic environment to emerge. These are, in turn, responsible for a stronger entrepreneurship rate and the rise of new business categories, that increases the provision of services for citizens. Thus, it can be concluded that limiting whole areas of cities to single uses affect the overall economic performance of such areas and, by extension, the quality of life of its citizens.

Regarding building regulations, urban planners usually set maximum standards for Floor Space Index and Ground Space Index in new buildings, or derived measurements, with different names but similar meanings. This research highlights that a high Ground Space Index actually contributes to a higher economic diversity, confirming Jacobs's assumption that higher compactness allows for more encounters and more economic activities flourish. Floor Space Index is seen as contributing for a higher economic complexity, although this is more likely associated to a specific building type occurring in wealthier areas of Belo Horizonte. The presence of more complex economic activities generates a more complex economy as a whole, being an indication that local governments should consider revisiting using these two concepts for limiting the possibility of building forms.

This research also highlights that local or even national governments can have their own role in complexifying local economies. Some of the business categories classified as complex commonly belong to the public sphere, such as higher education facilities, hospitals, and public administration services. It is valuable that the positioning of these facilities in cities take the underlying objective of fostering local economic complexity into account. Besides acting directly, governments could also promote the enhancement of specific industry categories considered to be more complex, the enhancement of categories that are shown to be more connected to others (enhancing diversity), or, ideally, prioritising multiple business categories that act together as networks for enhancing complexity. These are presented by this research as potentially more efficient manners of promoting economic performance than the common directed specialisation policies, usually aiming at single sectors (Hong & Xiao, 2016). As such, governments can improve an urban area's resilience from external economic shocks, protecting citizens against possible job losses derived from business closures in an area.

By bringing the assessment of economic diversity, complexity and performance down to spatial units within urban areas, this research highlights the importance of considering the intra-urban scales in the formulation of public policies. Some aspects of economic performance are affected by economic diversity more strongly in local scales, such as rates of innovation. This indicates that diversity spill-overs towards enhancing innovation are strongly influenced by geographical proximity and do not extend to more distant areas. Policies for enhancing diversity and complexity should, therefore, pay attention to localised planning, instead of large-scale interventions. On the other hand, economic complexity has been identified with an increasing influence on economic resilience with the increase of scale. This is a sign that policies for fostering economic complexity affect positively the resilience of broader areas, with its positive externalities reaching farther areas of the city.

Overall, this research has made use of innovative methods and has contributed to enhance current literature in the field of urban economics. The application of the Economic Complexity Index to intra-urban areas, including the service industry, bridges a gap between the economics of development and urban and regional sciences. It also contributes to summarising the quantification of urban theory assumptions and theories not often tested quantitatively.

LIST OF REFERENCES

Acs, Z. J., & Audretsch, D. B. (1988). Innovation in Large and Small Firms: An Empirical Analysis. *The American Economic Review*, *78*(4), 678–690. doi:10.2307/1811167

Adler, P., Florida, R., King, K., & Mellander, C. (2019). The city and high-tech startups: The spatial organization of Schumpeterian entrepreneurship. *Cities*, *87*(December 2018), 121–130. doi:10.1016/j.cities.2018.12.013

Andersson, M., Klaesson, J., & Larsson, J. P. (2016). How Local are Spatial Density Externalities? Neighbourhood Effects in Agglomeration Economies. *Regional Studies*, *50*(6), 1082–1095. doi:10.1080/00343404.2014.968119

Araldi, A., & Fusco, G. (2019). Retail Fabric Assessment: Describing retail patterns within urban space. *Cities*, *85*(September 2018), 51–62. doi:10.1016/j.cities.2018.11.025

Baeza, J. L., Cerrone, D., & Männigo, K. (2017). Comparing two methods for urban complexity calculation using the Shannon-Wiener Index. *WIT Transactions on Ecology and the Environment*, *226*(1), 369–378. doi:10.2495/SDP170321

Balland, P.-A. (2017). *Economic Geography in R: Introduction to the EconGeo package* (Papers in Evolutionary Economic Geography No. 17.09). Utrecht.

Balland, P.-A. (2020). EconGeo: Computing Key Indicators of the Spatial Distribution of Economic Activities. Retrieved from https://github.com/PABalland/EconGeo

Balland, P.-A., Jara-Figueroa, C., Petralia, S. G., Steijn, M. P. A., Rigby, D. L., & Hidalgo, C. A. (2020). Complex economic activities concentrate in large cities. *Nature Human Behaviour*, *4*(3), 248–254. doi:10.1038/s41562-019-0803-3

Bandara, R. (2016). Image Segmentation using Unsupervised Watershed Algorithm with an Over - segmentation Reduction Technique. *Lecture Notes in Information Technology*, 1–14. Retrieved from www.codeproject.com/Articles/751744/Image-Segmentation-using-Unsupervised-Watershed-Al

Batty, M. (2017). Diverse cities, successful cities. *Nature Human Behaviour*, *1*(0022), 2. doi:10.1038/s41562-016-0022

Batty, M., Besussi, E., Maat, K., & Harts, J. J. (2004). Representing Multifunctional Cities: Density and Diversity in Space and Time. *Built Environment*, *30*(4), 324–337. doi:10.2148/benv.30.4.324.57156

Berghauser Pont, M. Y., & Haupt, P. A. (2009). *Space, density and urban form*. University of Delft, Delft, NL. Retrieved from http://resolver.tudelft.nl/uuid:0e8cdd4d-80d0-4c4c-97dc-dbb9e5eee7c2

Bishop, A., & Mateos-Garcia, J. (2019). Exploring the Link Between Economic Complexity and Emergent Economic Activities. *National Institute Economic Review*, *249*(1), R47–R58. doi:10.1177/002795011924900114

Bobkova, E., Marcus, L., Berghauser Pont, M., Stavroulaki, I., & Bolin, D. (2019). Structure of Plot Systems and Economic Activity in Cities: Linking Plot Types to Retail and Food Services in London, Amsterdam and Stockholm. *Urban Science*, *3*(3), 66. doi:10.3390/urbansci3030066

Burlina, C., & Antonietti, R. (2020). *From variety to economic complexity: empirical evidence from Italian regions* (Papers in Evolutionary Economic Geography No. 19.30). Utrecht.

Bustos, S., Gomez, C., Hausmann, R., & Hidalgo, C. A. (2012). The Dynamics of Nestedness Predicts the Evolution of Industrial Ecosystems. *PLoS ONE*, *7*(11), e49393. doi:10.1371/journal.pone.0049393

Carol, H. (1960). The hierarchy of central functions within the city. *Annals of the Association of American Geographers*, *50*(4), 419–438. doi:10.1111/j.1467-8306.1960.tb00359.x

Chen, Y. (2014). Multifractals of central place systems: Models, dimension spectrums, and empirical analysis. *Physica A: Statistical Mechanics and Its Applications*, *402*, 266–282. doi:10.1016/j.physa.2014.01.061

Christaller, W. (1966). *Central Places in Southern Germany*. (C. W. Baskin, Ed.). Englewood Cliffs, NJ: Prentice Hall.

Cottineau, C., Finance, O., Hatna, E., Arcaute, E., & Batty, M. (2018). Defining urban clusters to detect agglomeration economies. *Environment and Planning B: Urban Analytics and City Science*, 239980831875514. doi:10.1177/2399808318755146

Cozzolino, S. (2019). The (anti) adaptive neighbourhoods. Embracing complexity and distribution of design control in the ordinary built environment. *Environment and Planning B: Urban Analytics and City Science*, *0*(0), 1–17. doi:10.1177/2399808319857451

Cutrini, E. (2009). Using entropy measures to disentangle regional from national localization patterns. *Regional Science and Urban Economics*, *39*(2), 243–250. doi:10.1016/j.regsciurbeco.2008.08.005

Feliciotti, A., Romice, O., & Porta, S. (2016). Design for change: five proxies for resilience in the urban form. *Open House International*, *41*(4), 23–30.

Feliciotti, A., Romice, O., & Porta, S. (2017). Urban regeneration, masterplans and resilience: The case of Gorbals, Glasgow. *Urban Morphology*, *21*(1), 61–79.

Fleischmann, M. (2019). momepy: Urban Morphology Measuring Toolkit. *Journal of Open Source Software*, *4*(43), 1807. doi:10.21105/joss.01807

Fleischmann, M., Feliciotti, A., Romice, O., & Porta, S. (2020). Morphological tessellation as a way of partitioning space: Improving consistency in urban morphology at the plot scale. *Computers, Environment and Urban Systems*, *80*, 101441. doi:10.1016/j.compenvurbsys.2019.101441

Florida, R., Adler, P., & Mellander, C. (2017). The city as innovation machine. *Regional Studies*, *51*(1), 86–96. doi:10.1080/00343404.2016.1255324

Folta, T. B., Cooper, A. C., & Baik, Y. S. (2006). Geographic cluster size and firm performance. *Journal of Business Venturing*, *21*(2), 217–242. doi:10.1016/j.jbusvent.2005.04.005

Forman, C., Goldfarb, A., & Greenstein, S. (2016). Agglomeration of Invention in the Bay Area: Not Just ICT. *American Economic Review*, *106*(5), 146–151. doi:10.1257/aer.p20161018

Frenken, K., Van Oort, F., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies*, *41*(5), 685–697. doi:10.1080/00343400601120296

Gilbert, B. A., McDougall, P. P., & Audretsch, D. B. (2006). New Venture Growth: A Review and Extension. *Journal of Management*, *32*(6), 926–950. doi:10.1177/0149206306293860

Glaeser, E. L. (2011). *Triumph of the city: how our greatest invention makes us richer, smarter, greener, healthier, and happier* (Kindle Ed.). New York: Penguin Publishing Group.

Glaeser, E. L., Kerr, W. R., & Ponzetto, G. A. M. (2010). Clusters of entrepreneurship. *Journal of Urban Economics*, *67*(1), 150–168. doi:10.1016/j.jue.2009.09.008

Gomez-Lievano, A., & Patterson-Lomba, O. (2018). *Uncovering the drivers behind urban economic complexity and their connection to urban economic performance*. Retrieved from https://arxiv.org/pdf/1812.02842.pdf

Gotway, C. A., & Young, L. J. (2002). Combining incompatible spatial data. *Journal of the American Statistical Association*, *97*(458), 632–648. doi:10.1198/016214502760047140

Government of Brazil. Access to Information Act, Pub. L. No. 12.527 (2011). Brazil. Retrieved from http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2011/Lei/L12527.htm

Hamidi, S., & Zandiatashbar, A. (2019). Does urban form matter for innovation productivity? A national multi-level study of the association between neighbourhood innovation capacity and urban sprawl. *Urban Studies*, *56*(8), 1576–1594. doi:10.1177/0042098018767002

Hausmann, R., & Hidalgo, C. A. (2011). The network structure of economic output. *Journal of Economic Growth*, *16*(4), 309–342. doi:10.1007/s10887-011-9071-4

Hidalgo, C. A. (2015). *Why Information Grows: The Evolution of Order, from Atoms to Economies*. Penguin Books.

Hidalgo, C. A., & Castañer, E. E. (2015). *The Amenity Space and The Evolution of Neighborhoods*. Retrieved from https://arxiv.org/ftp/arxiv/papers/1509/1509.02868.pdf

Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, *106*(26), 10570–10575. doi:10.1073/pnas.0900943106

Hidalgo, C. A., Klinger, B., Barabasi, A.-L., & Hausmann, R. (2007). The Product Space Conditions the Development of Nations. *Science*, *317*(5837), 482–487. doi:10.1126/science.1144581

Hong, S., & Xiao, Y. (2016). The influence of multiple specializations on economic performance in U.S. Metropolitan areas. *Sustainability (Switzerland)*, *8*(9), 1–16. doi:10.3390/su8090963

Hsu, W.-T. (2012). Central place theory and city size distribution. *Economic Journal*, *122*(563), 903–932. doi:10.1111/j.1468-0297.2012.02518.x.

Jacobs, J. (1961). *The Death and Life of Great American Cities*. New York: Vintage.

Jacobs, J. (1970). *The Economy of Cities* (1st ed.). New York: Vintage.

Jacobs, J. (1985). *Cities and the Wealth of Nations*. New York: Vintage Books.

Jost, L. (2006). Entropy and diversity. *OIKOS*, *113*(2), 363–375.

Kajtazi, B. (2007). Measuring Multifunctionality of Urban Area, 89. doi:10.1002/cpa.21520

KTH School of Architecture, & Chalmers Architecture and Spacescape AB. (2019). PST QGIS Plugin. doi:10.13140/RG.2.2.25718.55364

Liu, L., Peng, Z., Wu, H., Jiao, H., Yu, Y., & Zhao, J. (2018). Fast identification of urban sprawl based on K-means clustering with population density and local spatial entropy. *Sustainability (Switzerland)*, *10*(8). doi:10.3390/su10082683

Lynch, K. (1960). *The Image of the City*. The MIT Press.

Maraschin, C., & Krafta, R. (2013). Growth dynamic of retail locations: a methodological approach using a logistic model. *Cybergeo*, *2013*. doi:10.4000/cybergeo.26002

Martin, R., Sunley, P., Gardiner, B., & Tyler, P. (2016). How Regions React to Recessions: Resilience and the Role of Economic Structure. *Regional Studies*, *50*(4), 561–585. doi:10.1080/00343404.2015.1136410

Martin, R., Sunley, P., & Turner, D. (2002). Taking risks in regions: the geographical anatomy of Europe's emerging venture capital market. *Journal of Economic Geography*, *2*(2), 121–150. doi:10.1093/jeg/2.2.121

Mashhoodi, B., & Berghauser Pont, M. Y. (2011). Studying land-use distribution and mixed-use patterns in relation to density, accessibility and urban form. In *18th International Seminar on Urban Form* (pp. 1–19). Delft, NL.

Modica, M., & Reggiani, A. (2015). Spatial Economic Resilience: Overview and Perspectives. *Networks and Spatial Economics*, *15*(2), 211–233. doi:10.1007/s11067-014-9261-7

Monte-Mór, R. L. (2005). What is the urban in the contemporary world? *Cadernos de Saúde Pública*, *21*(3), 942–948. doi:10.1590/S0102-311X2005000300030

Mulligan, G. F., Partridge, M. D., & Carruthers, J. I. (2012). Central place theory and its reemergence in regional science. *Annals of Regional Science*, *48*(2), 405–431. doi:10.1007/s00168-011-0496-7

Murdoch, J. (2018). Specialized vs. diversified: The role of neighborhood economies in shrinking cities. *Cities*, *75*, 30–37. doi:10.1016/j.cities.2016.12.006

Nes, A. van, Pont, M. B., & Mashhoodi, B. (2012). Combination of Space Syntax with Spacematrix and the Mixed Use Index. The Rotterdam South test case. In *Eighth International Space Syntax Symposium* (pp. 1–29). Santiago de Chile, Chile. doi:urn:NBN:nl:ui:24-uuid:d865e5b6-519e-4959-8c3d-800054d1b351

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., … Wagner, H. (2019).

vegan: Community Ecology Package. Retrieved from https://cran.r-project.org/package=vegan

Pant, R., Barker, K., & Zobel, C. W. (2014). Static and dynamic metrics of economic resilience for interdependent infrastructure and industry sectors. *Reliability Engineering and System Safety*, *125*, 92–102. doi:10.1016/j.ress.2013.09.007

Quatraro, F., & Usai, S. (2017). Knowledge flows, externalities and innovation networks. *Regional Studies*, *51*(8), 1133–1137. doi:10.1080/00343404.2017.1337884

Rahman, M. S., & Islam, M. R. (2013). Counting objects in an image by marker controlled watershed segmentation and thresholding. In *3rd IEEE International Advance Computing Conference, IACC 2013* (pp. 1251–1256). IEEE. doi:10.1109/IAdCC.2013.6514407

Rappaport, J. (2008). Consumption amenities and city population density. *Regional Science and Urban Economics*, *38*(6), 533–552. doi:10.1016/j.regsciurbeco.2008.02.001

Ridd, M. K. (1995). Exploring a V-I-S (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities†. *International Journal of Remote Sensing*, *16*(12), 2165–2185. doi:10.1080/01431169508954549

Sassen, S. (2005). The global city: introducing a concept. *Brown Journal of World Affairs*, *11*(2), 27–43. Retrieved from https://research.vu.nl/en/publications/the-global-city-introducing-a-concept

Shutters, S. T., Muneepeerakul, R., & Lobo, J. (2015). Quantifying urban economic resilience through labour force interdependence. *Palgrave Communications*, *1*(1), 15010. doi:10.1057/palcomms.2015.10

Smit, M. J., Abreu, M. A., & de Groot, H. L. F. (2015). Micro-evidence on the determinants of innovation in the Netherlands: The relative importance of absorptive capacity and agglomeration externalities. *Papers in Regional Science*, *94*(2), 249–272. doi:10.1111/pirs.12068

Smith, A. (1776). *The Wealth of Nations* (Cannan Ed.). New York: Modern Library, 1937.

Sprague, C. D. (2018). *Resilience and the U.S. labor market: a cross-scale analysis on the role of industrial diversity and specialization*. Boise State University.

Vissers, G., & Dankbaar, B. (2013). Path dependence and path plasticity: textile cities in the Netherlands. *Zeitschrift Für Wirtschaftsgeographie*, *57*(1–2), 83–95. doi:10.1515/zfw.2013.0007

Xiao, Y., & Drucker, J. (2013). Does economic diversity enhance regional disaster resilience? *Journal of the American Planning Association*, *79*(2), 148–160. doi:10.1080/01944363.2013.882125

APPENDIX 01

For each one of the fields Type of Practise and Type of Facility's categories, it was defined whether the category fulfils the criterion condition (√), does not fulfil the condition (×), or not necessarily (–). An economic activity point is set to be discarded from the analysis if it is classified under a category that does not fulfil (×) at least one of the conditions. Table 11 and Table 12 depict, for each one of the categories, how the criteria were evaluated.

Table 11. Categories for Type of Practise and whether they fulfil the established criteria.

Type of practise	**Description**	**1**	**2**	**3**
At a fixed facility	Activity is practised in a determined establishment/building, within a property, for which the address coincides with the registered address.	√	√	–
Via internet	Activity is practised solely via the internet.	–	×	√
In a fixed spot, out of address	Activity is practised in a fixed spot, but outside the facility/building/headquarters: kiosks, selling stands, etc., for which the address does not coincide with the registered address.	×	×	√
Via postal services	Activity is practised with offers/purchases/hires via postal mail: selling through catalogues, portfolios, packages, mail bags, etc., despite the means used for service/product delivery.	–	×	√
Door-to-door, mobile posts or ambulatory	Activity is practised with a physical displacement of seller/supplier directly towards the clients' personal or professional domiciles: direct sales, street markets, street vendors, itinerant traders.	–	×	√
Teleshopping	Activity is practised with offers/purchases/hires via telephone.	–	×	√
Vending machines	Activity is practised using automatic/electronic vending machines: beverages machines, self-service, miscellaneous, etc.	√	√	×
Activity practised out of the facility	Activity is practised at the clients' addresses, not at the company's own registered address.	×	×	√

Table 12. Categories for Type of Facility and whether they fulfil the established criteria.

Type of facility		**Description**	**1**	**2**	**3**
Productive unit		Operational unit, when company practises the activities of producing or selling goods or services to third parties.	√	√	–
Auxiliary unit	Warehouse	Facility where company stores producer's goods for own use.	×	×	×
	Data processing centre	Informatic centre for company's internal use.	×	×	×
	Training centre	Personnel training and qualification centre, for company's internal use.	√	×	√

Type of facility		Description	1	2	3
	Storage	Facility where company stores its own goods, destined to either production or selling, within which sells are not made.	×	×	×
	Vehicle depot	Facility exclusive for parking the company's vehicles.	×	×	×
	Repairer	Facility where maintenance and repairing of company's fixed assets' goods.	×	×	–
	Exhibition spot	Facility for exhibition of company's own products, without commercial transactions, e.g., showrooms.	–	×	√
	Headquarters	Company's central administrative unit, board of directors, chief executives.	–	–	√
	Fuel supply	Exclusively for company's own vehicle fleet.	×	×	×
	Collection post	Facility for customer service, aiming at collecting products / materials / goods / equipment / information for further following them to productive unit, responsible for their analysis / processing / publication.	–	√	√

APPENDIX 02

Table 13. Descriptive statistics of all indicators calculated.

Level	*Dimension*	*Indicator*	*Abbreviation*	*2011*				*2015*				*2019*			
				Mean	*Std*	*Min*	*Max*	*Mean*	*Std*	*Min*	*Max*	*Mean*	*Std*	*Min*	*Max*
Local	Diversity	Richness Index	SRI	54.84	35.22	1.00	205.00	75.10	37.40	1.00	193.00	81.15	35.34	8.00	197.00
		Shannon's Entropy	SEI	5.22	1.19	0.00	7.10	5.76	0.90	0.00	7.21	5.91	0.72	2.30	7.17
		Simpson's Diversity	SDI	0.95	0.09	0.00	1.00	0.97	0.05	0.00	1.00	0.97	0.02	0.61	0.99
		Density-Diversity	DDI	15.40	17.09	0.00	204.28	19.87	18.58	0.00	212.19	21.41	18.15	0.96	189.86
		Economic Complexity Index	ECI	0.00	1.00	-2.36	4.65	0.00	1.00	-1.98	4.11	0.01	1.00	-1.98	3.31
	Morphology	Floor Space Index	FSI	1.32	0.91	0.07	10.76	1.39	0.97	0.23	11.64				
		Ground Space Index	GSI	0.60	0.10	0.11	0.91	0.62	0.10	0.11	0.93				
		Open Space Ratio	OSR	2.47	12.54	0.01	263.79	2.56	12.76	0.01	270.65				
		Building Size Diversity	BSD	0.95	0.80	0.09	7.58	0.94	0.79	0.22	7.43				
		Mixed-Usedness	MIX	0.42	0.26	0.00	0.99	0.46	0.27	0.00	0.99				
		Intersection Density (per há)	ISD	1.03	0.54	0.00	4.32	1.03	0.54	0.00	4.32				
		Average Block Size (m)	ABS	104.81	37.39	37.85	505.01	104.81	37.39	37.85	505.01				
		Address Fragmentation	ADF	2.29	1.14	1.17	13.78	2.29	1.14	1.17	13.78				
		Street Network Betweenness	SNB	1.03	0.05	0.62	1.35	1.03	0.05	0.62	1.35				
	Performance	Emergence of New Firms	ENF					116.16	240.59	0.00	5866.67	68.62	79.16	6.90	1900.00
		Company Size Diversity	CSD					0.69	0.08	0.05	0.75	0.71	0.06	0.08	0.75
		Predominance of Small Firms (%)	PSF					86.77	11.76	0.00	100.00	87.21	11.68	17.24	100.00
		Increase in Richness	INR					71.26	137.53	-86.67	2000.00	18.75	74.10	-28.57	1900.00
		Companies in New Categories (%)	CNC					40.95	18.37	2.91	100.00	25.81	14.95	1.61	100.00
		Rate of Business Closures (%)	RBC					18.43	13.00	0.00	100.00	39.58	9.88	0.00	75.00
Nbghd.	Diversity	Richness Index	SRI	109.34	45.36	5.00	234.00	136.77	46.35	27.00	262.00	141.52	42.49	36.00	267.00
		Shannon's Entropy	SEI	6.37	0.70	2.25	7.56	6.64	0.52	4.58	7.60	6.68	0.42	4.77	7.60
		Simpson's Diversity	SDI	0.98	0.02	0.78	0.99	0.98	0.01	0.93	0.99	0.98	0.01	0.94	0.99
		Density-Diversity	DDI	33.80	33.32	0.89	298.36	41.99	35.68	1.86	293.29	44.59	34.99	5.62	272.21
		Economic Complexity Index	ECI	0.00	1.00	-1.80	2.85	0.00	1.00	-1.60	2.88	0.00	1.00	-1.41	2.83
	Morphology	Floor Space Index	FSI	1.30	0.71	0.23	6.71	1.37	0.76	0.30	7.13				

Level	*Dimension*	*Indicator*	*Abbreviation*	*2011*				*2015*				*2019*			
				Mean	*Std*	*Min*	*Max*	*Mean*	*Std*	*Min*	*Max*	*Mean*	*Std*	*Min*	*Max*
		Ground Space Index	GSI	0.61	0.08	0.14	0.89	0.62	0.08	0.15	0.91				
		Open Space Ratio	OSR	1.94	4.12	0.13	39.85	2.01	4.17	0.15	39.85				
		Building Size Diversity	BSD	0.97	0.72	0.28	6.61	0.96	0.72	0.25	6.49				
		Mixed-Usedness	MIX	0.47	0.23	0.08	0.99	0.51	0.25	0.00	0.99				
		Intersection Density (per há)	ISD	1.00	0.43	0.10	2.79	1.00	0.43	0.10	2.79				
		Average Block Size (m)	ABS	100.57	22.43	45.31	201.35	100.57	22.43	45.31	201.35				
		Address Fragmentation	ADF	2.19	0.69	1.24	7.15	2.19	0.69	1.24	7.15				
		Street Network Betweenness	SNB	1.03	0.05	0.68	1.20	1.03	0.05	0.68	1.20				
	Performance	Emergence of New Firms	ENF					87.85	53.05	20.00	550.00	59.98	21.27	14.49	195.00
		Company Size Diversity	CSD					0.73	0.03	0.56	0.75	0.74	0.02	0.61	0.75
		Predominance of Small Firms (%)	PSF					86.87	9.75	28.99	98.47	87.01	10.41	25.76	97.51
		Increase in Richness	INR					33.03	37.23	-7.02	440.00	6.30	15.80	-15.31	155.56
		Companies in New Categories (%)	CNC					24.47	11.80	1.70	88.37	13.33	8.11	0.94	60.56
		Rate of Business Closures (%)	RBC					17.42	8.84	0.00	82.31	39.09	7.59	9.26	56.90
Comm.	Diversity	Richness Index	SRI	54.84	35.22	1.00	205.00	210.48	65.22	59.00	357.00	214.60	62.97	81.00	362.00
		Shannon's Entropy	SEI	7.02	0.55	4.93	7.93	7.09	0.46	5.14	7.86	7.06	0.38	5.76	7.71
		Simpson's Diversity	SDI	0.95	0.09	0.00	1.00	0.98	0.01	0.93	0.99	0.98	0.01	0.96	0.99
		Density-Diversity	DDI	70.44	78.53	2.00	524.70	83.46	84.01	4.95	525.20	87.68	81.94	6.89	497.58
		Economic Complexity Index	ECI	0.00	1.00	-1.28	2.36	0.00	1.00	-1.18	2.32	0.00	1.00	-1.32	2.38
	Morphology	Floor Space Index	FSI	1.22	0.48	0.73	3.27	1.28	0.52	0.74	3.52				
		Ground Space Index	GSI	0.60	0.06	0.45	0.76	0.62	0.06	0.47	0.78				
		Open Space Ratio	OSR	2.62	5.86	0.20	39.85	2.69	5.89	0.24	39.85				
		Building Size Diversity	BSD	0.95	0.58	0.52	4.12	0.95	0.59	0.51	4.15				
		Mixed-Usedness	MIX	0.48	0.22	0.13	0.98	0.54	0.23	0.16	0.99				
		Intersection Density (per há)	ISD	0.93	0.38	0.17	2.27	0.93	0.38	0.17	2.27				
		Average Block Size (m)	ABS	98.87	21.35	49.52	171.41	98.87	21.35	49.52	171.41				
		Address Fragmentation	ADF	2.18	0.52	1.38	4.69	2.18	0.52	1.38	4.69				
		Street Network Betweenness	SNB	1.02	0.05	0.84	1.13	1.02	0.05	0.84	1.13				
	Performance	Emergence of New Firms	ENF					88.03	35.64	39.23	212.00	59.46	14.67	33.59	104.40

Level	*Dimension*	*Indicator*	*Abbreviation*	*2011*				*2015*				*2019*			
				Mean	*Std*	*Min*	*Max*	*Mean*	*Std*	*Min*	*Max*	*Mean*	*Std*	*Min*	*Max*
		Company Size Diversity	CSD					0.74	0.01	0.71	0.75	0.75	0.00	0.74	0.75
		Predominance of Small Firms (%)	PSF					86.34	10.15	45.32	97.66	86.48	10.40	53.68	96.51
		Increase in Richness	INR					18.55	14.75	-2.24	88.89	3.20	8.45	-9.82	45.76
		Companies in New Categories (%)	CNC					12.85	7.88	1.54	42.51	5.90	4.80	0.46	29.44
		Rate of Business Closures (%)	RBC					16.90	6.82	2.55	45.92	39.12	7.62	10.24	51.64
Regional	Diversity	Richness Index	SRI	54.84	35.22	1.00	205.00	319.07	79.17	164.00	492.00	325.43	83.88	166.00	529.00
		Shannon's Entropy	SEI	5.22	1.19	0.00	7.10	7.50	0.26	6.84	7.82	7.42	0.25	6.81	7.74
		Simpson's Diversity	SDI	0.95	0.09	0.00	1.00	0.99	0.00	0.98	0.99	0.99	0.00	0.98	0.99
		Density-Diversity	DDI	15.40	17.09	0.00	204.28	184.40	223.40	15.84	858.82	192.44	217.66	19.66	833.72
		Economic Complexity Index	ECI	0.00	1.00	-2.36	4.65	0.00	1.00	-1.01	1.93	0.00	1.00	-1.08	1.78
	Morphology	Floor Space Index	FSI	1.13	0.23	0.88	1.63	1.19	0.24	0.91	1.70				
		Ground Space Index	GSI	0.58	0.04	0.48	0.65	0.60	0.04	0.51	0.67				
		Open Space Ratio	OSR	2.43	2.13	0.89	7.69	2.52	2.21	0.92	7.94				
		Building Size Diversity	BSD	0.90	0.30	0.60	1.48	0.90	0.30	0.59	1.49				
		Mixed-Usedness	MIX	0.48	0.19	0.19	0.90	0.56	0.18	0.29	0.91				
		Intersection Density (per há)	ISD	0.80	0.21	0.44	1.13	0.80	0.21	0.44	1.13				
		Average Block Size (m)	ABS	99.88	13.77	82.51	127.15	99.88	13.77	82.51	127.15				
		Address Fragmentation	ADF	2.06	0.21	1.64	2.41	2.06	0.21	1.64	2.41				
		Street Network Betweenness	SNB	1.02	0.06	0.84	1.10	1.02	0.06	0.84	1.10				
	Performance	Emergence of New Firms	ENF					84.14	28.37	55.11	154.35	59.22	11.55	43.00	82.10
		Company Size Diversity	CSD					0.75	0.00	0.75	0.75	0.75	0.00	0.75	0.75
		Predominance of Small Firms (%)	PSF					86.36	8.41	62.52	95.64	86.71	9.23	60.00	96.51
		Increase in Richness	INR					9.59	7.02	-4.80	26.15	1.92	4.42	-6.74	7.52
		Companies in New Categories (%)	CNC					5.17	4.14	1.21	17.36	1.74	1.54	0.09	5.86
		Rate of Business Closures (%)	RBC					17.96	4.64	9.89	28.05	39.28	6.03	23.95	49.56

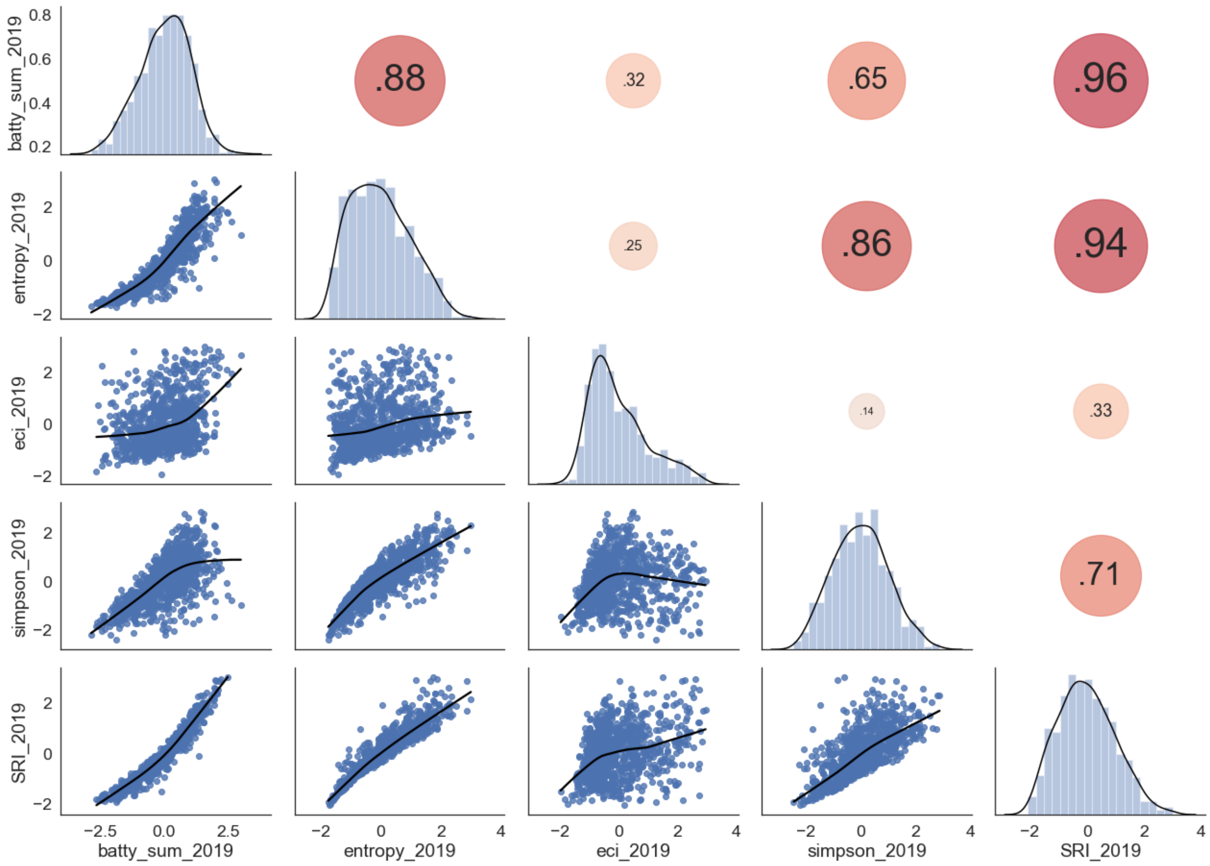


Figure 20. Sample of correlation testing for indicators within the scope of Economic Diversity

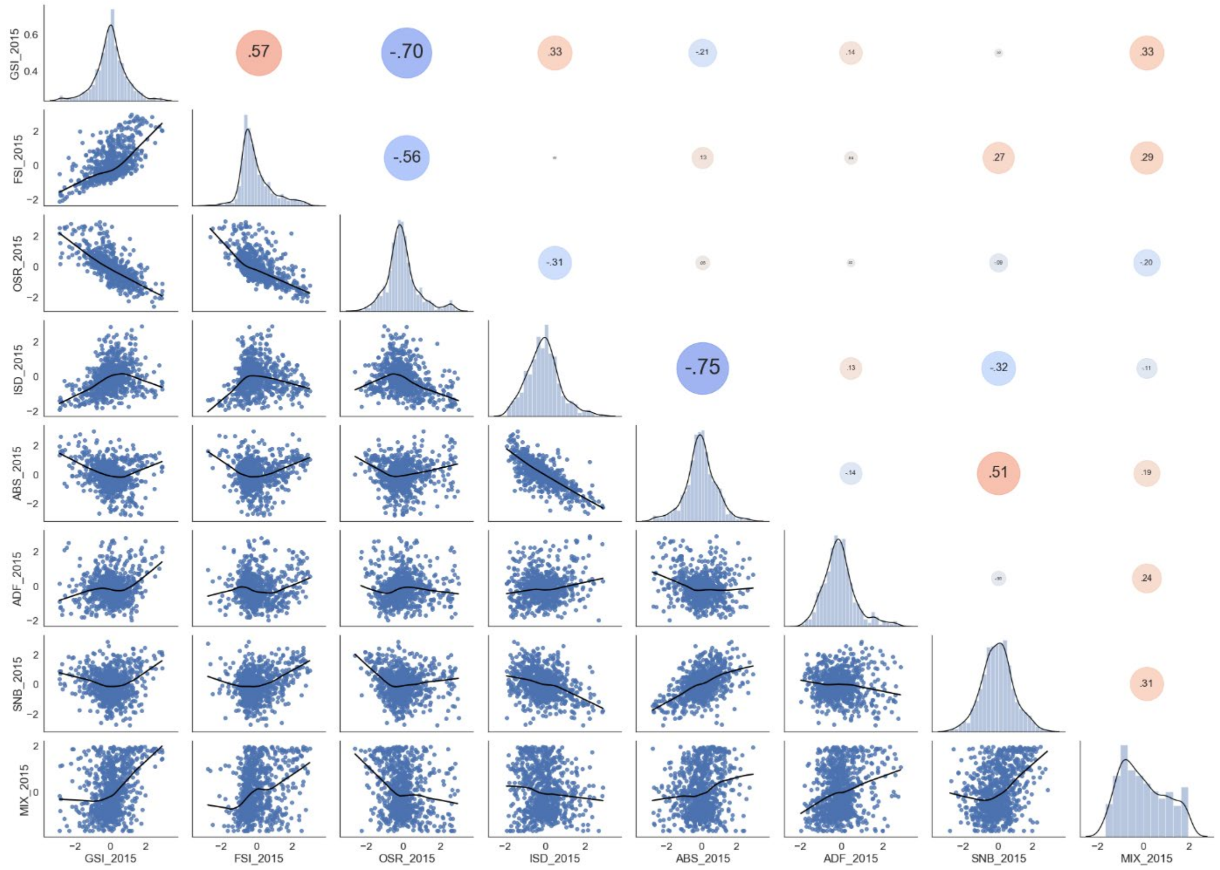


Figure 21. Sample of correlation testing for indicators under the scope of Urban Morphology

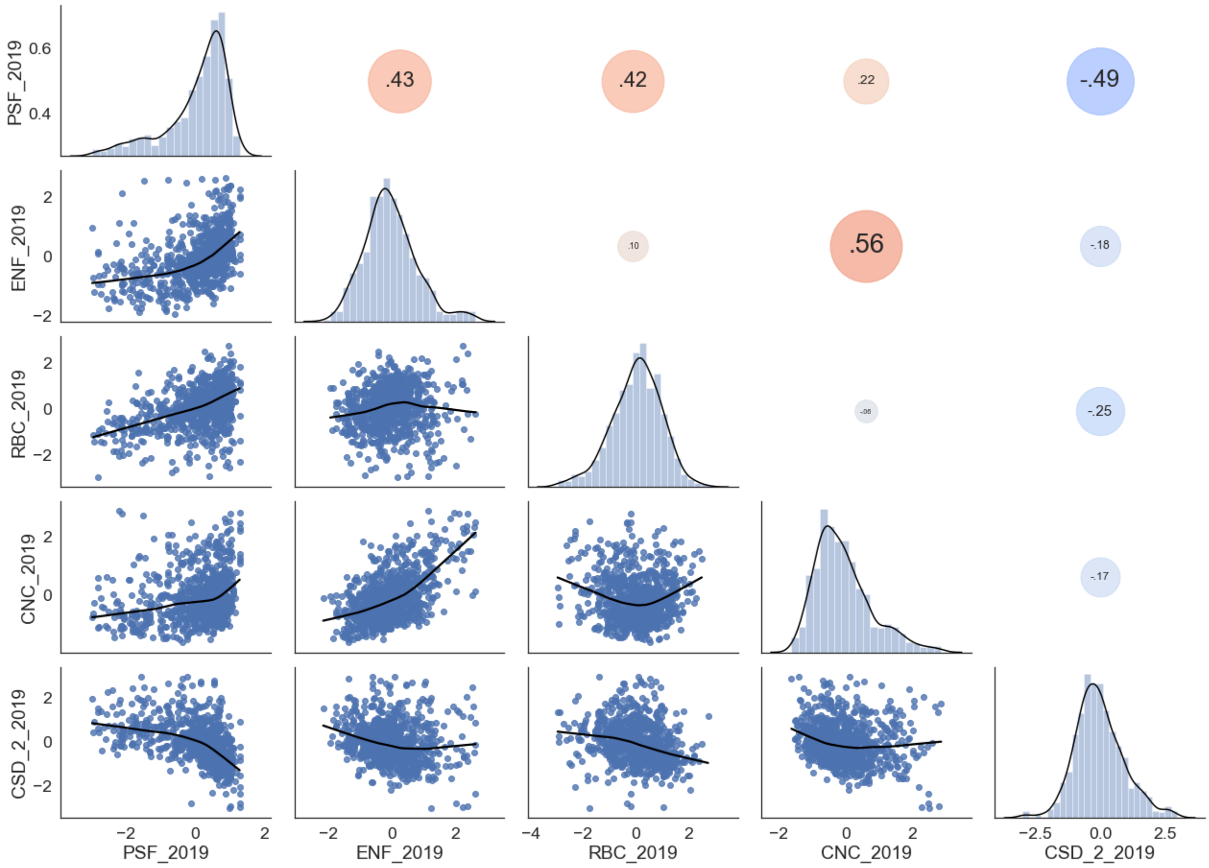


Figure 22. Sample of correlation testing for indicators under the scope of Economic Performance