Economic complexity and regional economic development: evidence from Brazil

PRELIMINARY DRAFT – 08/2022
ANPEC

Área 6 - Crescimento, Desenvolvimento Econômico e Instituições

João P. Romero – Cedeplar-UFMG
Elton Freitas – Cedeplar-UFMG
Fabrício Silveira – Cedeplar-UFMG
Gustavo Britto – Cedeplar-UFMG
Fernanda Cimini – Cedeplar-UFMG
Frederico G. Jayme Jr. – Cedeplar-UFMG

Abstract
The paper provides three contributions to the literature on economic complexity and regional development using data from Brazilian municipalities. First, it reports econometric tests of the relationship of regional economic complexity, calculated using employment data, and the logarithms of GDP per capita and of formal employment per capita. These tests, which are inspired in the ones carried out by Hausmann *et al.* (2014), transpose their results to the regional level and expand them to take into account the relationship between economic complexity and employment as well. Second, it proposes a new method to rank promising activities to be targeted by regional development policies, combining different indicators, as proposed by Hausmann *et al.* (2017), but using weights estimated through a principal component methodology. The results indicate that the proposed rule for formulation of smart diversification strategies performs very well when compared to regions that presented increases in their economic complexity. This methodology is illustrated using the example of the city of Belo Horizonte. Using the estimates of the relationship between economic complexity, income and employment, the paper presents simulations of the potential gains to be obtained following the proposed development strategies.
Keywords: Economic Complexity; Economic Growth; Regional Development.
J.E.L.: O11; O47; R11.

Resumo
O artigo apresenta três contribuições à literatura sobre complexidade econômica e desenvolvimento regional usando dados de municípios Brasileiros. Primeiro, o artigo reporta testes econométricos da relação entre complexidade econômica regional, calculada usando dados de emprego, e o logaritmo do PIB per capita e do emprego formal per capita. Estes testes, inspirados nos testes realizados por Hausmann *et al.* (2014), transpõem os resultados encontrados pelos autores para o nível regional, expandindo-os para levar também em consideração a relação da complexidade com o emprego. Segundo, o artigo propõe um novo método de ranquear atividades promissoras para serem fomentadas via políticas de desenvolvimento regional, combinando diferentes indicadores, como proposto por Hausmann *et al.* (2017), mas usando pesos estimados através da metodologia de componentes principais. Os resultados indicam que a regra proposta para formulação de estratégias de diversificação inteligentes tem bom desempenho se comparada às regiões que apresentaram crescimento da sua complexidade. A metodologia é ilustrada usando o exemplo da cidade de Belo Horizonte. Usando as estimativas da relação entre complexidade, renda e emprego, o artigo apresenta então simulações dos ganhos potenciais a serem obtidos ao serem implementadas as estratégias de desenvolvimento propostas.
Palavras-chave: Complexidade Econômica; Crescimento Econômico; Desenvolvimento Regional.
J.E.L.: O11; O47; R11.

Economic complexity and regional economic development: evidence from Brazil

1. Introduction

Economic development is intrinsically related to structural change. In its basic meaning, structural change means changing the productive structure of the economy. In the classical literature of economic development, this change means increasing the share of manufacturing in the economy, while reducing the share of agriculture (e.g. Schumpeter, 1934; Prebisch, 1961; Furtado, 1964; Hirschman, 1958; Kaldor, 1966). Structural change, therefore, involves leaning and mastering new economic activities.

As time went by, it became clear that increasing manufacturing production was not enough to guarantee economic development. As technology evolved, some sectors became more science-based than others and markets for manufacturing products changes markedly. While some manufacturing industries became more widespread across the globe, others remained heavily concentrated in a few companies and locations. Consequently, modern approaches to economic development started stressing the importance of moving into high-tech manufacturing for sustainable development (e.g. Nelson, 1992; Lundval, 1993; Romero and Britto, 2017).

With the rapid changes in science and technology observed in the past few decades, high-tech manufacturing industries became increasingly more heterogeneous, requiring highly specialized knowledge. The pathbreaking work of Hidalgo et al. (2007) explored fine-grained international trade data to build a network that interconnects products according to the probability of competitive co-production. This network, the *Product Space*, indicates the proximity of the productive knowledge required to produce each pair of goods. Moreover, it makes it clear that development is heavily path dependent due to the differences in accumulated knowledge between economies. As Britto et al. (2019) have shown, the shape of the *Product Space* has gradually changed through time, with clusters of products becoming clearer and more separated. This illustrates the increase in specialized knowledge that led to higher separations between manufacturing fields.

Exploring even further the information contained in disaggregate international trade data, Hidalgo and Hausmann (2009) showed that the ubiquity of the competitive production of different goods varies markedly. Furthermore, they also showed that the level of diversification of each economy is associated with its level of income per capita. Combining these two raw measures, the authors created the product complexity index (*PCI*) and the economic complexity index (*ECI*). The former indicates the amount of productive knowledge required to produce each good competitively. The latter indicates the amount of productive knowledge available in each economy.

Hausmann et al. (2014) provided evidence that indicates that increasing economic complexity predicts considerably higher growth rates of income per capita in the future, even after controlling for several additional variables. According to their estimates, an increase of one standard deviation in economic complexity is associated with a subsequent acceleration of 1.6 percent per year in the country's long-term growth rate.

Based on this literature, recent studies have been seeking to use indicators based on the economic complexity methodology to guide the formulation of development policies. Hausmann and Chauvin (2015) used a series of indicators constructed based on economic complexity and on relatedness between products to identify promising sectors for the development of Rwanda. Hausmann *et al.* (2017) used a similar methodology to identify the promising sectors for the development of Panama.

Nonetheless, it is not straightforward to adapt the economic complexity methodology to formulate regional development strategies. At the regional level exports are not as informative as they are at the national level, as sales to other regions within the same country are not be computed. Furthermore, economic interactions between neighbors are stronger at the regional level, making knowledge spillovers more relevant.

Due to these issues, economic complexity has been applied at the regional level using alternative measures of local knowledge. Balland et al. (2018), for example, use patent data to measure local technological knowledge, using the same methods proposed by Hidalgo et al. (2007) and Hausmann et al. (2014) to guide the formulation of regional smart specialization strategies. Alternatively, Gao et al. (2021) apply the economic complexity methodology using employment data from both China and Brazil, to show that knowledge spillovers are relevant at the regional level, and that improving transport infrastructure helps increasing these spillovers and the productive diversification they foster.

In this paper we provide four important contributions to the literature on economic complexity and regional development using data from Brazilian regions. First, we construct a network, the *Activity Space*, that links different activities based on the number of shared occupations. This network allows transposing the *Product Space* to the regional level using employment data. Second, we report econometric tests of the impact of regional economic complexity, calculated using employment data for Brazil, on the growth rates of GDP per capita and of formal employment share. These tests, which are inspired in the ones carried out by Hausmann *et al.* (2014), transpose their results to the regional level. To the best of our knowledge, this is the first paper to perform such tests. Third, we propose a new method to rank promising activities to be targeted by regional development policies, combining different indicators, as proposed by Hausmann *et al.* (2017), but using weights estimated using a principal component methodology. This methodology is illustrated using the example of the Brazilian city of Belo Horizonte. Combining the estimates of the impacts of economic complexity on income and employment, we present simulations of the potential gains to be obtained following the proposed development path.

2. Regional economic complexity and the *Activity Space*

2.1. Economic Complexity

The seminal paper of Hidalgo and Hausmann (2009) proposed to calculate products' and countries' complexity based on information on the diversification of economies and on the ubiquity of products. The level of diversification of each country is defined as the number of products the country produces with Revealed Comparative Advantage (*RCA*), while the level of ubiquity of each good is defined as the number of countries that produce the good with *RCA*. Formally:

$$RCA_{cp} = \frac{x_{cp}/\sum_p x_{cp}}{\sum_c x_{cp}/\sum_c \sum_p x_{cp}} \tag{1}$$

$$Diversification = k_{c,0} = \sum_p M_{cp} \tag{2}$$

$$Ubiquity = k_{p,0} = \sum_c M_{cp} \tag{3}$$

where x denotes the export quantum, while subscripts c and p denote country and product, respectively. M is a *dummy* variable which equals one if country c exports the good p with *RCA*, and zero otherwise. A *RCA* above one indicates that the country is competitive in the production of the good, while the opposite holds if the index is below one.

Using these measures, Hidalgo and Hausmann (2009) provided evidence that there is a strong positive correlation between income per capita and diversification. In addition, they showed that diversification and ubiquity are negatively correlated, which suggests that countries that are more diversified tend to produce goods that are less ubiquitous.

According to Hidalgo and Hausmann (2009), the competitive production of different types of goods requires different capabilities. Hence, the capabilities present in a country determine the goods it can produce and how difficult it is for the country to start producing goods that require different (or additional) capabilities. As a result, the range of goods a country can produce competitively and the level of complexity of these goods indicates the capabilities a country possesses. The negative correlation between diversification and ubiquity corroborates the idea that countries that possess a high number of capabilities: (i) are more diversified, since they are able to produce a high number of goods competitively; and (ii) produce goods that have lower ubiquity, since the higher number of capabilities enables them to produce more complex goods.

Based on the information from these indexes, Hidalgo and Hausmann (2009) calculate a Product Complexity Index (*PCI*) and an Economic Complexity Index (*ECI*). The intuition for combining the two indexes is straightforward. On the one hand, a country with high diversification is considered less complex if the products it produces competitively (with *RCA*) present high ubiquity. On the other hand, a product with low ubiquity is considered more complex if it is produced by countries that are very diversified. Consequently, by repeating this process and performing continuous iterations between the two indexes it is possible to extract progressively more refined information about the economic complexity of each product and country. Formally:

$$k_{c,N} = (1/k_{c,0}) \sum_p M_{cp} k_{p,N-1} \tag{4}$$

$$k_{p,N} = (1/k_{p,0}) \sum_c M_{cp} k_{c,N-1} \tag{5}$$

where N denotes the number of iterations.

Substituting (4) into (5) yields:

$$k_{c,N} = \sum_{c'} \widetilde{M}_{cc'} k_{c',N-2} \tag{6}$$

where $\widetilde{M}_{cc'} = \sum_p (M_{cp} M_{c'p})/(k_{c,0}\, k_{p,0})$ and c' denotes other countries besides c.

Equation (6) is satisfied when $k_{c,N} = k_{c,N-2} = 1$, which is the eigenvector associated with the highest eigenvalue of $\widetilde{M}_{cc'}$. However, since this eigenvector is formed of ones, he is uninformative. Hence, the eigenvector associated with the second highest eigenvalue ($\vec{K}$) is used to capture highest part of the system's variance. Thus, *ECI* is calculated as:

$$ECI = (\vec{K} - < \vec{K} >)/sd(\vec{K}) \tag{7}$$

where < > denotes the average, and *sd* denotes the standard deviation.

The same procedure is used to calculate *PCI*, but now substituting (5) into (4) and using the eigenvector associated with the second highest eigenvalue ($\vec{Q}$) of $\widetilde{M}_{pp'}$:

$$PCI = (\vec{Q} - < \vec{Q} >)/sd(\vec{Q}) \tag{8}$$

Following their seminal results, a series of papers explored other effects of economic complexity. Hausmann et al. (2014) provided evidence that indicates that increasing economic complexity predicts considerably higher growth rates and levels of GDP per capita in the future, even after controlling for a series of additional variables. Hartmann et al (2017), found evidence that economic complexity is correlated with lower income inequality. Mealy and Teytelboym (2020) and Romero and Gramkow (2021) found strong evidence indicating that increasing economic complexity contributes to reduce greenhouse gas emissions and other environmental impacts.

2.2. Product Space

The pathbreaking paper of Hidalgo *et al.* (2007) investigated whether the sectoral composition of each country's competitive exports influences the path, the costs and the speed of it's structural transformations. In this paper, therefore, the authors explore the idea that each country's productive structure influences it's growth and development possibilities, stressing the path-dependence of knowledge and capabilities accumulation.

Hidalgo *et al.* (2007) established how close products are in terms of the capabilities required for their competitive production using the conditional probabilities of exporting each pair of goods with *RCA*. Hence, this method assumes that the probability of producing two products that require similar capabilities for their competitive production is higher than the probability of producing two goods that require different capabilities.

The authors explored the large amount of information in international trade data to calculate the *proximity* between goods as the probability of a country exporting product p with RCA given that it exports product j with RCA as well. The level of proximity between two products (p and j) is given by the authors as:

$$\varphi_{p,j} = min\{P(RCA_j = 1|RCA_p = 1), P(RCA_p = 1|RCA_j = 1)\} \quad (9)$$

Adopting a threshold value for proximity, the authors established the active links between products, creating a network that they called *Product Space*. In this network, therefore, products that use similar capabilities for their competitive production tend to for clusters. Moreover, complex products tend to be located towards the centre of the network, while least complex product tend to locate more towards the fringes of the network.

Using the *Product Space*, Hidalgo *et al.* (2007) showed that, on average, less developed countries produce goods with a limited number of links. This restricts these countries' diversification possibilities, making in more costly for these countries to move towards the production of more complex products, while the opposite holds true for developed countries. Hence, the authors' seminal paper provided three important empirical contributions to the economic development literature. They showed that different countries face different opportunities to diversify their economies and increasing their economic growth, given their distinct productive structures and associated capabilities. Consequently, they highlighted that structural change is highly path dependent. And finally, these results point out that achieving competitiveness in the production of complex goods takes time, since it requires learning new capabilities (Hidalgo *et al.*, 2007: 487).

Seeking to explore the implicit information contained in the *Product Space*, Hausmann et al (2014) developed an indicator that measures the ease of acquiring competitiveness in any given industry as a function of existing capacities in the economy. This indicator would help to identify new diversification possibilities based on the implied costs associated with the acquisition of the new capabilities required for performing new activities competitively.

Assuming that close products use similar productive capacities, Hausmann et al (2014) proposed an index that measures the ease of competitive production of a given good as a function of the competitive production of nearby goods, which serves as a proxy for existing capabilities. This index, called Product Density Index (*PDI*), measures the proximity of a given product to the country's current production structure (products with *RCA*), indicating the difficulty (or cost) of achieving *RCA* in this product. Hence, this measure reflects the amount of new productive knowledge that a region needs to acquire to be able to produce and export a particular product with *RCA*. In other words, the smaller the *PDI*, the more capabilities will have to be acquired and the longer and more difficult/costly will be the process of achieving *RCA* in this product. In this way, products that the country exports without *RCA* but that have a high *PDI* appear as products with high potential to be fostered.

PDI is calculated as the sum of the proximities ($\varphi_{p,j}$) of the products that the country exports with *RCA* in relation to the product p. The index is normalized by the sum of the proximities between all products in the network in relation to the product p. Formally:

$$PDI_{pct} = \frac{\sum_p M_{ict}\,\varphi_{pi}}{\sum_p \varphi_{pi}} \quad (8)$$

Hausmann et al (2014) proposed also a second indicator, called Opportunity Gain Index (*OGI*). This index measures the gains that achieving *RCA* in any given good brings in terms of facilitating the production of more complex goods that were not previously produced/exported competitively in this economy. Formally:

$$OGI_{pct} = \sum_p \frac{(1 - M_{cit})\varphi_{pi} PCI_{it}}{\sum_p \varphi_{pi}} - (1 - PDI_{pct})PCI_{\ pt} \quad (9)$$

A high value in the *OGI*, therefore, indicates that the product under investigation is closer to complex products. Hence, this index can be used to devise policies that aim to increase an economy's economic complexity considering several rounds of diversification into progressively more complex goods.

2.3. Activity Space

At the regional level using export data to measure economic complexity is highly problematic, since transactions between regions within the same country are not computed. Furthermore, economic interactions between neighbors are stronger at the regional level, making knowledge spillovers more relevant. In addition, in regions services play a more prominent role, so that it becomes more relevant to take these activities into account.

To solve these issues, some studies have been using employment or patent data instead of trade data to calculate the indicators of economic complexity. Using employment data has one additional advantage: it makes possible to use information on the number of occupations within companies or regions to measure proximity through occupation similarities.

According to Farjoun (1994), companies diversify through networks of industries that interrelate according to the resources they use. Thus, it is important to observe the similarities between the resources being used (for example, the level of education in different occupations) in order to understand the diversification patterns of companies and regions. From these *Resource-Related Industry Groups*, companies are able to share and transfer similar resources to benefit from and encourage diversification processes (Farjoun, 1994, p. 188).

Following Freitas (2019), from the concept of co-occupation it is possible to estimate the proximity of activities with similar jobs and build the complexity indicators through employment data. First, we define the indicator of effective occupations (*EO*), analogous to the revealed comparative advantage index (*RCA*), as the basis for calculating the complexity indicators using employment data. Formally:

$$EO_{s,o} = \frac{emp_{s,o}/emp_s}{emp_o/emp} \tag{10}$$

where $emp_{s,o}$ is the employment of occupation o in sector s and emp_s is the total employment of sector s in the country. In addition, emp_o is the total employment of the occupation o in the country and emp is the total employment in the country.

Thus, if the *EO* indicator is equal or greater than one, the participation of occupation o in sector s is greater than the participation of occupation o in the country. Hence, the sector in question effectively employs such an occupation. Otherwise, if *EO* is less than one, the conclusion is that the sector does not effectively employs this occupation in the analyzed location.

Using the *EO* indicator it is possible to calculate the proximity between activities and for an *Activity Space*. Proximity is calculated as the probability of an industry employing a certain occupation, given that another industry also employs that occupation. This represents a different way, although similar, of measuring the similarities between industries in terms of occupations. Thus, following Freitas (2019), equation (9) can be adapted to establish the relationship between activities s and i replacing *RCA* by *EO*.

The proximity between activities calculated according to similarities in the resources used by companies has a similar interpretation in relation to the concept of co-occurrence of productive capacities in nearby industries, as proposed by Hidalgo *et al.* (2007). In fact, one of the stylized facts presented by Hidalgo and Hausmann (2009) is the positive correlation between the number of occupations and countries' economic complexity.

Finally, it is important to note that using the *EO* indicator it is also possible to calculate the economic complexity of each region and of each activity following the same methodology described in the previous section.

2.4. Economic Complexity at the regional level

The literature on economic complexity and regional development has been growing rapidly in the last few years. Economic complexity has been applied at the regional level using alternative measures of local knowledge. Balland et al. (2018), for example, use patent data to measure local technological knowledge, using the same methods proposed by Hidalgo et al. (2007) and Hausmann et al. (2014) to guide the formulation of regional smart specialization strategies. Alternatively, Gao et al. (2021) apply the economic complexity methodology using employment data from both China and Brazil, to show that knowledge spillovers are relevant at the regional level, and that improving transport infrastructure helps increasing these spillovers and the productive diversification they foster.

The issue of structural change in regional and urban economics gained prominence through the reassessment of the works of Jane Jacobs and Alfred Marshall. Glaeser *et al.* (1992) provided important contributions to the framework of agglomeration economies, which previously focused on the effects of localization and urban size economies, by investigating the economic importance of urban diversity. This focus on the so-called Jacobs' externalities can be considered the first attempt to assess the effect of the local industrial structure on economic development. Henderson *et al.* (1995) took another step towards investigating structural change and growth studying whether externalities were more important in sustaining traditional industries than in attracting new industries. They found that new industries, especially the high-tech, entered diversified cities where Jacobs externalities were available, while mature industries benefited more from location externalities generated in more specialized cities. According to Simões and Freitas (2014), Jacobs externalities are more relevant in sectors of high technological intensity, while sectors with low and medium technological intensity are more benefited in medium-sized urban centers, relatively less diversified.

Jacobs' main argument for new industries benefiting from diversified urban economies was that urban diversity prompts the division of labor in the city. However, division of labor contributes to urban growth due to its effect on opportunities for innovation, and not so much due to its effect on technical efficiency. This fits very well into the Schumpeterian concept of innovation as successful new combinations of productive forces, i.e., old ideas.

Cognitive theory, however, emphasizes a trade-off between diversity and similarity. Although cognitive proximity (overlapping competencies) facilitates communication between agents, only the ones who do not share overlapping competencies and knowledge can really offer something new to be learned (Nooteboom, 2000).

Hence, the fact that social learning may require an optimal level of cognitive distance may explain why, after several empirical studies, the evidence on the effects of Jacobs externalities is still inconclusive (De Groot *et al.*, 2009). Regional knowledge spillovers only happen between certain industries, since a more effective communication is often hampered by the cognitive distance between these. Recently, several authors (Almeida & Kogut, 1999; Boschma & Frenken, 2009; 2011; Gilsing et al., 2007; Menzel, 2008) have suggested that industries are more likely to learn from each other when they are technologically related. Thus, a broad set of technologically related industries in a region should be more beneficial for growth than a diversified set of industries in a broad range of technological areas, given that it is the combination of distance and cognitive proximity that brings together the positive sides of diversity and similarity across industries.

Frenken *et al*. (2007) argue that regions with a greater degree of variety in related industries present more learning opportunities and, consequently, a higher dissemination of local knowledge. Using data for the Dutch economy, the authors show that regions with a higher degree of "related variety" often present higher employment growth. The same result was also found for other countries (Essletzbichler, 2005; Bishop & Gripaios, 2009).

Boschma and Iammarino (2009) argue that related variety can also flow from one region to another through commercial links between industries. Using regional trade data, the authors show that inter-regional knowledge flows are, in fact, associated with regional employment growth, when these come from industries related to industries in the region.

In these studies, the industrial base of a region is treated as a stable property. This makes sense in the short term because the industrial composition of a regional economy changes slowly over time. However, it is likely that the relationship between regional industries not only drives the incremental growth of existing industries through agglomeration economies but may be responsible for more radical changes in the regional productive structure. In fact, industrial kinship can be an important factor in attracting new industries to the region and in the disappearance of unrelated industries. This is a fundamental aspect because it sheds light on how the Schumpeterian process of creative destruction develops regionally over the long term. Similarly, understanding how new regional growth paths emerge has been repeatedly raised by economic geographers as one of the most intriguing and challenging questions of the field (Scott, 1988; Storper & Walker, 1989; Martin & Sunley, 2006). After all, the industrial history of regions is expected to affect how regional structures create new activities over time, and how they transform and restructure of their economies.

Bathelt & Boggs (2003) and Glaeser (2005) show that new local industries are often related to activities already established in the region. In addition, recently, there is more systematic evidence showing that territories are more likely to expand and diversify towards sectors that are closely related to their existing activities (Hausmann & Klinger, 2007; Hidalgo *et al*., 2007). Focusing on changes in export portfolios over time, Hausmann & Klinger (2007) showed that countries expanded their export mix, moving towards products that were related to their current export agenda, which implies that a country's position in the product space affects their opportunities for diversification. As a result, rich countries specialized in the more densely connected parts of the product space presented more opportunities to sustain economic growth than poorer countries.

Boschma & Frenken (2009) call the process by which new activities arise from technologically related industries by “regional ramifications”. The reason this process of regional branching takes place is that new industries can connect to existing ones through various knowledge transfer mechanisms. These mechanisms are: (i) diversification of firms; (ii) entrepreneurship in the form of spinoffs; (iii) mobility of workers; (iv) social networks. The branching process is essentially a regional phenomenon, as these mechanisms operate primarily – but not exclusively – at the regional level, that is, within subnational regions rather than across regions.

In their diversification strategies, firms tend to develop their previously existing competencies. The reason for this is that, as Nelson & Winter (1982) argue, intra-firm diversification is not simple, as companies seeking new markets and new technologies face fundamental uncertainties. However, companies try to limit these uncertainties and avoid large switching costs by carrying out local search processes in the technological sense, that is, aimed at technologies and markets similar to those in which the companies became known. Likewise, Penrose (1959) conceives firm growth as a progressive process of related diversification, in which firms diversify towards products that are technologically related to their current products. This opinion is supported by the fact that mergers and acquisitions present higher levels of performance when connecting companies with related technological knowledge bases (Piscitello, 2004; Cassiman *et al.*, 2005). Since new divisions of companies are often established within existing facilities, the internal diversification of companies is often not only local, in cognitive terms, but also in geographic terms.

In sum, there are good reasons why firm-level-related diversification (through internal and external growth) is geographically biased, although systematic empirical evidence for such a hypothesis seems to be scarcely discussed.

Regional diversification through entrepreneurship generally occurs when new companies in an emerging industry are created by entrepreneurs who have previously gained knowledge and experience in a related industry in the same region. There is considerable evidence that these companies benefit economically from the experiences acquired by entrepreneurs in related industries, reflecting in their greater chances of survival (Klepper, 2007). Longitudinal studies also confirm that these experienced entrepreneurs play a crucial role in the regional diversification process. Boschma & Wenting (2007) show that, in the early stages of development of the UK automobile industry, companies had a higher survival rate when those responsible had already worked in related industries, such as bicycle and bus assemblers or in the area of mechanical engineering, and when their regions stood out for the strong presence of these related industries. Kia's case also reflects this very well. Founded in 1944 in South Korea as a bicycle manufacturer, it later went on to produce military vehicles and equipment.

Regional diversification through labor mobility has not yet been so explored. Worker mobility is often considered as a key mechanism of knowledge diffusion (Almeida & Kogut 1999; Heuermann 2009), but until recently little attention had been paid to *spillovers* between firms in related industries with respect to labor mobility. Boschma *et al*. (2009) provide empirical evidence that the economic effects of workflows cannot be properly assessed without paying attention to how the knowledge flows are related to existing knowledge bases in companies. They point out that the entry of employees with skills related to the plant's knowledge base was positively associated with productivity increases, while the hiring new employees with skills already available at the plant was negatively associated with the productivity measure. However, the study of the role of labor mobility in the constitutive phases of an industry deserves further development. If, in fact, labor mobility induces industrial ramification, this phenomenon is likely a regional phenomenon, as most workers who change job remain in the same region (Boschma *et al*., 2009; Timmersmans & Boschma, 2013).

Social media can be another source of regional diversification. They are considered an important channel for the dissemination of knowledge and learning among companies (Powell *et al*., 1996, Sorenson *et al*., 2006; Ter Wal, 2009). However, the importance of networks for innovations, and thus for the development of new economic activities, may depend on the degree of technological kinship between network partners. It is likely that there is an optimal level of cognitive proximity between network partners in order to stimulate new ideas, and at the same time allow effective communication (Boschma & Frenken, 2009). Studies on networks of alliances between companies show that new knowledge is developed when actors bring different, but related, competences (Gilsing *et al*., 2007). Breschi & Lissoni (2003) state that social networks tend to be highly localized and that they can contribute to the process of regional diversification.

The implications of the above should not be underestimated. First, the industrial relationship at the regional level should affect how Schumpeter's creative destruction process will shape the economic landscape, i.e, the kinship between industries influences both incoming and existing industries that are going to leave a region. Second, the rise and fall of industries are conditioned to regional industrial structures established in the past, and this is supported by the notion of regional trajectory dependence (Rigby & Essletzbichler, 1997). Third, the trajectory dependency process implies that there is some degree of coherence in the region's industrial profile. However, this coherence is constantly redefined through the process of creative destruction. The entry of new industries in a region, although technologically related to existing local industries, is likely to inject new variety into the region, which reduces technological coherence. In contrast, exit from existing industries increases the industrial coherence of regions, because unrelated industries are more likely to be selected, leading to a decrease in variety.

3. Empirical Investigation

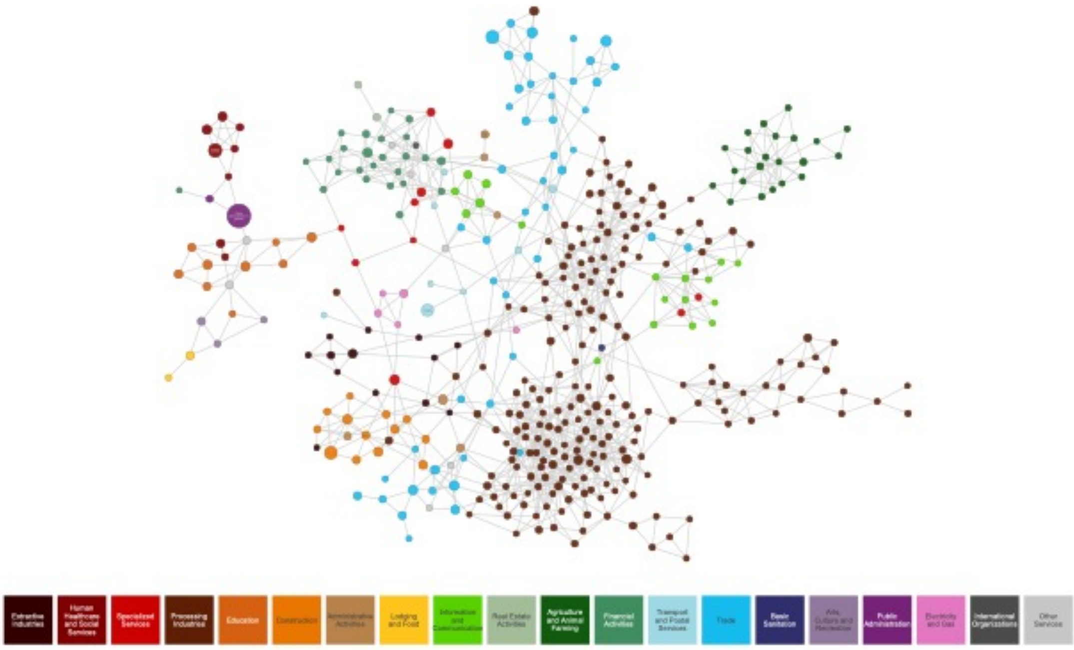
3.1. Database

The empirical investigation presented in this paper is based on employment data for the municipalities and microregions of Brazil over the period 2006-2017. Employment data used for the construction of the *Activities Space* and of the economic complexity indicators of local activities come from RAIS (Annual Report of Social Information), of the Brazilian Ministry of Employment. The database provides information on the number of employees by economic activity classified according to the CNAE (version 2.0) classification and by occupation within each activity at both regional levels explored here. GDP per capita was gathered from the Brazilian Institute of Geography and Statistics (IBGE). The CNAE classification comprises 672 activity classes from all sectors (services, manufacturing and agriculture). Moreover, employment is divided into 596 occupations within each activity. The database covers the 5570 Brazilian municipalities, which are grouped into 558 microregions.

3.2. Brazilian Activity Space

Using the employment information from the RAIS database it is possible to build the *Activity Space* for Brazil following the methodology presented in section 2.3. The network, presented in Figure 1, shows that manufacturing activities (in brown) and modern services (in dark green) are more connected and localized more towards the center of the network, highlighting their importance for diversification and development.

Figure 1: *Activity Space* for Brazil



Source: Auhtors own elaboration based on employment data from RAIS.

3.3. Econometric Specification

To analyze the impact of the *ECI* on economic growth at the regional level, we estimate a set of regressions in which the dependent variable is either the annualized growth rate of GDP per capita, or the growth rate of the share of formal employment in total population.

The estimated equation for GDP per capita follows the specification used by Hausmann et al (2014), but in a panel data form:

$$\log(y)_{i,t} = \alpha + f_i + f_t + \beta_1 ECI_{i,t-1} + \beta_2 L.ECI_{i,t-1} * \log(y)_{i,t-1} + \beta_3 \log(y)_{i,t-1} + \beta_i \log(x)_{i,t-1} + \varepsilon \qquad (11)$$

where y is the GDP per capita, f are fixed effects for individuals i (municipalities/microregion) and periods t (annual dummies), α is the constant, ε is the residuals. Among the explanatory variables are the initial GDP per capita (log) and a multiplicative variable between *ECI* and GDP per capita. The former seeks to capture the effect of the hypothesis of convergence or technological catch-up. The multiplicative term seeks to capture the non-linearity of the effect of *ECI* on the growth rate. Hypothetically, this effect is negative because the potential gains from increasing *ECI* also reduce with the increase in GDP per capita and *ECI* over time. Finally, x indicates three control variables introduced in the tests: (i) population; (ii) the percentage of workers with secondary education; and (iii) the percentage of workers with tertiary education.

The explanatory variables are introduced with a lag to assess if they predict increases in GDP per capita in the subsequent period. Although this does not solve potential simultaneity, it provides a stronger indication of the relevance of each variable.

The impact of *ECI* on the evolution of formal employment was evaluated estimating the following equation, analogous to equation (11):

$$\log(l)_{i,t} = \alpha + f_i + f_t + \beta_1 ECI_{i,t-1} + \beta_2 ECI_{i,t-1} * log\,(l)_{i,t-1} + \beta_3 log\,(l)_{i,t-1} + \beta_i \log(x)_{i,t-1} + \varepsilon \qquad (12)$$

where l is the ratio of formal jobs in the economy in relation to population. The interpretation of the variables is the same as in equation (11).

3.4. Main Results

The estimation results for equation (11) are presented in Table 1. Columns (i) to (iv) present the results for municipalities. ECI is always positive and highly significant. Most importantly, the models corroborate all assumed hypotheses, including very similar parameters for the ECI at the municipality level to those of the seminal study by Hausmann *et al.* (2014) at the country level and using export data. This result is important as it validates the use of employment data in the assessment of economic complexity. The interaction between ECI and GDP per capita is only significant when population is introduced, and its effect, although negative (as expected) is very small. It is interesting to note that population enters with a positive sign, which indicates that at the regional level, increasing the city's population is associated with future increase in its GDP per capita. The lag of GDP per capita is positive, but lower than one, which corroborates the hypothesis of technological absorption when the log of GDP per capita is used as the dependent variable. The share of employees with tertiary education is not significant.

Table 1: Economic Complexity and GDP per capita

	Municipalities				Microregions			
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
ECI, lag	0.021***	0.064**	0.074***	0.073***	0.030***	0.185***	0.177***	0.175***
	(0.004)	(0.027)	(0.026)	(0.027)	(0.009)	(0.054)	(0.052)	(0.053)
Log of GDP per	0.580***	0.580***	0.584***	0.582***	0.628***	0.624***	0.635***	0.635***
capita, lag	(0.012)	(0.012)	(0.012)	(0.012)	(0.029)	(0.030)	(0.030)	(0.030)
(ECI * Log of GDP		-0.005	-0.006**	-0.006*		-0.014***	-0.013***	-0.013**
per capita), lag		(0.003)	(0.003)	(0.003)		(0.005)	(0.005)	(0.005)
Log of Population, lag			0.046***	0.041**			0.045	0.045
			(0.017)	(0.017)			(0.047)	(0.047)
Log of Perc. emp. with				0.000				0.002
tert. educ., lag				(0.001)				(0.010)
Constant	3.782***	3.781***	3.632***	3.661***	4.188***	4.235***	3.972***	3.975***
	(0.104)	(0.103)	(0.128)	(0.127)	(0.321)	(0.331)	(0.380)	(0.381)
Observations	61213	61213	61213	60975	6138	6138	6138	6138
Adjusted R2	0.870	0.870	0.870	0.870	0.914	0.914	0.914	0.914

Note: The dependent variable is the Log of GDP per capita. Robust standard errors in parenthesis. All models are estimated introducing region and year fixed effects. Significance: *=10%; **=5%; ***=1%.
Source: Authors' elaboration.

Columns (v) to (viii) of Table 1 present the results for microregions. The results are similar, but the coefficient of ECI is larger as well as of the interaction with GRP per capita, while population loses its significance. These estimates reinforce the robustness of the results found for municipalities.

The estimation results for equation (12) are presented in Table 2. Once again columns (i) to (iv) present the results for municipalities. ECI is positive and highly significant in column (i), but when the interaction of ECI with Employment per capita is introduced, it enters with a positive and significant coefficient, while the coefficient of ECI becomes negative and not significant. This is probably due to a high correlation between the two variables. When the interaction is replaced by other control variables, the effect of ECI on employment per capita becomes positive and significant once again. Population enters with a negative sign, which indicates that at the regional level, increasing the city's population is associated with future decrease in Employment per capita. Contrary to what was expected, the share of employment with tertiary education is negatively associated with Employment per capita.

Table 2: Economic Complexity and Employment per capita

	Municipalities				Microregions			
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
ECI, lag	0.064***	-0.086	0.066***	0.058***	0.037***	-0.054	0.039***	0.039***
	(0.009)	(0.058)	(0.009)	(0.009)	(0.011)	(0.115)	(0.012)	(0.012)
Log of GDP per capita, lag	0.230***	0.255***	0.223***	0.230***	0.399***	0.404***	0.348***	0.348***
	(0.017)	(0.015)	(0.018)	(0.015)	(0.050)	(0.052)	(0.058)	(0.058)
(ECI * Log of GDP per capita), lag		0.034***				0.014		
		(0.013)				(0.018)		
Log of Population, lag			-0.198***	-0.191***			-0.217***	-0.217***
			(0.037)	(0.033)			(0.064)	(0.064)
Log of Perc. emp. with tert. educ., lag				-0.015***				0.000
				(0.005)				(0.011)
Constant	3.488***	3.360***	4.011***	3.930***	4.097***	4.056***	5.122***	5.123***
	(0.076)	(0.069)	(0.140)	(0.129)	(0.332)	(0.351)	(0.565)	(0.564)
Observations	61215	61215	61215	60977	6138	6138	6138	6138
Adjusted R2	0.163	0.164	0.164	0.166	0.454	0.454	0.462	0.462

Note: The dependent variable is the Log of Employment per capita. Robust standard errors in parenthesis. All models are estimated introducing region and year fixed effects. Significance: *=10%; **=5%; ***=1%.
Source: Authors' elaboration.

Columns (v) to (viii) of Table 2 present the results for microregions. The results are once again similar, but the coefficient of ECI is smaller and the interaction with Employment per capita enters with a positive coefficient but not significant. Population is still negative and significant, while the share of employment with tertiary education loses its significance. Overall, therefore, it is possible to consider the estimates for microregions slightly more robust. This is not unexpected, since these regions are defined based on the economic interaction within them.

The results presented in Table 2 indicate the fundamental role of ECI in increasing the share of formal employment in the labor market. In other words, it is possible to interpret that the increase in the complexity of the municipality or microregion is associated with the increase in the quality of jobs generated in the location.

4. Complexity based diversification strategies

4.1. Economic Diversification Score

Taking the economic complexity literature as reference, several studies have sought to use the indicators discussed in the previous section to devise diversification policies. Hausmann and Chauvin (2015), for example, used a series of indicators constructed based on economic complexity and on relatedness between products to identify promising sectors for the development of Rwanda. Hausmann *et al.* (2017) and Romero and Freitas (2018) used a similar methodology to identify the promising sectors for the development of Panama and Brazil, respectively.

These studies gathered a series of variables related to three relevant dimensions related to each product not exported with *RCA*: (i) *Markets*; (ii) *Capabilities*; and (iii) *Gains* (in terms of economic complexity). The variables were then normalized and put together in a weighted sum to generate a score of promising industries to be fostered through development policies. The main problem with this methodology is that the weights attributed to each variable are completely arbitrary. Hence, in this paper, we seek to solve this problem and improve the economic complexity smart diversification score by using a Principal Component Analysis (PCA) method.

Brazilian foreign trade data were gathered from SECEX and UN COMTRADE classified by product according to the Mercosur Common Nomenclature (NCM) classification. A correspondence table provided by IBGE was used to associate NCM products to economic activities classified by CNAE (version 2.0). Four-year averages of the variables were calculated to smooth short-term fluctuations and reduce eventual measurement errors. Year 2006 was dropped from the analysis.

The procedure carried out to estimate the weights of the variables followed a series of steps. First, sub-samples of municipalities/microregions were defined based on the characteristics of their productive structures in terms or High/Low Density and Complexity. Second, the weights attributed to each variable within each dimension (*Capabilities*, *Markets* and *Gains*) were defined by applying the PCA method. Third, after defining these weights, two scores were generated, one with homogeneous weights, and the other with the weights calculated via PCA. Fourth, these scores were used to try to analyze the proportion of sectors that gained *RCA* in municipalities/microregions that showed an increase in their *ECI* during the period evaluated predicted by each score, in order to validate their quality.

In this paper, for each dimension a PCA was performed. It was decided to keep the "n" components generated that cumulatively explained at least 80% of the variance of each one of the dimensions. Principal Component Analysis (PCA) has the main objective of explaining the structure of variance and covariance of a random vector, composed of p-random variables, through the construction of linear combinations of the original variables. It allows the researcher to reorient the data so that the first few dimensions explain as much information as possible. The individual main components are linear combinations of random variables X1, X2, Xn. Geometrically, these linear combinations represent the selection of a new coordinate system obtained by rotating the original system with X1, X2, Xn as the coordinate axes. Hence, PCAs are used to discover and interpret dependencies that exist between variables and to examine relationships that may exist between individuals.

To assess the improvements proposed in this research for the identification of promising activities, we built three versions of the *EDS*, considering: (i) all variables with equal weight; (ii) all variables with weights defined by the PCA method; (iii) eliminating some of the variables and considering the weights defined by the PCA method.

Table 2: Variables considered in the Economic Diversification Scores (*EDSs*) and their weights

Dimensions	Weights	Indicators	Models tested		
			All variables with equal weights	All variables with weights via PCA	Selected variables with weights via ACP
Capabilities	0.33	Employment	0.333	0.359	0.359
		Revealed Comparative Advantage (*RCA*)	0.333	0.583	0.583
		Employment growth rate	0.333	0.058	0.058
Markets	0.33	Imports (municipality/microregion)	0.111	0.104	0.088
		Revealed Comparative Disadvantage (*RCD*) (municipality/microregion)	0.111	0.04	0.157
		Import growth (municipality/microregion)	0.111	0.042	0.173
		Imports (state)	0.111	0.12	-
		Revealed Comparative Disadvantage (*RCD*) (state)	0.111	0.093	-
		Import growth (state)	0.111	0.216	0.583
		Imports (country)	0.111	0.132	-
		Revealed Comparative Disadvantage (*RCD*) (country)	0.111	0.102	-
		Import growth (country)	0.111	0.15	-
Gains	0.33	Product Density Index (*PDI*)	0.25	0.216	0.216
		Product Complexity Index (*PCI*)	0.25	0.293	0.293
		Opportunity Gain Index (*OGI*)	0.25	0.271	0.271
		Number of connections in the network	0.25	0.22	0.22

Source: Authors own elaboration.

Table 2 shows the variables considered to calculate the Economic Diversification Score (*EDS*) and the results of the weights in the three versions indicated above. The PCA was carried out considering the sample of 593 municipalities belonging to the High Density (average *PDI*>median *PDI*) and High Complexity (*ECI*>0) quadrant and which presented growth in the *ECI* in the three periods analyzed in this research. To calculate the weights, the variables referring to the most recent period available in the database were used, from 2015 to 2018.

To assess the predictive capacity of the *EDSs* proposed here in identifying promising activities, we performed a simple validation test. First, we created rankings for the sectors identified as promising by each of the three versions of the *EDS* in period 1. Then, we selected the industries in which the municipalities that have increased their *ECI* did not have *RCA* in period 1 (2007 to 2010) and achieved *RCA* in period 3 (2015 to 2018). Finally, from the number of activities found in the previous step, we verified how many of them were well ranked in each of the rankings built by the three versions of the *EDS*. The limit was defined by the number of sectors that transitioned from *RCA*<1 to *RCA*>1. For example, if a municipality X had 14 sectors that reached *RCA*>1 in period 3, we evaluated the first 14 sectors indicated in each of the rankings built to identify how many of the activities that achieved *RCA* were identified as promising by each rule. Similarly, if a municipality Y had 30 sectors that reached *RCA*>1 in period 3, we verified the first 30 activities indicated in each of the rankings. From these comparisons we calculated the average rate of success of each score as the ratio of the number of activities found in the final step of the process in relation to the total number of new activities with *RCA* (for each municipality or microregion).

Table 3 shows the mean percentages of correct predictions for each of the scores. Column 1 shows the percentages of the 3 different rules if applied to all municipalities. All 3 rules have a similar rate of correct predictions, around 30%, which indicates that about 1 in every 3 activities that actually obtained *RCA* in period 3 were targeted as promising by the rules. Column 2 presents the percentages only for the sectors that gained *RCA* in period 3 and that had *PCI*>0, that is, above average. As can be seen, in this case the results improve considerably: rule 3 presents a correctly predicts 41.4% of the activities, 3.4 percentage points better than rule 1, and 2 percentage points better than rule 2. The same pattern is observed in columns 3 and 4. The difference is that now the rates are evaluated considering only the municipalities with High Density and High Complexity, with was the subsample used to calculate weights using PCA.

Table 3: Validation tests of the EDSs

	Correct predictions			
	Test 1	Test 2	Test 3	Test 4
	All municipalities	All municipalities with *PCI*>0	All municipalities in the same group as Belo Horizonte	All municipalities in the same group as Belo Horizonte with *PCI*>0
Ranking 1	31.60%	38.00%	30.60%	35.70%
Ranking 2	30.50%	39.40%	30.00%	37.00%
Ranking 3	30.20%	41.40%	29.10%	37.00%
Number of observations	4440	1033	635	362

Source: Authors own elaboration.

The results suggest the importance of the refinements carried out in the research, indicating that the use of weights calculated via PCA with sub-samples of municipalities/microregions similar to the region under investigation increased the proportion of sectors predicted as promising by the scores and that effectively achieved RCA in municipalities/microregions were an increase in economic complexity was observed.

4.2. Alternative strategies

To improve even further the smart diversification strategies stablishing following the *ODS* proposed in the previous section, activities were classified according to a series of strategies, following a classification proposed by the Brazilian National Export Plan (2015-2018), namely: (i) *Maintenance*, (ii) *Consolidation*, (iii) *Recovery*, and (iv) *Betting*. The scores were used to rank the most promising sectors within each of these strategies. The purpose of this ranking is to bridge the gap between the proposed methodology for the selection of promising sectors and different smart diversification strategies, based on some extra information about characteristics of local competitiveness and of sectoral market dynamics.

The four strategies presented in Table 4 offer a comprehensive view of the products and activities with global, national and municipal competitiveness according to the smart diversification score proposed in the previous section. Through the typology presented in Table 4 it becomes possible to outline short, medium, and long-term development strategies.

Table 4: Smart diversification strategies – subgroups

Subgroups	Definition	Parameters
Maintenance	Sectors that are well positioned in the market and have a comfortable situation in relation to their main competitors	*RCA* >= 1,5 and municipal and national employment growth > 0
Consolidation	Sectors that are not yet consolidated but are growing at a pace close to or above that of their competitors.	0,5 =< *RCA* < 1,5 and Municipal and national employment growth > 0
Recovery	Sectors that have not yet consolidated or products that were once consolidated but have been reducing their market share.	*RCA* >= 0,5 and employment growth < 0 and national > 0
Bets	Sectors whose participation is very low, but whose exports are growing in the market.	0 < *RCA* < 0,5 and municipal and national employment growth > 0

Note: Based on the National Export Plan 2015-2018. Activities that did not fit into subgroups were discarded.

Source: Authors own elaboration.

In this sense, short-term strategies should focus on strengthening activities classified in the *Recovery* and *Maintenance* categories, since it encompasses sectors that were once competitive in the municipality, but are now declining, and sectors that are currently competitive in the region. The *Consolidation* strategy, in its turn, should be associated with medium-term development strategies, since it focuses on sectors that the municipality already has a certain level of competitiveness, but does not yet have *RCA*. Finally, the *Betting* strategy would be associated with long-term strategies, as it considers sectors in which the region's competitiveness is still low.

5. The case of Belo Horizonte

To illustrate how the Economic Diversification Score (*EDS*) can be used to help devising regional development policies, the *EDS* calculated using some selected variables and weights estimated using PCA was applied to the case of the city of Belo Horizonte. The city, located in the state of Minas Gerais, in the center of Brazil, is among the top 10 cities of the country both in terms of GDP per capita and economic complexity. Moreover, its microregion covers a large industrial area were one of the main factories of the car manufacturer Fiat is located.

5.1. Identifying promising activities for Belo Horizonte

Figure 2 shows the *Activity Space* of the microregion of Belo Horizonte in 2018. The dots colored mark the activities in which the region is competitive, i.e. with *RCA*>1. In this figure it is possible to identify four clusters: (1) public services; (2) modern services; (3) trade; (4) manufacturing; (5) construction.

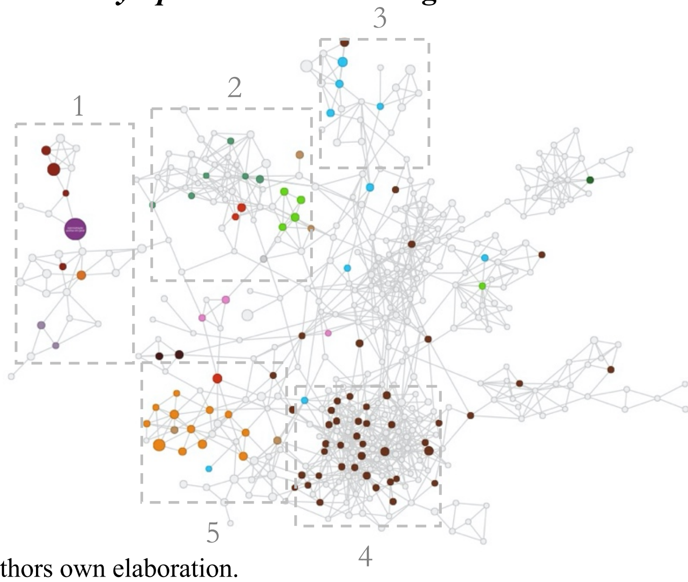
Although we aim to identity promising activities for the municipality of Belo Horizonte, it is important to consider the activities of its microregion in order to take into account the sectors that are competitive in nearby municipalities.

Using the *EDS* presented in section 4.1 it was possible to identify the top 10 most promising activities within each the diversification strategy presented in section 4.2. Figure 3 shows the position of these 40 promising activities in the *Activity Space*. The figure shows that the Bets (in yellow) are in general associated with industrial activities, which tend to present higher complexity, while the remaining activities identified are more spread around the network.

To identify promising macro-sectors, it is also possible to group the 10 most promising activities within each diversification strategy both for Belo Horizonte and its microregion into more aggregated CNAE sectors, at 2 digits.

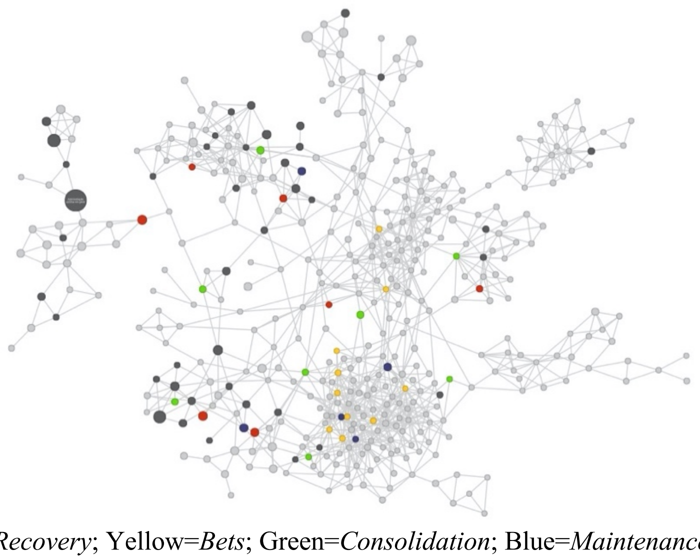
Among the promising activities, the macro-sectors *Machines*, *Electronics* and *Vehicles* encompasses 8 activities. Among these activities, 6 are from the *Bets* strategy. Next comes the *Chemical-Pharmaceutical* (3 activities) sector, and *Metallurgy* (2 activities). Moreover, several specialized services associated with production are identified as promising, such as *Infrastructure*, *IT* and *Machine Maintenance* (6 activities). The results are very similar both for the municipality and the microregion.

Figure 2: *Activity Space* of the microregion of Belo Horizonte in 2018



Source: Authors own elaboration.

Figure 3: Promising activities for the municipality of Belo Horizonte



Notes: Red=*Recovery*; Yellow=*Bets*; Green=*Consolidation*; Blue=*Maintenance*; Dark-Grey=VCR>1.
Source: Authors own elaboration.

The message that emerges from this analysis of promising sectors classified into macro-sectors, therefore, is that the *Machinery*, *Chemicals* and *Metal-Mechanical* sectors are the ones that concentrate the greatest number of promising activities. In addition, the importance of developing the sector of *Modern Services* linked to production is also highlighted.

5.2. Projecting the diversification gains of Belo Horizonte

After identifying the promising activities for Belo Horizonte, it is also possible to calculate how much would the city's economic complexity increase if it achieved *RCA* in some of the targeted activities. Then, using the estimates of the impact of *ECI* on the growth rates of GDP per capita and employment presented in section 3 it is possible to estimate by how much would these variables increase in response to the increase in the city's *ECI*.

Table 5 presents the potential gains (average marginal effects) of different diversification strategies. Only the *Consolidation* and *Bets* strategies were considered, as well as the combination of both. The *Maintenance* strategy is presented to show the inverse relationship, measuring the negative effects of losing *RCA* in these activities.

The estimated effects of *RCA* loss in *Maintenance* activities indicates that losing competitiveness in these 7 sectors could lead to a 3.41% decrease in the municipal *ECI*, which would result in a reduction of 0.7 percentage points in the municipality's growth rate and 15 percentage points in the growth rate of the formal employment ratio.

At the other end of the analysis, the acquisition of *RCA* in the 10 products of the *Bets* strategy would bring a 7.9% gain in *ECI* and an increase in the municipal GDP per capita growth rate of around 1.6 percentage points. This effect is even greater in the formal labor market, increasing the formal employment growth rate by almost 35 percentage points.

The results presented in this section demonstrate the importance of devising short-term and long-term development strategies. More than that, the results show that the opportunity cost for the municipality of Belo Horizonte is considerably higher than that of the microregion in all the strategies presented, which should be carefully looked at by local public managers.

Table 5: Projected gains in GDP per capita and employment by acquiring *RCA* in the indicated promising activities: municipality and microregion of Belo Horizonte

	Municipality				Microregion			
	*Maintenance**	*Consolidation*	*Bets*	Promising sectors	*Maintenance**	*Consolidation*	*Bets*	Promising sectors
Total products in the category	7	5	10	15	10	4	10	14
Projected ECI	2.741	2.901	3.062	3.113	2.836	2.990	3.078	3.176
ECI gain (%)	-3.41%	2.23%	7.90%	9.70%	-1.59%	3.75%	6.81%	10.22%
Rate of change of GDP per capita (%)	-0.70%	0.46%	1.63%	2.00%	-0.21%	0.50%	0.90%	1.35%
Rate of change of employment (%)	-15.03%	9.84%	34.78%	42.71%	-2.55%	6.03%	10.94%	16.41%
Counterfactual level of employment (gain)**	-0.22%	0.14%	0.50%	0.61%	-0.17%	0.40%	0.73%	1.09%
Counterfactual GDP per capita (gain)***	-0.57%	0.37%	1.32%	1.63%	-0.31%	0.73%	1.32%	1.98%

Notes: The *ECI* for Belo Horizonte in 2019 was 2.84 and for the microregion 2.88. * The Mantenance strategy shows the loss if the municipality loses RCV in the sectors. ** Average marginal effect of *ECI* on GDP growth estimated at 0.073 and on employment growth at 0.557. 1 Counterfactual GDP per capita was calculated from the average marginal effect of *ECI* on GDP per capita estimated at 0.06 for the municipality and 0.07 for the region. 2 Counterfactual formal employment level was calculated from the average marginal effect of *ECI* on GDP per capita estimated at 0.02 for the municipality and 0.04 for the region.
Source: Authors own elaboration based on data from RAIS and IBGE.

Finally, it is important to note that the increase in complexity has effects not only on growth, but also on exports, as pointed out by Romero and Britto (2017), on inequality, as pointed out by Hartmann *et al.* (2017) and on greenhouse gas emission, as shown by Romero and Gramkow (2021). The combination of this evidence, therefore, supports the need for smart diversification strategies to overcome the bottlenecks faced by peripheral regions, reinforcing the importance of structural change for economic development. The challenge is to create the right incentives for economic diversification to take place in the desired direction.

6. Concluding remarks

The literature on economic complexity and regional development has been growing rapidly in the last few years. Economic complexity has been applied at the regional level using alternative measures of local knowledge. Balland et al. (2018), for example, use patent data to measure local technological knowledge, using the same methods proposed by Hidalgo et al. (2007) and Hausmann et al. (2014) to guide the formulation of regional smart specialization strategies. Alternatively, Gao et al. (2021) apply the economic complexity methodology using employment data from both China and Brazil, to show that knowledge spillovers are relevant at the regional level, and that improving transport infrastructure helps increasing these spillovers and the productive diversification they foster.

This paper sought to contribute to the literature on economic complexity and regional development using data from Brazilian microregions in three ways. First, it reported econometric tests of the impact of regional economic complexity, calculated using employment data for Brazil, on the growth rates of GDP per capita and of formal employment per capita. These tests, which are inspired in the ones carried out by Hausmann *et al.* (2014), transpose their results to the regional level while expanding them to show that economic complexity is associated with employment as well. Second, it proposed a new method to rank promising activities to be targeted by regional development policies, combining different indicators, as proposed by Hausmann *et al.* (2017), but using weights estimated using a principal component methodology. Third, it showed that the proposed rule for devising smart diversification strategies performs very well when put to test against regions that presented increases in their economic complexity. This methodology is illustrated using the example of the Brazilian city of Belo Horizonte. Using the estimates of the relationship between economic complexity, income and employment, the paper presents simulations of the potential gains to be obtained following the proposed development strategies.

The evidence presented in this paper, therefore, provides important contributions for the formulation of regional development policies. More specifically, it provides an interesting framework for identifying promising activities to be candidates to be targeted by policymakers to promote regional economic development.

References

Almeida, P.; Kogut, B (1999) Localization of knowledge and the mobility of engineers in regional networks, *Management Science*, v.45, n.7 p.905-917.

Balland, P.A.; Jara-Figueroa, C.; Petralia, S.; Steijn, M.; Rigby, D.; Hidalgo, C. (2018) Complex Economic Activities Concentrate in Large Cities. *Papers in Evolutionary Economic Geography*, n.18.

Bathelt, H.; J.S. Boggs. (2003) Towards a Reconceptualization of Regional Development Paths: Is Leipzig's Media Cluster a Continuation of or a Rupture with the Past? *Economic Geography*, v.79, n.3, p.265-293.

Bishop, P.; Gripaios, P. (2009) Spatial Externalities, Relatedness and Sector Employment Growth in Great Britain. *Regional Studies*, v.44, n.4, p.443-454.

Boschma, R.; Eriksson R.; Lindgren U. (2009) How does labour mobility affect the performance of plants? The importance of relatedness and geographical proximity. *Journal of Economic Geography*, v.9, n.2, p.169-190.

Boschma, R.; Frenken, K. (2009) Technological relatedness and regional branching. *Papers in Evolutionary Economic Geography*, n.0907.

Boschma, R..; Iammarino S. (2009) Related variety, trade linkages and regional growth in Italy, *Economic Geography*, v.85, n.3, p.289-311.

Boschma, R.A.; Wenting, R. (2007) The spatial evolution of the British automobile industry, *Industrial and Corporate Change*, v.16, n.2, p.213-238.

Breschi, S.; Lissoni, F. (2003) Mobility and Social Networks: Localized Knowledge Spillovers Revisited. *CESPRI Working Paper*, n.142.

Britto, G.; Romero, J. P.; Freitas, E.; Coelho, C. (2019) The great divide: economic complexity and development paths in Brazil and the Republic of Korea. *Cepal Review*, 127, p. 191-213.

Cassiman, B.; Colombo, M.G.; Garrone, P.; Veugelers, R. (2005) The impact of M&A on the R&D process. An empirical analysis of the role of technological and market relatedness, *Research Policy*, v.34, n.2, p.195-220.

De Groot, H.L.F.; Poot, J.; Smit, M. J. (2009) Agglomeration externalities, innovation and regional growth: theoretical perspectives and meta-analysis. In Capello, R; Nijkamp, J. *Handbook of Regional Growth and Development Theories.* Edward Elgar Pub.

Essletzbichler, J. (2005) Diversity, stability and regional growth in the United States, 1975-2002, *Papers in Evolutionary Economic Geography*, n.0513.

Essletzbichler, J. (2015) Relatedness, regional branching and technological cohesion

Farjoun, M. (1994) Beyond Industry Boundaries: Human Expertise, Diversification and Resource-Related Industry Groups. *Organization Science*. v.5, n.2, p.185-199.

Frenken, K.; Van Oort, F.G.; Verburg, T. (2007) Related variety, unrelated variety and regional economic growth, *Regional Studies*, v. 41, n.5, p.685-697.

Furtado, C. (1964) *Development and Underdevelopment*, Berkeley: University of California Press.

Gao, J.; Jun, B., Pentland, A. S.; Zhou, T.; Hidalgo, C. A. (2021) Spillovers across industries and regions in China's regional economic diversification, *Regional Studies*, p. 1-17.

Gilsing, V.; Nooteboom, B.; Vanhaverbeke, W.; Duysters, G.; Van Den Oord, A. (2007) Network embeddedness and the exploration of novel technologies. Technological distance, betweenness centrality and density, *Research Policy*, v.37, n.10, p.1717-1731.

Glaeser, E.L. (2005) Reinventing Boston: 1630-2003, *Journal of Economic Geography*, v.5, n.2, p.119-153.

Glaeser, E.L.; Kallal, H.D.; Scheinkman, J.A.; Schleifer A. (1992) Growth in Cities, *Journal of Political Economy*, v.100, n.6, p.1126-1152.

Hartmann, D.; Guevara, M. R.; Jara-Figueroa, C.; Aristaran, M.; Hidalgo, C. A. (2017) Linking Economic Complexity, Institutions and Income Inequality, *World Development*, 93, p. 75-93.

Hausmann, R.; Chauvin, J. (2015) Moving to the adjacent possible: discovering paths of export diversification in Rwanda, *Center for International Development (CID) Faculty Working Paper*, n. 294, Harvard University.

Hausmann, R.; Hidalgo C.A.; Bustos, S.; Coscia, M.; Chung, S.; Jimenez, J.; Simões, A.; Yildirim, M. A (2014) *The Atlas of Economics Complexity – Mapping Paths to prosperity*. Puritan Press, p. 364.

Hausmann, R.; Santos, M. A.; Obach, J. (2017) Appraising the Economic Potential of Panama: Policy Recommendations for Sustainable and Inclusive Growth, *Center for International Development (CID) Faculty Working Paper*, n. 334, Harvard University.

Henderson, J.V.; Kuncoro, A.; Turner, M. (1995) Industrial development in cities, *Journal of Political Economy*, v.103, n5, p.1067-1085.

Heuermann, D.F. (2009) Reinventing the skilled region. Human capital externalities and industrial change. *Working paper, University of Trier*, Trier.

Hidalgo, C.; Hausmann, R. (2009) The building blocks of economic complexity, *Proceedings of the National Academy of Sciences*, v.106, n.26, p.10570–10575.

Hidalgo, C.A.; Klinger, B.; Barabási, A.-L.; Hausmann, R. (2007) The Product Space Conditions the Development of Nations, *Science*, v.317, n.5837, p.482-487.

Hirschman, A. O. (1958) *The Strategy of Economic Development*. New Haven: Yale University Press.

Kaldor, N. (1966) *Causes of the Slow Rate of Economic Growth of the United Kingdom*. London: Cambridge University Press.

Lundvall, B. A. (Ed.) (1992). *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning*. London: Pinter.

Markusen, A. (1985) *Profit Cycles, Oligopoly and Regional Development*. Cambridge: MIT Press.

Martin, R.; Sunley, P. (2006) Path dependence and regional economic evolution, *Journal of Economic Geography*, v.6, n.4, p.395-437.

Mealy, P.; Teytelboym, A. (2020) Economic complexity and the green economy, *Research Policy*, p. 1-24.

Menzel, M. (2008) Dynamic Proximities – Changing Relations by Creating and Bridging Distances. *Papers in Evolutionary Economic Geography*, n.0816.

Nelson, R. R. (Ed.). (1993). *National innovation systems: a comparative analysis*. Oxford: Oxford University Press.

Nelson, R.R.; Winter, S. G. (1982) *An Evolutionary Theory of Economic Change*. Cambridge, MA and London: The Belknap Press.

Nooteboom, B. (2000) *Learning and innovation in organizations and economies*. Oxford: Oxford University Press.

Penrose, E. (1959) *The Theory of the Growth of the Firm*. Oxford: Oxford University Press.

Piscitello, L. (2004) Corporate diversification, coherence and economic performance, *Industrial and Corporate Change*, v.13, n.5, p.757-787.

Powell, Ww.; Koput, K.; Smith-Doerr, L. (1996) Interorganizational collaboration and the locus of innovation: networks of learning in biotechnology, *Administrative Science Quarterly*, v.41, n.1, p.116-145.

Prebisch, R. (1962) The economic development of Latin America and its principal problems", *Economic Bulletin for Latin America*, 7(1), p. 1-19.

Rigby, D.L.; Essletzbichler, J. (1997) Evolution, process variety, and regional trajectories of technological change in US manufacturing, *Economic Geography*, v.73, n.3, p.269-284.

Romero, J. P., & Britto, G. (2017). Increasing returns to scale, technological catch-up and research intensity: endogenising the Verdoorn coefficient. *Cambridge Journal of Economics*, 41, 391-412.

Romero, J. P.; Freitas, E. (2018) Setores promissores para o desenvolvimento do Brasil: complexidade e espaço do produto como instrumentos de política. In: Mônica Viegas, Eduardo Albuquerque. (Org.). *Alternativas para uma crise de múltiplas dimensões*. 1ed.Belo Horizonte: Cedeplar-UFMG, p. 358-374.

Romero, J. P.; Gramkow, C. (2021) Economic Complexity and Greenhouse Gas Emissions. *World Development*, 139, p. 1-18.

Schumpeter, J. (1934) *The Theory of Economic Development*. Cambridge, MA: Harvard University Press.

Scott, A.J. (1988) *New Industrial Spaces: Flexible Production Organization and Regional Development in North America and Western Europe*. London: Pion.

Simões, R.F.; Freitas, E.F. (2014) Urban Attributes and Regional Differences in Productivity: Evidence from the External Economics of Brazilian Micro-regions from 2000 – 2010, *International Journal of Economics*, v.1, n.2, p.30-44.

Sorenson, O.; Rivkin, J.W.; Fleming, L. (2006) Complexity, networks and knowledge flow. *Research Policy*, v.35, n.7, p.994-1017.

Storper, M.; Walker, R. (1989) *The Capitalist Imperative: Territory, Technology and Industrial Growth*. New York: Basil Blackwell.

Ter Wal, A.L.J. (2009) The spatial dynamics of the inventor network in German biotechnology: Geographical proximity versus triadic closure. *Papers in Evolutionary Economic Geography*, n. 1102, Utrecht: Department of Economic Geography.