Taylor & Francis
Taylor & Francis Group

Regional Studies

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/cres20

Agricultural economic complexity and regional inequalities: a new approach using census data from Brazil

Alexandre de Queiroz Stein, Gustavo Britto & Victor Medeiros

To cite this article: Alexandre de Queiroz Stein, Gustavo Britto & Victor Medeiros (07 Nov 2024): Agricultural economic complexity and regional inequalities: a new approach using census data from Brazil, Regional Studies, DOI: 10.1080/00343404.2024.2411345

To link to this article: https://doi.org/10.1080/00343404.2024.2411345

REGIONAL STUDIES
https://doi.org/10.1080/00343404.2024.2411345

Routledge Taylor & Francis Group
Regional Studies Association

Check for updates

Agricultural economic complexity and regional inequalities: a new approach using census data from Brazil

Alexandre de Queiroz Stein[a], Gustavo Britto[a] and Victor Medeiros[a]

ABSTRACT
The paper presents an original application of economic complexity to the agricultural sector aiming to determine its suitability to measure and depict regional structural inequalities across Brazilian regions. Using census data, we calculate agricultural complexity indexes for products (PCI-Agro) and regions (ECI-Agro) and the agricultural product space. We show that higher PCI-Agro is associated with higher levels of productivity, education, income and human development. Also, we found that ECI-Agro is associated with higher productivity in agriculture. The econometric evidence points to an inverted 'U'-shape curve relationship between ECI-Agro and agricultural productivity, which has significant policy implications.

KEYWORDS
agricultural economic complexity; agricultural product space; structural heterogeneity; regional development

JEL O10, O13, O21, R11, R12
HISTORY Received 3 April 2024; in revised form 20 September 2024

1. INTRODUCTION

The economic literature has long emphasised the critical role played by agriculture's development within the general economic development process. Ever since Arthur Lewis's early contributions in the context of dualism, the imperative of a backward agricultural sector eventually catching up with the modern manufacturing sector has been the focus of economic research (Lewis et al., 1954). However, most of the debate on structural change over the decades, with notable exceptions, has focused on the transformative role played by the relative growth of modern sectors vis-à-vis backward ones (Gollin et al., 2014; Janvry & Sadoulet, 2020; McMillan & Rodrik, 2011).

In the context of development studies, a wide variety of works can be found covering a myriad of relevant socio-economic aspects related to the connection between agricultural development and economic development as a whole (Bezemer & Headey, 2008; Eberhardt & Vollrath, 2018; German et al., 2020). However, although the importance of the modernisation of agricultural production has been acknowledged, a more in-depth analysis of detailed characteristics has been carried out mainly by the specialised literature, and the proper characterisation of the structural heterogeneity in the agricultural sector is still outdated by a dualistic view of agriculture (Ruzzante et al., 2021; Timmer, 2002).

Studies focusing on the manufacturing industry and the role of diversification and structural change have followed a very different path. These studies escaped from dualistic approaches to development by increasing the levels of disaggregation and going deeper into empirical analysis, which reshaped the field. From the overwhelming and all-encompassing importance of the national manufacturing industry in classical development theories, the literature moved on to the relevance of embedded knowledge in regions, be it the form of the technological intensity of the production of goods and services or in the more recent form of more knowledge-intensive and complex products and sectors in a process of smart specialisation (Balland et al., 2018; Hidalgo, 2023; Restrepo et al., 2022).

In the last decade, studies that draw from the economic complexity perspective originally developed by Hidalgo et al. (2007), Hidalgo & Hausmann (2009a) and Hausmann et al. (2014) have branched to several different areas and themes, either directly through the evaluation of the relationship between economic complexity indexes (ECIs) and various socio-economic characteristics or through the still very fertile exploration of the principle of relatedness under the lens of economic geography (Hidalgo, 2021; Hidalgo et al., 2018). The agricultural sector, however, has been poorly included in recent empirical development of economic complexity methods.

CONTACT Alexandre de Queiroz Stein queiroz.stein@gmail.com
[a]Center for Development and Regional Planning (Cedeplar), Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil
Supplemental data for this article can be accessed online at https://doi.org/10.1080/00343404.2024.2411345

It is particularly noteworthy that studies focused on this sector are absent even in special issues dedicated to economic complexity (Balland et al., 2022; Cimoli et al., 2016).

One important reason for the lack of studies about economic complexity and agriculture is the lack of suitable data. The vast majority of the literature uses trade data, which mixes manufacturing and primary goods. However, trade data fail to capture internal market dynamics, which is essential information, given that a significant part of agricultural production flows to internal markets. More recently, with the expansion of the use of firm-level data based on formal employment, services started to be considered together with industrial goods. However, the assessment of sectors characterised by high levels of informality, such as agriculture, is still a major challenge. Besides, it is still an open question whether the same logic that associates capabilities and productive knowledge with economic complexity applies to the agricultural sector. This sector has considerable differences in the nature of its productive processes, that is, the prominent necessity of specific natural conditions, different work relations in rural areas, the existence of production for subsistence and the perishable nature of many goods, among others, that can directly affect the geography of agriculture.

To tackle these substantial gaps in the literature, this paper adapts and applies the methods developed by the literature on economic complexity to the agricultural sector at the regional level, using the most comprehensive dataset (census) for domestic agricultural production. The underlying hypothesis is that different levels of diversification of the local output, together with the relative 'rarity' of individual goods, can give a meaningful approximation to a region's implicit capacity to combine knowledge and productive factors to produce a basket of agricultural goods. The result is an index to measure and characterise agricultural structural heterogeneity upon which an analytical framework can be built to inform the design policies to foster productive diversification and regional development.

The empirical strategy follows the steps of the known literature dedicated to manufacturing goods. This proof of concept is carried out using data on agricultural production from the latest Brazilian Agricultural Census. These data are particularly suited to apply the economic complexity methodology given that they include the production of more than 300 goods from over 5 million rural establishments, which were aggregated in more than 500 regions. Moreover, crucially, census data are not restricted to goods produced for external markets and include all types of production, including informal and subsistence activities. Hence, data coverage goes well beyond that of more common economic complexity studies, albeit being far less frequently updated. Such a regional approach, based on the production of goods that are destined not only for exports but also for local markets, is also a novelty in the economic complexity literature.

In practical terms, we first calculate the agricultural ECIs for regions and agricultural products. The second step is to calculate similarity measures between agricultural products and to build an agricultural product space. Third, we apply a network clustering method and use its results to characterise the productive structure of regions. These three steps are essential to evaluate if the methodology is suitable given the nature of the sector. The final step is to test empirically the association between the agricultural ECI, the clusters in the product space and key sectoral economic indicators, with an emphasis on the sector's productivity. We show that the productive knowledge measured by the agricultural economic complexity indexes – ECI-Agro and PCI-Agro – are associated with critical socio-economic indicators, such as education, technology, human development levels, productivity and the future growth of regions. Besides, the product space allows us to connect these indicators with the specific types of agricultural production in different regions.

2. AGRICULTURE AND ECONOMIC COMPLEXITY

Like the country's entire socio-economic fabric, the Brazilian agricultural sector has been characterised historically by a high degree of heterogeneity. The concept of structural heterogeneity to characterise the country's agricultural production has become common, as very high and very low productivity sectors and activities coexist persistently through time (Pinto, 1970). Vast chasms between high-productivity, exporting-producing activities and low-productivity, internal market and subsistence-producing ones have persisted over time and across the territory. Over time, these inequalities have endured several waves of productive modernisation, which transformed a deeply agricultural and rural country into a mostly services, manufacturing-led, urban one.

The literature regarding the agricultural sector explores different aspects to portray the structural heterogeneity in the country. Gasques et al. (2010), Felema et al. (2013) and Vieira-Filho and Fishlow (2017) explore differences in agriculture productivity along the Brazilian territory. Souza et al. (2018) and Marconato et al. (2012) present an analysis of technology diffusion in Brazilian agriculture. Regarding productive diversification and specialisation, Santos and Vieira-Filho (2012) present an analysis based on the 12 most important kinds of agricultural products, emphasising the productive concentration in the country. Sambuichi et al. (2014, 2016) also make an interesting approach to productive diversity based on establishments and family farming. Besides the economic studies in Brazil on structural heterogeneity in agriculture, a vast and relevant national literature on the field of economic sociology emphasises that the highly heterogeneous variety of livelihood ways in Brazilian rural areas, especially in family farming, is related to the causes and persistence of structural heterogeneity in the country (Ellis, 2000; Niederle, 2018; Ploeg et al., 2012, 2022; Schneider & Niederle, 2010; Vennet et al., 2017).

Despite the important body of work available, there are no consistent empirical and methodological frameworks capable of resuming the different aspects of the productive heterogeneity among regions. Moreover, there is no toolkit to inform the design of policies to foster regional development in agriculture that considers each region's particularities. Hence, finding a more accurate way to measure and represent the structural heterogeneity in agriculture is still a challenge. There is no consensus about which indicators and methods are more suitable to represent the diversity across regions and to illuminate possible development paths. Here, we propose that an economic complexity approach for agriculture can deal with this issue.

The economic complexity perspective emphasises how local capabilities and productive knowledge play a central role in determining which set of goods and services can be produced by a region. It has been an important framework for applied economics studies in development in the last 15 years. Its empirical toolkit, centred around the ECI and the product space, has been extensively used to analyse changes in countries' productive structures. The seminal work of Hidalgo et al. (2007) and Hidalgo & Hausmann (2009a) gave origin to important studies on structural change and growth at the country level (Britto et al., 2019; Felipe et al., 2012; Stojkoski et al., 2016, 2023; Tacchella et al., 2018). Moreover, it has also been rather prolific in showing how the level of economic complexity relates to a multitude of socio-economic indicators, such as income inequality (Fawaz & Rahnama-Moghadamm, 2019; Hartmann et al., 2017; Morais et al., 2021; Zhu et al., 2020), labour market characteristics (Sbardella et al., 2017), innovation and intellectual property rights (Sweet & Eterovic, 2019; Sweet & Maggio, 2015) and environmental issues (Can & Gozgor, 2017; Lapatinas et al., 2021; Neagu & Teodoru, 2019; Romero & Gramkow, 2021).

Recently, this approach has been used to better understand different aspects of urban and regional economies, especially employment, innovation and technological specialisation patterns. Balland and Rigby (2017) and Balland et al. (2018, 2020) bring significant contributions when using patent data to understand the regional perspective of knowledge flows and specialisation patterns under the approach of economic geography and the relatedness principle. Montresor and Quatraro (2020) go deep into this approach to understand the vital issue of specialisation in green technologies. On the other hand, Freitas et al. (2024), Chávez et al. (2017), Fritz and Manduca (2021), Gomez-Lievano and Patterson-Lomba (2021) and Hane-Weijman et al. (2022) use employment data to deal with diversification matters, estimate economic complexity and also discuss growth in regions and cities. Reynolds et al. (2018), on the other hand, use input–output data in a subnational approach to calculate the economic complexity of Australian regions.

The economic complexity applications until now, however, have not been able to deal with the heterogeneity in agriculture. The original product space and complexity index are calculated using exclusive data on global and country exports. To paint a general picture, this approach has been very successful in ranking countries as well as correlating ECIs with a wide variety of socio-economic indicators. However, as the analysis moves from a more macro-perspective to a subnational (regional) level, the product space network and the ECI based on trade data lose utility. The fact that the vast majority of goods and services produced are destined for local markets rather than foreign ones introduces a significant bias in the analysis. Hence, the use of trade data can underestimate the level of economic complexity of countries in which the production is mainly focused on internal consumption, as is the case of Brazilian agriculture.

There are consistent studies using employment data to analyse economic complexity at the regional level. Still, considering the lack of information about products besides the export databases, there are no studies from a regional perspective using product data. This means that it is still a challenge to understand if the product space is a valuable tool to map and represent the productive structure at sectoral and regional levels. This is particularly relevant for studies that go beyond manufacturing and include agriculture and services, where the level of employment informality is much higher. For these sectors, even the use of employment data to calculate ECIs at the firm level still tends to render biased results. In the work that follows, we test the adaptation of economic complexity instruments to regional data on agricultural production to provide effective indicators to understand structural heterogeneity in the agricultural sector.

3. DATA: BRAZILIAN AGRICULTURAL CENSUS

In order to analyse the complexity of the Brazilian agricultural sector, we used the Brazilian Agricultural Census of 2017 from the Brazilian Institute of Geography and Statistics (IBGE). The census contains information on more than 5 million rural establishments across the territory, which embraces 'any production unit dedicated, either wholly or partially, to agricultural, forestry, or aquaculture activities, regardless of its size' (IBGE, 2019, p. 9). The census is the most complete source of information on Brazilian agriculture, and other fields of economics and agrarian sciences mostly use its data (e.g., Cardille & Foley, 2003; Rada et al., 2019; Ferreira & Féres, 2020; Depaula, 2020; Neves et al. 2021; Oliveira et al. 2022; Junior et al., 2022).

The Agricultural Census has rural establishments as its informational unit. The census questionnaire captures detailed information on the characteristics of each establishment, such as its area, the number of people employed, the production value, the destination of the production, the number and type of machines used in the establishment (such as tractors, harvesters, seeders, etc.), the techniques used for soil preparation, the existence of electric power, if the producer has legal possession of the land, if the establishment is formally registered, the use of fertilizers and pesticides, the existence of organic agriculture, etc. Besides that, the census allows the discrimination of some of these variables by the kind of products, among which is the production value by type of product. This information is available for more than 300 products and is key to infer productive knowledge in regions using the economic complexity approach.

Complementary, the census also gathers information about the establishment managers, the people who work in the establishment and those who live in the establishment. It captures information such as gender, educational level, race, age, marital status and formality of employment relations, among others. All these data are publicly available for 10 different aggregation levels ranging from municipalities (5570 units), which is the most granular aggregation level, through nine other intermediate regional aggregation levels, including the five large administrative regions of the country. Besides that, aggregated data for the country itself are also available. As a census, the data are of good quality for the vast majority of the municipalities. In sum, by providing detailed information on establishments and the individuals living and working within them for a broad range of regions, the census data open the possibilities for a myriad of studies about productive and social development in the rural side of Brazil.

The census is the only database in Brazil that includes reliable data on agricultural production for all municipalities in Brazil with considerable product disaggregation. Other databases, such as the Municipal Agricultural Survey (*Pesquisa Agrícola Municipal*) and the Municipal Livestock Survey (*Pesquisa Pecuária Municipal*) do not map the entirety of Brazilian regions, have a reduced number of products and present a high number of missing observations.

In this paper, we used data on production value for 312 different products aggregated in 510 immediate regions. Each immediate region is an aggregation of municipalities, and the Brazilian urban network gives the definition of which municipalities compose one immediate region. The network is built based on the identification of hub cities and the municipalities that are polarised by them regarding economic and population flows. The information about the flows and the network of cities is provided by the own IBGE's survey called 'Regions of Influence of Cities' (*Região de Influência das Cidades – REGIC*).[1] We did not use the municipality level – which is the most disaggregated level – to avoid informational loss given by the non-identification criteria of the survey, which removes information of any kind that had fewer than three respondents on the chosen regional level (IBGE, 2019). Using the second disaggregation level allowed the capture of a high data variability with low informational loss. To make it clear how spatially granular the data are, Figure A1 in the supplemental data online shows an example of the municipality level of disaggregation, as well as that used in this study (immediate regions) where the spatial distribution of the gross production value (GPV) of agriculture is presented.

The Agricultural Census effectively solves a major issue of data availability for agriculture, given that all other usual databases used in economic complexity approaches have considerable problems regarding the sector. Export data are unable to capture the dynamics of the domestic market, which accounts for the most significant share of Brazilian agricultural production. The census solves this problem by collecting data on all establishments, regardless of whether or not they are formal, if they are family farming producing for local markets or exported-oriented, if they are illegal land occupations or occupied by social movements or even indigenous lands.

Another limitation of highly disaggregated studies in agriculture is the absence of standardisation in the classification of agricultural products or activities. The census allows us to gather information on the production value of 312 products of all kinds across the country, including those produced for self-consumption or local markets. It will enable us to include the least developed regions in the analysis, which usually do not figure in other datasets. It provides a more accurate view of regional disparities in the analyses.

The main limitation of the census is twofold. The first is its declaratory nature. The main datasets in complexity approaches are usually administrative records, which are ideal. However, the high informality in the sector, its diversity of markets and the vast expanse of Brazilian territory make it impracticable to establish regular administrative records of agricultural activities at the desegregation level offered by the census. Second, the IBGE does not make the census micro-data available for public use, which implies significant work in collecting and standardising the information.

4. AGRICULTURAL ECONOMIC COMPLEXITY INDEXES: METHODOLOGY AND FINDINGS

4.1. Methodology for ECI-Agro and PCI-Agro

In order to estimate the agricultural ECIs, we follow the canonical approach of Hidalgo and Hausmann (2009a) and Hausmann et al. (2014), but with regions, denoted by r, in the place of countries and gross production value (GPV) of agricultural products, denoted by p, instead of exported value. The definition of which regions are specialised in each type of agricultural product is based on the revealed comparative advantage indicator (RCA) of Balassa (1965). Formally it is defined in equation (1):

$$RCA_{rp} = \frac{X_{rp}}{\sum_p X_{rp}} / \frac{\sum_r X_{rp}}{\sum_r \sum_p X_{rp}} \tag{1}$$

where X_{rp} is the production value of product p in region r.

From the RCA indicator of a product in a region, we build a binary M_{rp} matrix with r rows and p columns, where $M_{rp} = 1$ if $RCA_{rp} \geq 1$, and 0 otherwise. From M_{rp}, the measures of diversification ($K_{r,0}$) and ubiquity ($K_{p,0}$) were obtained by the sum of the rows and columns of the M_{rp} matrix, respectively, as presented in equations (2) and (3):

$$K_{r,0} = \sum_{p} M_{rp} \tag{2}$$

$$K_{p,0} = \sum_{r} M_{rp} \tag{3}$$

The iteration between ubiquity and diversity is the methodological key to assessing productive knowledge. If alone, diversification provides little information on the quality of a region's capabilities, and ubiquity does not say much about the amount of capabilities a product contains. A region may be highly diversified in very simple products, making several products from a restricted knowledge base. In contrast, a product can have low ubiquity because of a particular natural condition that is not associated with knowledge sophistication.

In order to solve it, the method of reflexes was applied, as proposed by Hidalgo and Hausmann (2009a), deriving the agricultural economic complexity index (ECI-Agro) and the agricultural product complexity index (PCI-Agro), which are based on an iterative method of mutual control between ubiquity and complexity. Formally, the generalised equations of the N iteration is given by equations (4) and (5):

$$k_{r,N} = \frac{1}{k_{r,0}} \sum_{p} M_{rp} \cdot k_{p,N-1} \tag{4}$$

$$k_{p,N} = \frac{1}{k_{p,0}} \sum_{r} M_{rp} \cdot k_{r,N-1} \tag{5}$$

Solving these equations for a chosen N will result in two symmetric matrices of products-by-products and regions-by-regions which condense information about the data, similarly to multivariate analysis methods such as PCA or factorial analysis. The complexity indexes are obtained by normalising the second autovectors of the matrices by their standard deviations from the average.[2]

4.2. Results: heterogeneity across regions

The economic complexity approach is grounded in the understanding that economies are complex systems of information organisation, and the goods produced within a region are construed as the outcome of this process of informational organisation (Hidalgo, 2015). A product results from the combination of individuals' knowledge (explicit knowledge) and know-how (tacit knowledge) within the region, whose interactions occur through various means, particularly within markets. Therefore, products inherently encapsulate fragments of distinct knowledge scattered throughout society, the so-called capabilities. Economic complexity methods posit that the assessment of an economy's portfolio of products has the potential to illuminate the accumulated productive knowledge within that economy and that this knowledge is closely related to its economic development. The intricate question behind our study is assessing if the economic complexity approach to knowledge and its relation with economic development holds when analysing the agricultural sector. In other words, if the index captures the heterogeneity of Brazilian agriculture.

The results show that this is the case. The ECI-Agro depicts the very high levels of productive heterogeneity across the country. Figure 1 shows a gradient of complexity levels in the map for each region. It is clear that even though a north–south pattern arises, the dualistic depiction of the structural heterogeneity of the sector as low- and high-productivity goods is limited. There is a significantly wider range of levels when economic complexity is taken into consideration. This is true for the country as a whole and, in particular, for regions within different states, even the higher complexity ones in the southern part of the country. The stark intra-regional heterogeneity can also be seen in the density distribution of ECI-Agro, by the crossover between the tail end of the higher complexity production of the North and that of the lower complexity production of the South.

Structural heterogeneity is defined by high productivity differentials in the same economy, and the agricultural ECI captures this essential information. The ECI-Agro is positively associated with the GPV per worker. Similar results were found when testing for other productivity measures such as GPV per establishment, GPV per agricultural usable area, agricultural added value per worker and population. These results are aligned with economic complexity studies that show a relation between ECI and gross domestic product (GDP) per capita (Hidalgo & Hausmann, 2009a; Felipe et al., 2012).

When it comes to agriculture, the first (and naive) approach is to think that climatic and geographical conditions essentially determine production structures. Although these conditions play a fundamental role, a number of other factors are essential, such as funding availability, technical assistance, access to machinery, available production techniques and accessible market channels for products, etc. These characteristics are part of the productive knowledge and know-how embedded in each establishment and region.

An essential step in the proof of concept for ECI-Agro is to ensure it captures the productive knowledge associated with agriculture instead of the natural conditions of regions. The positive relation between ECI-Agro and machinery per worker shows how complexity is related to different capital/labour ratio levels in agriculture. It is an example that ECI-Agro captures more than climatic conditions. On the other hand, the not-perfect relation between those two variables, especially for the North and Northeast regions, shows that ECI-Agro captures more than mechanisation in agriculture, dismissing the usual (green revolution-related) approach that development in agriculture is necessarily a matter of modernisation towards homogeneous mechanised agriculture in

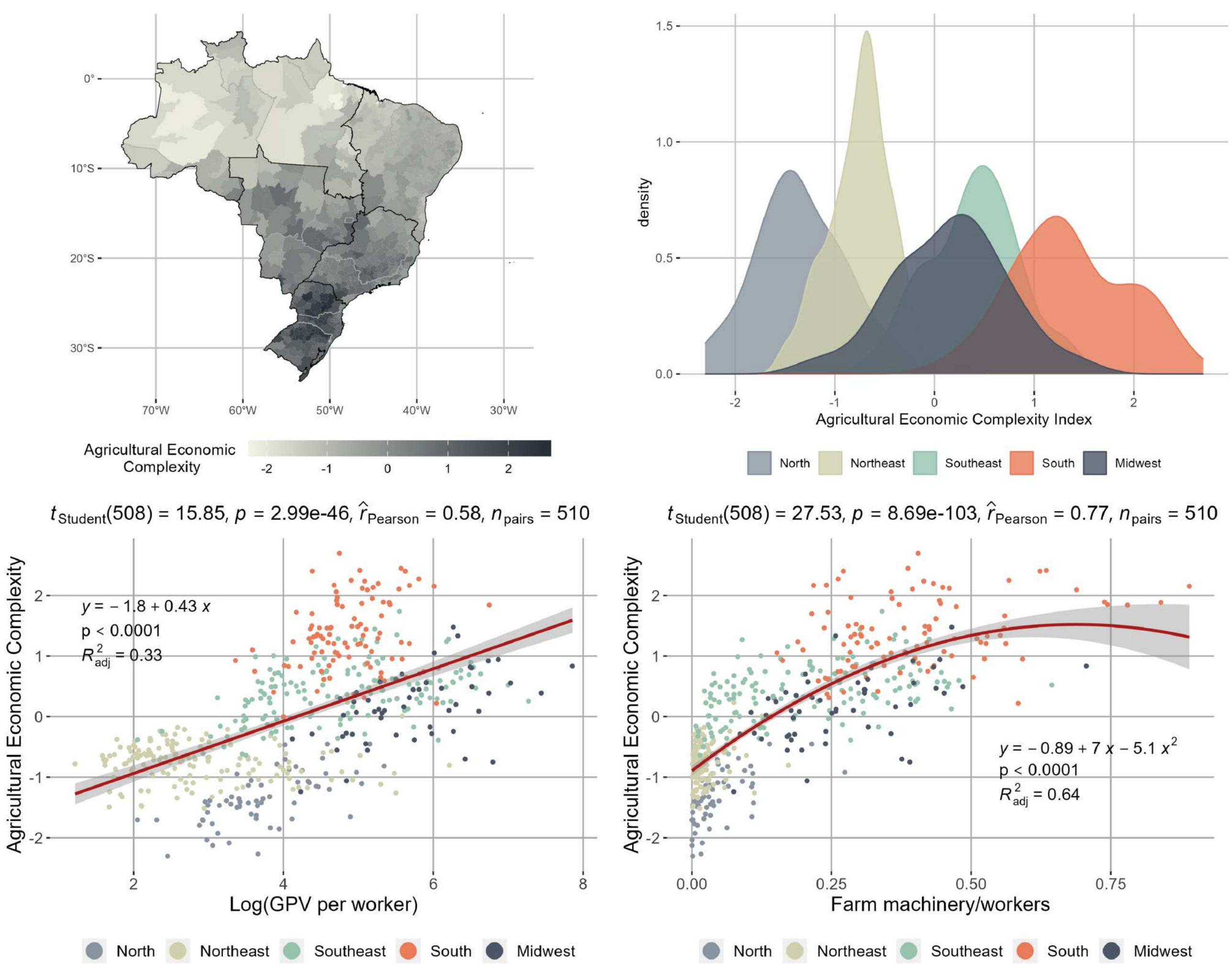

Figure 1. Agricultural economic complexity index (ECI): regional heterogeneity.
Source: Authors' own elaboration based on the Agricultural Census (IBGE).

all regions. In the econometric exercise, we will return to this question.

5. AGRICULTURAL PRODUCT SPACE: METHODOLOGY AND FINDINGS

5.1. Methodology

The product space is a network where each node represents one product, and the connection between two nodes represents the similarity of capabilities embedded in a pair of products (Hausmann et al., 2014; Hidalgo et al., 2007). It is more feasible for an economy to diversify into products or activities similar to its knowledge base. This diversification idea is synthesised as the 'principle of relatedness', which can be defined as 'an empirical principle describing the probability that a region enters (or exits) an economic activity as a function of the number of related activities present in that location' (Hidalgo et al., 2018, p. 452).

The principle of relatedness can be seen in multiple adaptations of the product space made in the last years. Some examples are the technologies spaces from patent data (Boschma et al., 2015), research spaces in physics from publication data (Chinazzi et al., 2019), occupational and skills networks from employment data (Alabdulkareem et al., 2018; Muneepeerakul et al., 2013). Despite sector-specific adaptations of the product space that have been made before, such as the sector-level manufacturing industries space of Neffke et al. (2011), our proposition of structuring a product-level network for agriculture is a novelty in the literature.

The agricultural product space was built following the methodology proposed by Hidalgo and Hausmann (2009a). We estimated the proximity between a pair of products using the usual co-location criteria, in which the higher the number of regions that produce those two products, the higher the similarity between them. Formally, in order to estimate the proximities between products, we multiply the transposed $(M_{rp})^T$ matrix by its original, obtaining a product by product square matrix M_{pp} This is expressed as:

$$M_{pp} = (M_{rp})^T . M_{rp} \tag{6}$$

The proximity measure (ρ_{pp}) between products p and p' is given by the number of regions producing both products (M_{pp}) divided by the biggest ubiquity between the goods

($max(M_{pp}; M_{pp})$):

$$\rho_{pp} = \frac{M_{pp}}{max(M_{pp}; M_{pp})} \tag{7}$$

This process results in a weighted symmetric adjacency matrix which is the base of the product space network. However, simulations show that randomly assigning the RCA values in the original (M_{rp}) binary matrix hardly will render values > 0.3.[3] Removing the edges above these thresholds tends to provide a clear representation of the relations between products. Hence, we defined the agricultural product space as a network where two nodes are connected by edges following two criteria: (1) the edge is part of the graph's maximum spanning tree or (2) the similarity between them is > 0.3: $\rho_{pp'} \geq 0.3$.[4] The first condition assures that the network does not include isolated groups of products, and the second reduces the presence of non-significant connections.

Finally, we applied the Leiden network clustering method of Traag et al. (2019) to the agricultural product space in order to identify homogeneous groups of products with similar productive capabilities. As will be evident below, this is an important step that allows us to interpret the results and connect economic complexity in agriculture to structural change and economic development.

5.2. Results: product groups and regional specialisation patterns

The agricultural product space allows the association of products' characteristics with regional specialisation patterns. Similar activities are expected to be closer in the product space. Considering that the proximity measure is calculated using co-location criteria, it is likely that clusters in the product space will reflect production patterns in groups of regions. This characteristic allows us to identify production patterns across regions. It provides a more detailed characterisation of the productive structure than the usual methods in agriculture studies, which typically use only a sample of the most relevant products of the regions.

The Brazilian agricultural product space has 312 nodes representing products and 2336 undirected edges connecting the nodes by similarity criteria with no isolated components (Figure 2).[5] The clustering algorithm allowed the identification of 12 groups based on the similarity of products. Four clusters, which will be the focus of the analysis, account for almost 75% of the products. For clarity, each group was named after their level of product complexity, given by the agricultural product complexity index (PCI-Agro) we estimated. We called them low-, medium-low-, medium-high- and high-complexity groups, while the other clusters were grouped as 'others'. The groups have statistically significant differences in their PCI-Agro levels, except between the medium-low-complexity group and other non-classified groups.

Due to the co-location criteria used to compute the similarity between products – where products are closer if they are produced in the same regions – product clusters are connected to groups of regions by definition and carry information about them. The low-complexity group contains typical products of the northern regions, especially those from the native forest, such as *babaçu*, *açaí*, *pequi*, cocoa, latex (rubber), *bacuri*, etc. Some of them, such as *açaí*, cocoa and latex, are products of high historical importance for the sociopolitical formation of states such as Acre, Amazonas and Pará. It is interesting to note that these products are mostly plant extraction, which is usually very labour intensive. Yet, they require particular conditions for their production given by edaphoclimatic conditions. For this reason, these products are less ubiquitous than those of the other groups. However, like the classic example of diamonds, their low ubiquity is not caused by the requirement of extraordinary productive knowledge. In Figure 3, the left side depicts the distribution of comparative advantage in the low-complexity products group, common in the North, partly covered by the Amazon forest, and the Northeast of the country.

Medium-low-complexity products are the most ubiquitous group. It includes products that are historically important in Brazilian formation, such as mules (Klein, 1990), and staple foods in the country such as bovine cattle, milk, chicken, cassava, beans, fresh corn on the cob, and basic garden products such as lettuce, collard, tomato and bell pepper. All these products have low-capability requirements and can be found across the Brazilian territory. Figure 4 shows that its relative specialisation is more pronounced in the poorer regions, such as the Northeast and North. This indicates that these regions have not transitioned towards more sophisticated products.

The medium-high- and high-complexity groups of products (Figures 5 and 6) have particular spatial distributions. The first is located mainly near the largest Brazilian urban agglomerations. The logic of the spatial distribution is related to the higher perishability of the goods within this group. It includes not-so-usual garden products such as broccoli, cauliflower, zucchini, cucumber, radish, endive, arugula, etc. Those products do not demand high technical capabilities to be produced on a low scale, but in order for a region to gain comparative advantages in its production, a much higher scale, well-established market channels and proper levels of demand are necessary. Besides, many of these products demand investment in controlled environments (greenhouses) and fertilizers. Hence, even though there is an apparent degree of similarity between medium-low and medium-high products, the latter includes products characterised by the need for larger scale and market-integrated production, while the first comprises a lower scale and often subsistence production. The specialisation pattern in medium-high PCI-Agro products closely follows the population distribution across the country, with the exception of the Brazilian Midwest.

Finally, the high-complexity group of products includes the highly mechanised type of production, such as soybean, corn, rye and oats, many of which are export-oriented productions. This type of production is

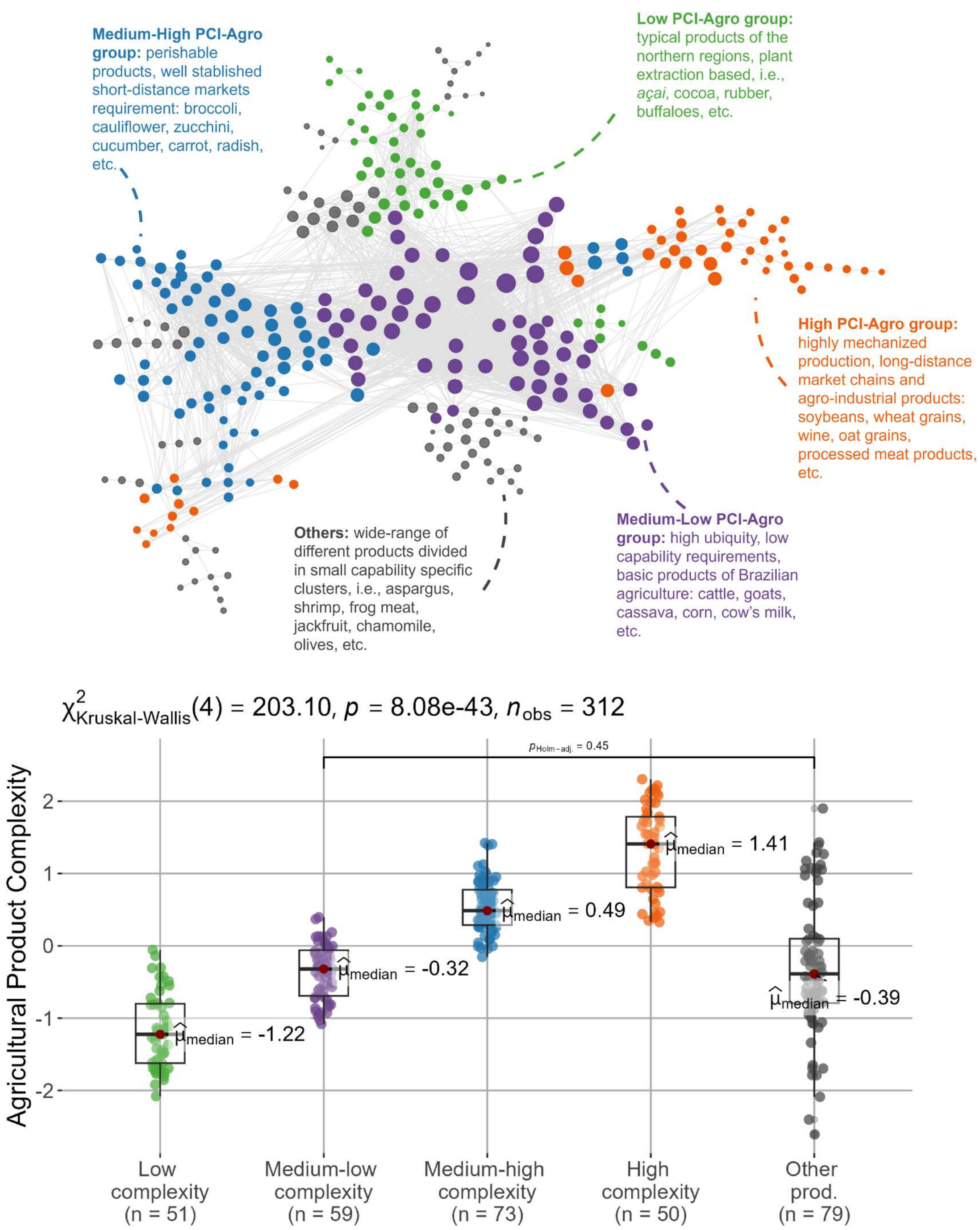

Figure 2. Agricultural product space and product complexity levels.
Note: Complete agricultural product space, nodes = 312, edges = 2336, size = products ubiquity and colours = clusters.
Source: Authors' own elaboration based on the Agricultural Census (IBGE).

clearly more concentrated in the South region, but not limited to it, having an important presence in the Midwest region and in the *Matopiba* region.[6]

The group also includes typical southern products (such as *erva-mate*, grapes and tobacco), which depend on climatic conditions but have highly developed production chains and result in important manufactured goods. The group also includes several agro-industrial products such as breads, sausages and milk creams. These products are produced in rural establishments and commercialised in the local markets, which is compatible with the market-integrated family farming that characterises an essential part of these regions. It is also possible to notice the presence of fruits such as apples, pears, kiwis, plums and peaches, which demand several non-trivial techniques in their production process. In sum, this group includes higher tech agricultural products that are part of longer production and market chains, all of which demand the combination of several different kinds of capabilities and justify their complexity levels.

A careful analysis of the clusters allows us to understand how they represent the Brazilian agricultural

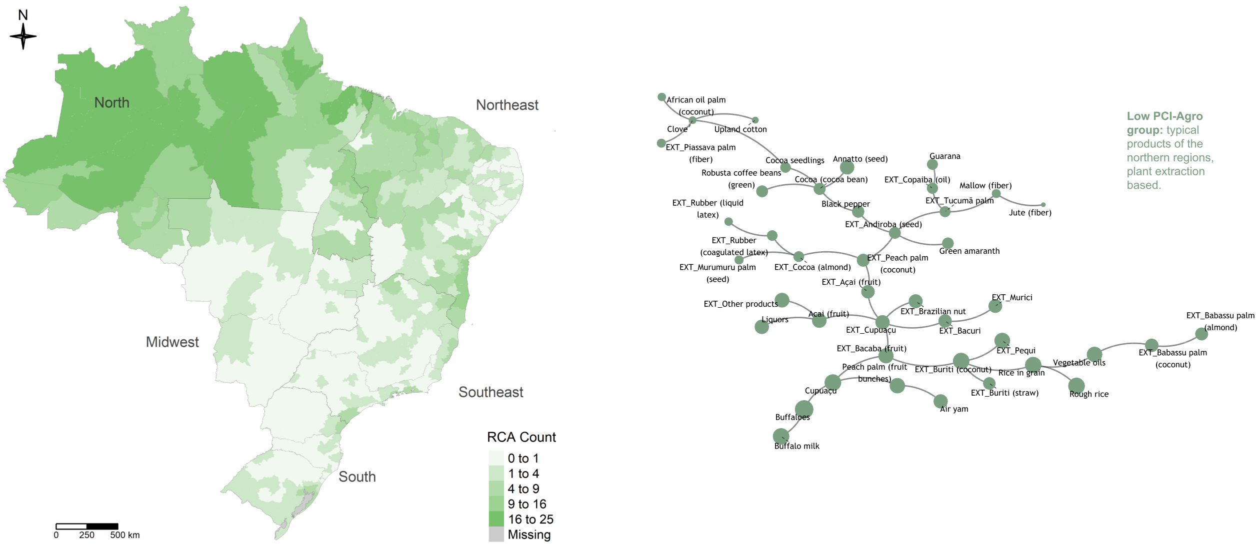

Figure 3. Low-complexity group products spatial distribution.
Note: (left) RCA regional distribution; and (right) low PCI-Agro cluster, size = ubiquity. The prefix 'EXT' in the labels indicates plant extraction products in order to differentiate from the same products obtained by cultivation.
Source: Authors' own elaboration based on the Agricultural Census (IBGE).

productive structure and characterise both low- and high-economic complexity regions. The internal homogeneity of groups of products in terms of complexity alone points to three important characteristics already mapped in the economic complexity literature but not explored for agriculture specifically:

- The agricultural product space reflects regional specialisation patterns in which regions producing a low-complexity product also produce other low-complexity products rather than combine low- and high-complexity products.
- Producing less complex products tends to make it easier to diversify the production into other low-complexity products rather than towards higher complexity products.
- Due to the previous items, a 'hard-to-change' core–periphery productive structure can be identified, where some regions account for complex products and others for less complex products.

The agricultural productive structure allows us to link regions where agriculture is not sophisticated in technical terms – such as North and Northeast – to

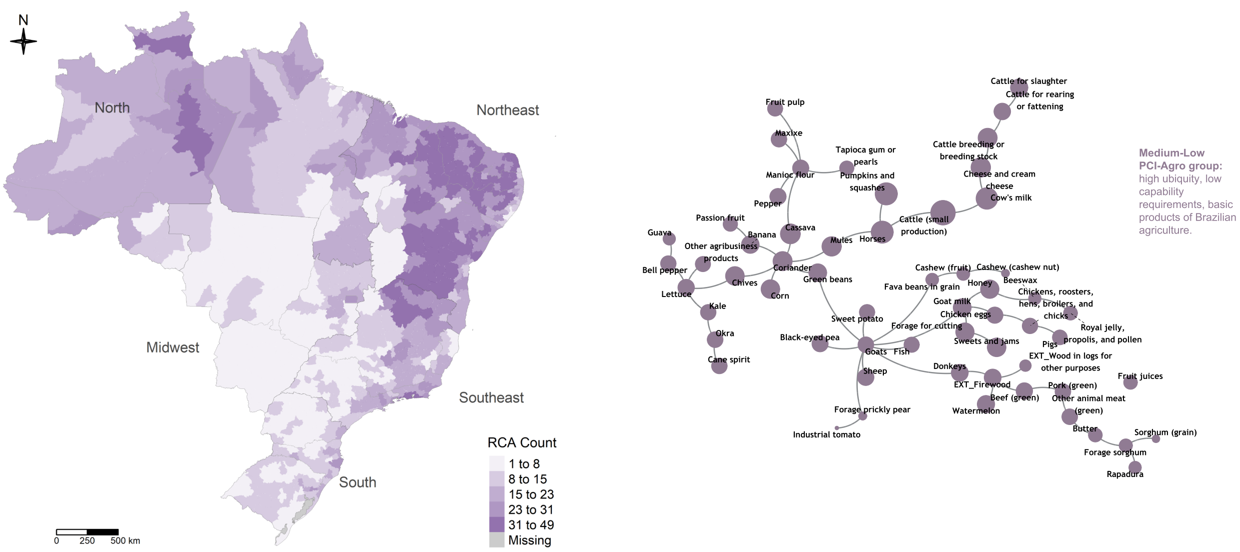

Figure 4. Medium-low-complexity products spatial distribution.
Note: (left) RCA regional distribution; and (right) medium-low PCI-Agro cluster, size = ubiquity.
Source: Authors' own elaboration based on the Agricultural Census (IBGE).

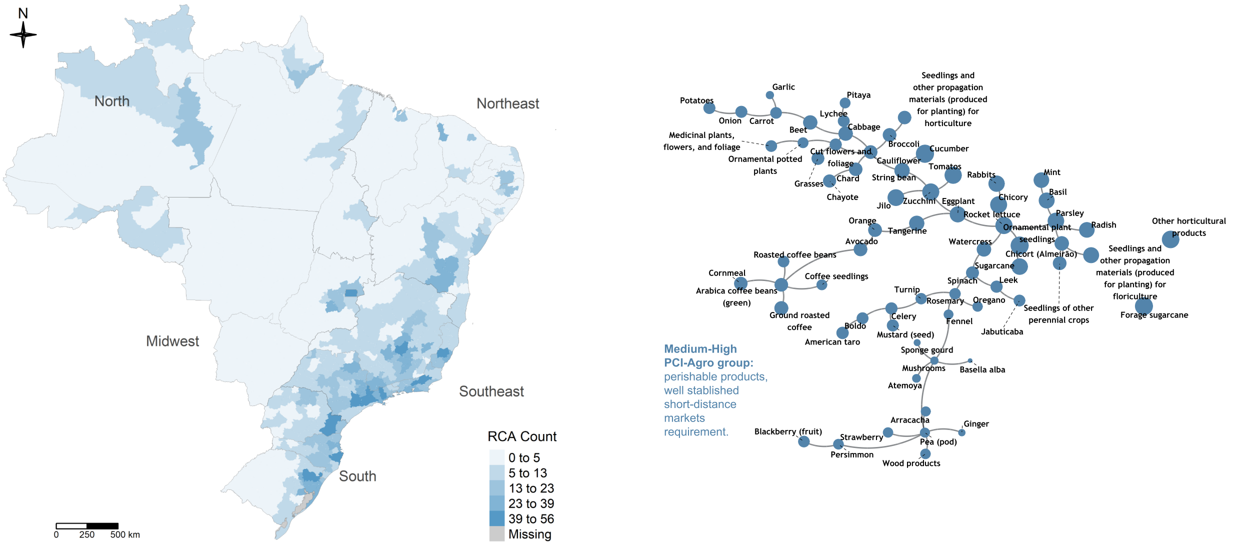

Figure 5. Medium-high-complexity products spatial distribution.
Note: (left) RCA regional distribution; and (right) medium-high PCI-Agro cluster, size = ubiquity.
Source: Authors' own elaboration based on the Agricultural Census (IBGE).

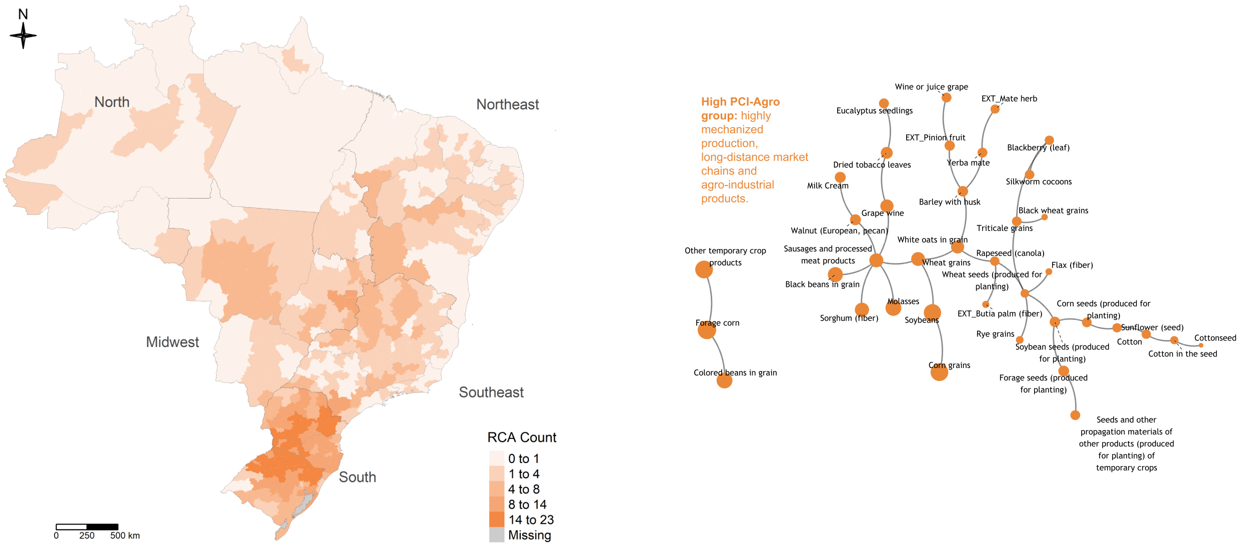

Figure 6. High complexity products spatial distribution.
Note: (left) RCA regional distribution; and (right) high PCI-Agro cluster, size = ubiquity.
Source: Authors' own elaboration based on the Agricultural Census (IBGE).

the kind of products that are important in these places and visualise what kind of products the more sophisticated regions – such as South and Southeast – diversified into to reach higher development levels. In addition, knowing the similarity between products allows us to find more feasible paths for productive diversification towards more complex products and so enhance regional economic development regarding the particularities of each region. It is a significant contribution to drawing policies for the agricultural sector and do not repeat the mistake made by the proposals of the green revolution, which implied the homogenisation of agricultural production aiming to improve productivity levels. By choosing the already existing productive knowledge and conditions of the region as a starting point for structural change, the economic complexity perspective for development brings new tools to help in the development process of regions.

5.3. Connecting the nature of agricultural products with development indicators

The economic complexity approach allows for an intuitive and straightforward depiction of the productive structure of countries and regions. It provides a relatively easy way to differentiate highly developed industrial regions with sophisticated services sectors from underdeveloped regions, usually characterised by agricultural products or products based on natural resources. However, in a

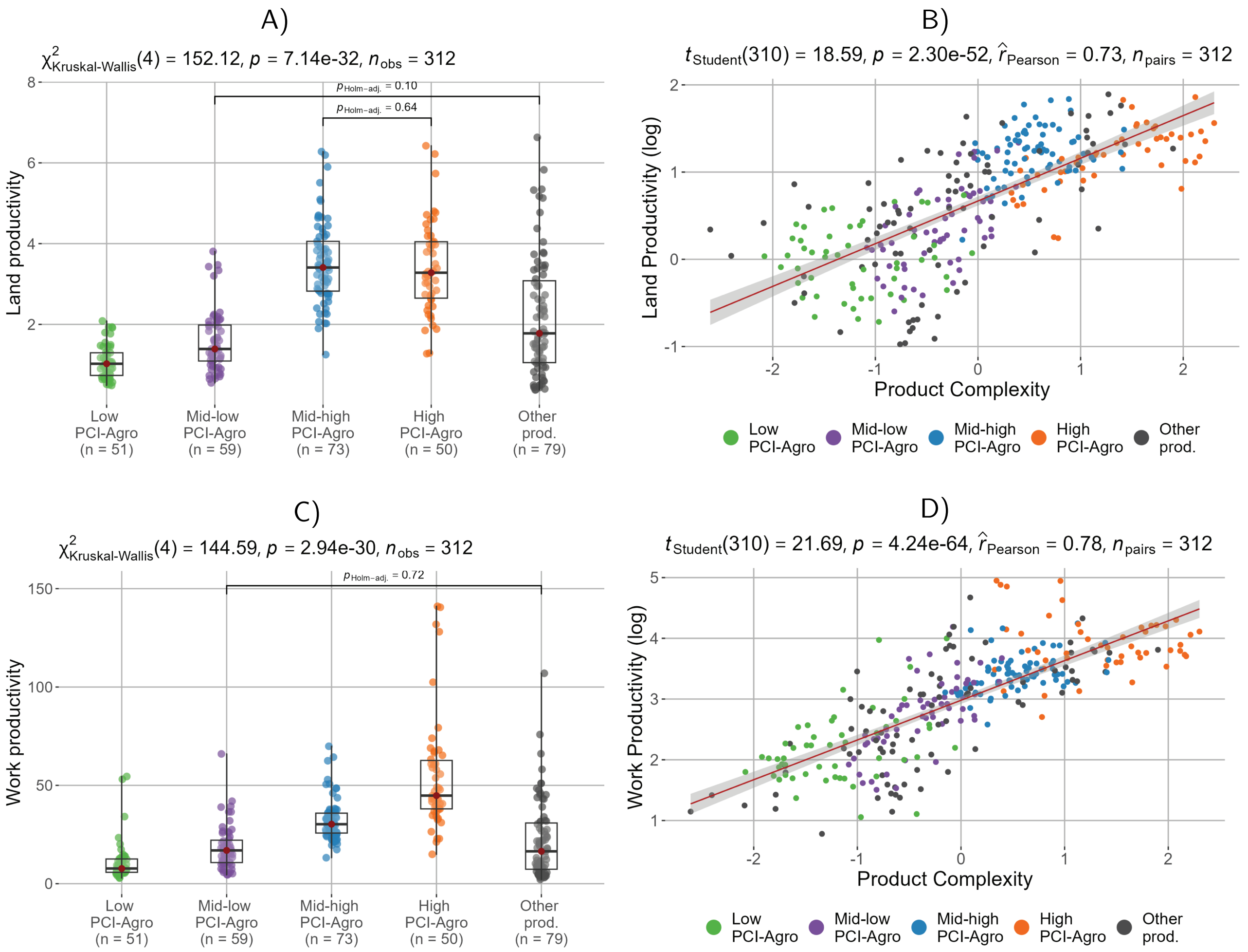

Figure 7. Relation between productivity and product complexity: (A) gross production value (GPV) per usable area (ha); (B) GPV by usable area (ha); (C) GPV per person working in agriculture; and (D) GPV per occupied person working in agriculture.
Note: The box plot shows where group median comparisons do not indicate significant statistical differences.
Source: Authors' own elaboration based on the Agricultural Census (IBGE).

sector-specific case such as agriculture, the interpretation of the results is not as straightforward. For instance, it is harder to know whether a diversified family farming region with well-established institutional markets for small-scale production has a more complex knowledge base and productive structure than a region based on large-scale grain production for external markets. Besides, there is no technological classification available specifically for agricultural products, such as that offered by Lall (2000) for industrial products. Neither there is consensus in the literature regarding what kind of agricultural production leads to higher levels of economic development in the long run.

In order to address this issue and check the relevance of the product complexity index, we opted for the strategy of empirically verifying the relation between products, complexity and socio-economic indicators. To do so, we used the method proposed by Hartmann et al. (2017) when considering income inequality. It consists of computing the weighted average of the social and economic indicators regarding the regions that produce each product p, using W_{rp} (share of product p in region r that produces p) as weights:

$$\bar{Y}_p = \frac{\sum_r W_{rp} Y_r}{\sum_r W_{rp}} \tag{8}$$

where:

$$W_{rp} = M_{rp} \frac{X_{rp}}{\sum_{p'} X_{rp'}} \tag{9}$$

where Y_r stands for the value of each variable of interest Y in region r. It provides a link between products and economic indicators. It allows us to test if there are differences between the relation of products and development indicators for groups of products in the product space, considering their economic complexity levels. As highlighted by Hidalgo (2021), this approach also allows the creation of counterfactual levels for each variable for an economy, given possible changes in its product portfolio and, by doing so, to find interesting diversification paths in order to enhance development.

The combination of the product complexity indexes, product space and the spatial distribution of the product

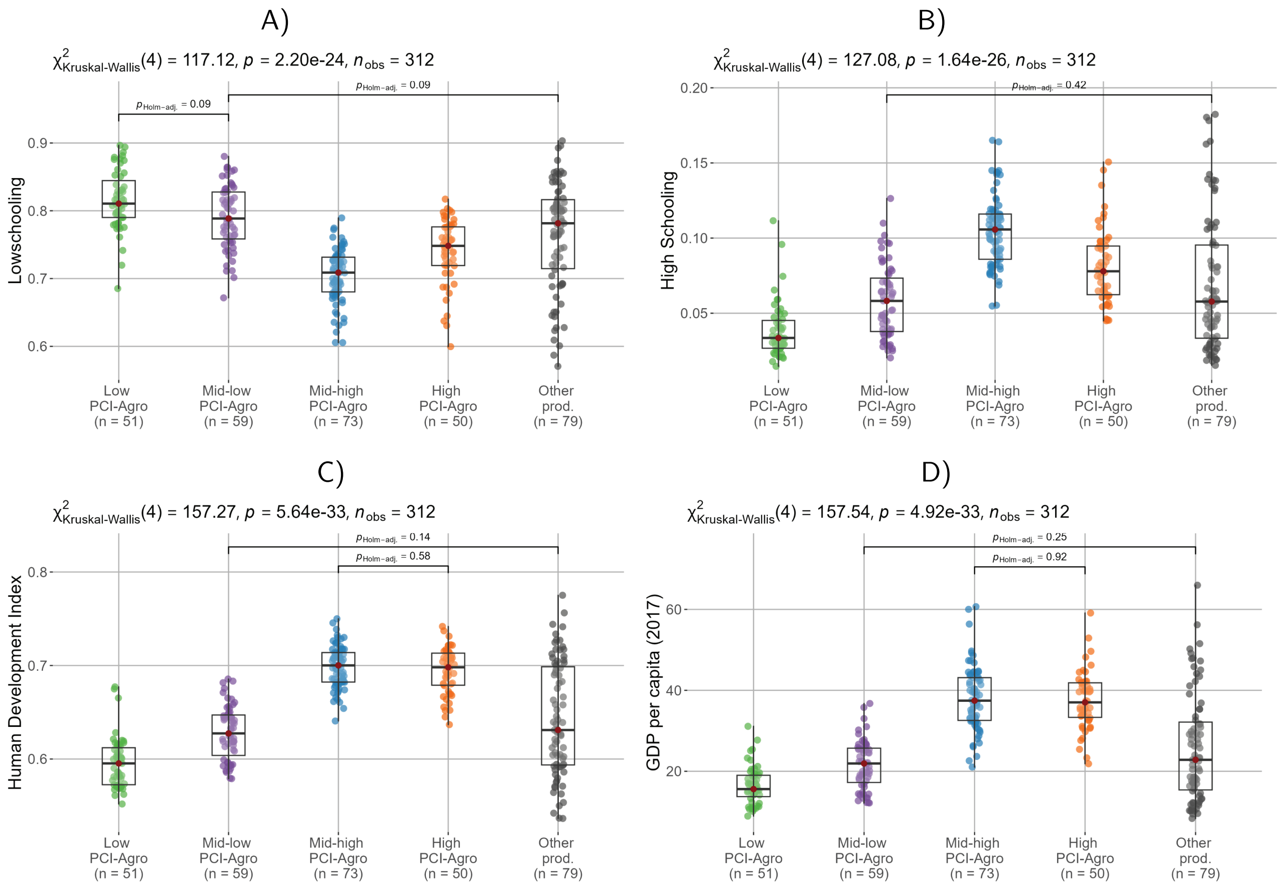

Figure 8. Association between product complexity and socio-economic variables: (A) low educational level; (B) high educational level; (C) human development index; and (D) gross domestic product (GDP) per capita.
Source: Authors' own elaboration based on the Agricultural Census (IBGE).

clusters allows us to characterise the agricultural productive structure. It offers an interesting perspective on productive development in agriculture. The concept of structural heterogeneity, however, includes the assessment of the productivity associated with the productive structure. Our approach to agriculture allows us to connect both perspectives by associating the difference in the region-productivity levels with the kind of products the regions produce. In the plots of this section, we show how the 312 products used in this study are associated with economic development indicators.

The higher the product complexity, the higher the productivity of the regions where they are produced. Two measures of productivity were made by the ratio between agricultural GPV and (1) total workers (occupied people) in agriculture; and (2) usable area for agriculture.[7] Medium-high- and high-complexity groups of products show significant statistical differences from medium-low- and low-complexity groups in terms of productivity per area and work productivity (Figure 7). It suggests that changes in the productive structure through diversification towards higher complexity products are associated with higher productivity levels.

These results show that different kinds of production are associated with different productivity levels. It corroborates the classic structuralist perspective that closing the gap between developed and underdeveloped economies demands significant changes in the productive structure toward sectors capable of diffuse technical progress. In the economic complexity literature, this vision is marked by the motto 'what you export matters' (Hidalgo et al., 2007). Here we agree with that motto and advocate that inside a specific sector such as agriculture, 'what you produce matters' to improve productivity levels.

Besides the economic dimension, we show that more complex agricultural products are also associated with higher levels of education, human development and GDP per capita (Figure 8). Education measures consider the ratio between the number of establishment managers with high/low educational levels and the total amount of managers in the regions. Products in medium-high- and high-complexity groups show significant differences in the educational levels of the property managers compared with low- and medium-low-complexity groups. The same happens when testing for statistical differences in human development indexes of the regions where products are made.[8] The same pattern emerges for GDP per capita. These findings demand future econometric exercises in order to isolate confounding effects and assess the causality relations between the variables, opening a research agenda focused on intersectional agriculture relations.

In highly structurally heterogeneous countries such as Brazil, the traditional proposition that developing industries and promoting innovation is the way to break the underdevelopment trap meets the barrier created by the feasibility of promoting structural changes in less developed regions. Even the economic complexity perspective lacks proper tools to assess feasible diversification paths in regions that have only agriculture production as a knowledge basis. The exercise presented here provides a first step to overcoming these issues by showing ways to plan productive diversification policies for underdeveloped agriculture-based regions.

6. ECI-AGRO AND PRODUCTIVITY: RESULTS FROM ECONOMETRIC ESTIMATIONS

6.1. Basic models and results

The descriptive evidence provided above shows how the economic complexity approach offers an important look at agricultural structural heterogeneity. The following exploratory econometric exercises test if productive knowledge in agriculture is related to productivity and growth. We reinforce this proof-of-concept exercise with an approach similar to that provided by Hausmann et al. (2014). Two sets of regressions were carried out. First, we estimated the relationship between ECI-Agro and agricultural productivity, controlled by other confounding variables. Second, the contribution of ECI-Agro to future growth in the sector was taken into consideration.

For the first exercise, we tested the relationship between ECI-Agro and four measures of productivity: (1) agricultural GPV per agricultural usable area and (2) worker, (3) agricultural added value per capita, and (4) GDP per capita. The first two regressions use data from the Agricultural Census, and the other two use the Municipal GDP Survey data from IBGE.[9] While the first two provide more accurate data, given the detailed nature of the Agricultural Census, the third measure can be used for subsequent years, as we test in the GDP growth rate exercise. Also, as the GPD accounts for the whole economy, it allows us to estimate the importance of ECI-Agro beyond the sector itself. The model has the following specifications:

$$\ln(Y_r) = \beta_0 + \beta_1 * ECIAgro_r + \beta_2 * ECIAgro_r * ECI_r + \beta'_k X'_{kr} + u_k \quad (10)$$

where Y is a vector of productivity measures, X' is a vector of k control variables, u is an idiosyncratic error term, β_0 is a constant term, β_k is a vector of parameters to be estimated for each of the control variables, and $r = 1, \ldots, 510$ represents the regions. We are interested in β_1 and β_2, which capture the ECI-Agro effect and the interaction effect between ECI-Agro and non-agricultural ECI on productivity levels in 2017.

In the second model, the growth equation is specified as follows:

$$\Delta\ln(Y_r) = \beta_0 + \beta_1 * ECIAgro_r + \beta_2 * ECIAgro_r * ECI_r + \beta'_k X'_{kr} + u_k \quad (11)$$

where Δ represents the change in agricultural GDP or GDP per capita between 2017 and 2020.[10] We repeat the same set of control variables as in equation (10), adding the initial level of our dependent variable as control.

As shown in Table 1, higher levels of ECI-Agro are significantly associated with higher productivity in all regressions. The dependent variables in models are in log form and all the complexity indexes are normalised between 0 and 1. In regressions (1) and (3), if not considering the effect of the interaction between ECI-Agro and ECI and using just the linear term, the results show that a 1% increase in ECI-Agro is positively associated with a 1.24–2.26% higher agricultural productivity level.

The interaction term points to an inverted 'U'-shape of the marginal effects of ECI-Agro in productivity; however, only in regression (2) is the interaction term significant. In order to consider the interaction term, we can take the fixed value of 1 for the non-agricultural ECI, which is the maximum it can assume and, therefore, the minimal marginal effects of ECI-Agro. In this case, the marginal effect of ECI-Agro on productivity ranges from −0.13 (regression 2) to 1.29 (regression 3). For any value of non-agricultural ECI < 0.89, all marginal effects are positive. As expected, regression (4) shows that ECI-Agro is positively associated with just a 0.5% change in GDP per capita once this measure refers to all economic sectors in the region and might be more related to complexity in other sectors, such as manufacturing and services.

The results confirm the importance of productive knowledge in agriculture. The regressions were controlled by variables that could confound the effect of productive knowledge. Non-agricultural ECI comes into the models in order to isolate the effect of more developed industrial and services sectors because, as shown in Figure A3 in the supplemental data online, there seems to exist an important unexplored relation between ECI-Agro, the ECI of services and industries and productivity.[11] Human capital controls the role of explicit dimensions of knowledge in agriculture. The measure of agricultural machines per worker captures the capital endowment of regions. Population controls for the size of the economies. The land Gini index and average establishment size capture institutional aspects in rural areas, while taxes and the human development index capture social and institutional aspects of regions as a whole. Finally, we used the latitude and longitude of regions' centroids in order to isolate climatic and geographical effects on agriculture. Besides that, all regressions are controlled for state-fixed effects to isolate specific state policies that can play a role in productivity. Population, taxes, added value and GDP

Table 1. Association between agricultural complexity and productivity, 2017: ordinary least squares (OLS) regressions.

	1. GPV/land area	2. GPV/ worker	3. Agricultural added value per capita	4. GDP per capita	5. ΔAgricultural added value per capita
ECI-Agro	1.2497***	1.2283***	2.2629***	0.5061**	0.8279***
	(0.4495)	(0.3804)	(0.5205)	(0.2298)	0.31
Non-Agricultural ECI	1.0481**	1.2034***	−0.2326	1.3601***	−0.4246**
	(0.4398)	(0.3852)	(0.5592)	(0.2710)	0.20
Taxes (% GDP)	−1.6869**	−1.0584	−5.2485***	1.6588**	0.4751
	(0.7939)	(0.6979)	(1.3840)	(0.8280)	0.37
Population	−0.0208	−0.0162	−0.3365***	−0.0385*	−0.004
	(0.0358)	(0.0301)	(0.0511)	(0.0199)	0.02
Human	3.5566***	2.9283***	−0.9372	4.0983***	−1.2279***
Development	(1.0996)	(0.9064)	(1.5677)	(0.7254)	0.45
Land Gini Id.	1.2600***	0.8375***	0.6227	0.0031	−0.3504***
	(0.3073)	(0.2586)	(0.3912)	(0.1549)	0.13
Human Capital	3.2272***	1.8523***	0.0890	−0.2870	−0.6459**
	(0.6664)	(0.5711)	(0.8889)	(0.3594)	0.30
Machines per	0.3993***	0.3997***	0.4199***	0.0347	0.0868***
Worker	(0.0480)	(0.0426)	(0.0529)	(0.0239)	0.02
Avg. Establishment	−0.6987***	0.2294***	0.0786	0.0818***	0.014
Size	(0.0441)	(0.0393)	(0.0586)	(0.0232)	0.02
Latitude	0.0144	−0.0040	0.0343*	−0.0127	0.0133*
	(0.0144)	(0.0121)	(0.0195)	(0.0083)	0.01
Longitude	−0.0324***	−0.0360***	−0.0828***	−0.0153**	−0.0047
	(0.0125)	(0.0108)	(0.0168)	(0.0066)	0.01
ECI-Agro*Non-	−0.6274	−1.3680**	−0.9678	−0.5547	−0.0485
agricultural ECI	(0.7267)	(0.6259)	(0.8953)	(0.3736)	0.37
Agric. A.V. PC					−0.0064
(2017)					0.05
Agric. A.V. PC					−0.1408*
(2017)*ECI-					0.08
Agro					
Observations	510	510	510	510	510
R^2	0.81	0.91	0.75	0.85	0.527

Note: Robust standard errors are reported in parentheses. *0.1, **0.05, ***0.01. All regressions include state fixed-effects.
Source: Authors.

were obtained from IBGE. Land Gini was estimated using the number of establishments and area from the Agricultural Census. In sum, we can strongly argue that ECI-Agro captures a tacit dimension of productive knowledge in the agricultural sector that is related to productivity.

In the second model, the objective is to test the relationship between ECI-Agro and the sectoral growth of agriculture, given the current level of productivity. Our aim is very similar to the exercise presented in Hausmann et al. (2014), which tests for the relationship between economic complexity indicators and future growth. In *The Atlas of Economic Complexity* (2014), the intention is to test if the ECI is capturing natural endowments, foreign commerce openness or the size of the economy. Our version focuses on showing that ECI-Agro does not reflect natural conditions, capital endowment, institutional aspects, or the regions' industrial and services productive structure.

Regression (5), presented in both Table 1 and Table A2 in the supplemental data online, shows the results for the second econometric model, where our dependent variable is the growth in agricultural added value per capita of regions between 2017 and 2020. The interaction term between added value and ECI-Agro intends to capture the heterogeneous effects that agricultural complexity has in regions with different productivity levels. Across the different specifications tested, a 1% increase in agricultural economic complexity is associated with an increase in agricultural productivity that varies from 0.17% to 0.82% over a three-year period. The interaction term between ECI-Agro and non-agro-ECI was not significant, and its coefficient does not impact the marginal effect of ECI-Agro much. These results suggest a persistent correlation between ECI-Agro and productivity growth.

The results shown in Table 1 provide evidence that ECI-Agro is strongly and positively correlated with productivity. It opens new avenues of investigation for future research, such as the causal impacts of ECI-Agro on productivity growth as well as its relationship with other variables. It could provide important and original results, especially in contexts where a clear policy orientation to foster agricultural complexity can be noticed, in the same vein as recent studies have done for the manufacturing sector (Queiroz et al., 2023).

As agricultural complexity and productivity may be correlated with a similar range of control variables, omitted variable bias likely does exist. We tried to overcome this bias by including various independent variables related to economic, social, institutional and geographical attributes. In addition, the causality between agricultural complexity and productivity seems to be hard to identify. We expected a reverse causality wherein agricultural complexity causes productivity growth, but the opposite direction might also be observed. This endogeneity issue might be further explored by employing instrumental variable approaches in a more extensive panel data setting, whereby region-fixed effects could be combined with instruments providing quasi-random sources of variation to isolate the true effect of agricultural complexity on productivity outcomes.

Unfortunately, no census data are available to build a complete panel that can estimate the long-term effects of ECI-Agro on agricultural growth. Besides, the model presented has very period-specific characteristics, especially considering that the pandemic period is included. However, it will pave the way for similar applications in agriculture in the future when more data become available. It also provides a template for applications using data from other countries, with different datasets, in order to test the methodology presented here. It brings up an important research agenda for the next years on the relevance of productive structure and structural change in agriculture for the development of countries and regions.

6.2. Not a panacea: notes on non-linearity, diversity and productivity

Even if the effects of ECI-Agro on productivity are robust on the aggregate, the high structural heterogeneity among Brazilian regions requires that possible non-linear effects of economic complexity on agricultural productivity be accounted for. Moreover, considering that land is the key factor in agricultural production and the substitution of one type of product for another is the usual form of structural transformation in the sector, the relation between accumulated productive knowledge and diversity is not necessarily linear. These elements may result in heterogeneous effects across regions. Morais et al. (2021), for instance, showed how the effect of economic complexity on inequality in Brazil follows an inverted 'U'-shape. Departing from lower levels of economic complexity, income inequality initially worsens, improving as economic complexity reaches higher levels. This effect is more acute for less developed states.

The results obtained by the introduction of the square term of ECI-Agro help to highlight the specificities of agriculture in relation to other sectors. More specifically, the process of productive diversification can follow different patterns compared with that of manufacturing. In general, the evolution of the productive structure follows a pattern where a new and more complex product is added to the region's portfolio, resulting from the combination of existing productive knowledge, thus expanding its diversification, economic complexity and productivity simultaneously. Over time, some of the products can completely disappear as processes of diversification or specialisation settle in. However, a characteristic of structural change in agriculture is that introducing a new product can immediately substitute the production of another, especially when a limited endowment of land is available. Hence, once a region acquires the capabilities for a more complex product, either by related diversification or by jumps in product space, it may eliminate the production of pre-existing goods. Hence, diversification may decrease local productive diversity.

To understand how this characteristic affects the relationship between economic complexity and productivity in this sector, consider two different scenarios. The first is where a region obtains the capabilities to produce a more complex product and replace other products. The development of agriculture in the Amazon is an example of this trend. There, the forest is cleared, and its diversity, whether in native extraction products or potential goods related to biodiversity and bioeconomy, is usually replaced by soybean monoculture or cattle ranching. In the case of substituting with soybean plantations, this process increases economic complexity by adding a product with a higher level of productive sophistication and high productivity while simultaneously reducing the region's level of diversification.[12]

The second scenario is the opposite. This seems to be the case in several regions in the Northeast, especially in the *Cerrado* biome. Many Northeastern regions have a predominance of mid-low-complexity products and are marked by highly diversified production of small-scale basic products, usually destined for subsistence or local markets. Higher productivity levels can be achieved by reducing their diversification and focusing resources on fewer products that bring gains of scale, even if these products have lower economic complexity than the average of the region. Thus, in this case, a reduction in economic complexity may come together with productivity gains.

In sum, these possibilities imply three important considerations: (1) if regions replace more than one agricultural product to give room to one predominant product, it may not be possible to observe a clear positive relation between economic complexity and diversification as is usual in other sectors such as manufacturing and services;

Table 2. Non-linear ECI-Agro and densities effects on productivity.

	GPV/ha			GPV/occupied person		
Dependent variable	**1**	**2**	**3**	**4**	**5**	**6**
ECI-Agro	1.2497***	−1.6903*	2.9175***	1.2283***	−0.9074	2.5028***
	(0.45)	(0.92)	(0.85)	(0.38)	(0.76)	(0.72)
Non-Agric. ECI	1.0481**	1.5893***	1.36***	1.2034***	1.5966***	1.535***
	(0.44)	(0.44)	(0.4)	(0.39)	(0.39)	(0.35)
ECI-Agro*Non-Agric. ECI	−0.6274	−1.8365**	−1.446**	−1.368**	−2.2463***	−2.2132***
	(0.73)	(0.72)	(0.73)	(0.63)	(0.64)	(0.64)
ECI-Agro^2		3.041***	−1.0843		2.2091***	−0.8695
		(0.74)	(0.75)		(0.63)	(0.65)
Density (Low PCI-Agro)			3.6157***			3.0803***
			(0.97)			(0.82)
Density (Mid-Low PCI-Agro)			−7.8512***			−5.8242***
			(0.68)			(0.6)
Density (Mid-High PCI-Agro)			2.3689***			2.6072***
			(0.86)			(0.7)
Density (High PCI-Agro)			−0.2309			0.2009
			(0.71)			(0.64)
Density (Other Products)			0.2906			−2.1844
			(1.99)			(1.66)
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	510	510	510	510	510	510
R^2	0.811	0.818	0.887	0.913	0.915	0.944

Note: Robust standard errors are reported in parentheses. *0.1, **0.05, ***0.01. The regressions also include the same control variables as the previous models: Population, Human Development Index, Land Gini, Human Capital, Machines per Worker, Avg. Establishment Size and Geographical Coordinates.
Source: Authors.

(2) consequently, in some cases, agricultural complexity more closely related to the type of the products in the region and its level of sophistication than it is about the diversification levels of the region; and (3) the relationship between economic complexity and productivity in agriculture is mediated by the level of diversification and the nature of the products.

The question that arises in the relation between agricultural economic complexity and productivity is if, between two regions that have their production concentrated in fewer goods, the one producing more complex products tends to have higher productivity. In other words, it is necessary to account for the effects of the diversity of the products and their characteristics.

To this end, along with the quadratic term of ECI-Agro, we introduced density variables for the regions in the different clusters of the product space. The density indicator measures how close a region's productive structure is to a specific product x by counting the number of adjacent products to x in the product space in which the region has RCA (Hidalgo, 2021; Hidalgo et al., 2007) and weighting it by its similarities. The same can be applied to capture the density of the region to a group of products. Thus, in practical terms, if a region has a high density of products of the group of low PCI-Agro, for instance, it means that this region is well diversified in products of that group.

As shown in Table 2, when controlled by the density indicators, the linear term of ECI-Agro becomes positive and significant, and the non-linear relation loses significance. Also, the coefficients of the density variables show that, everything else being constant, higher diversification in the groups of low PCI-Agro and mid-high PCI-Agro products is significantly related to higher productivity levels.

While the density in high PCI-Agro products is not significant, likely because its effects are captured by agricultural ECI,[13] the density in mid-low PCI-Agro confirms that, in some regions, reducing the diversification in small-scale products can boost productivity, as discussed above.

As our regression includes a square term of ECI-Agro and an interaction term between ECI-Agro and non-agricultural ECI (which, in this scenario, is significant), it presents marginal effects that vary for different ECI-Agro levels. Figure 9 shows the marginal effects of ECI-Agro on land productivity from model 3. The marginal effect is mostly positive and diminishes to near zero in regions with high complexity. It implies an inverted 'U'-shape of the effect of ECI-Agro on productivity. Moreover, by defining the levels of non-agricultural ECI as the mean of Brazilian administrative regions, it is possible to note that the higher the marginal effects, the lower the non-agro ECI is. For example, considering the ECI average of the South region (0.57) and an ECI-Agro of 0.5, the marginal effect on land productivity is around 1%.

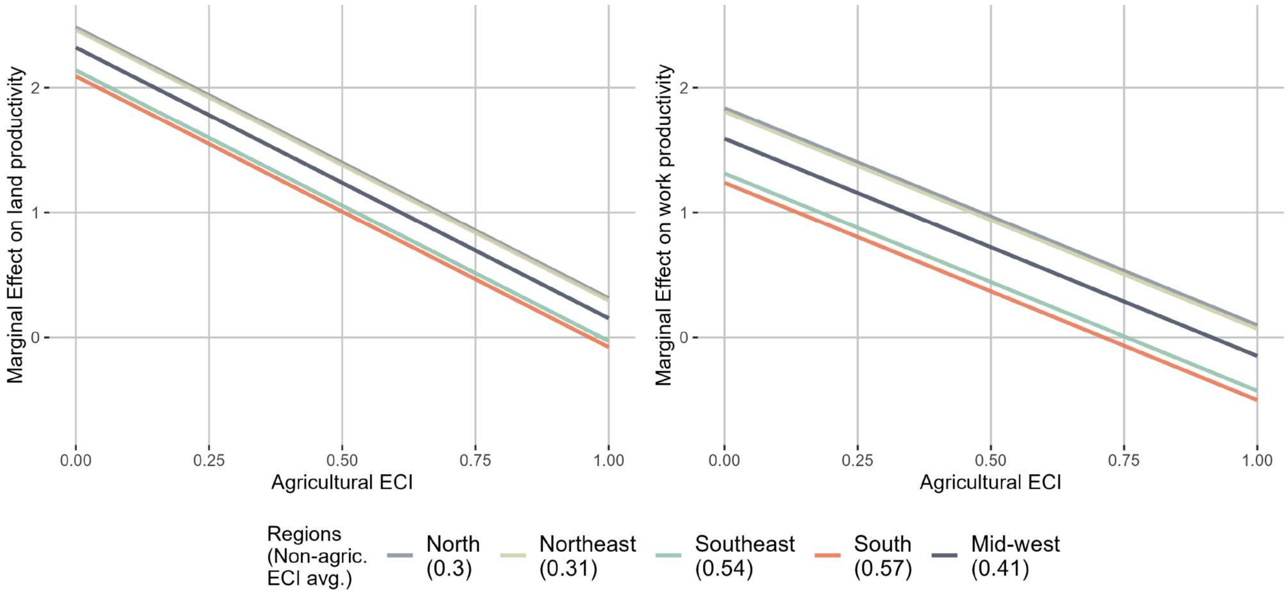

Figure 9. Marginal effects of ECI-Agro on productivity.
Source: Authors' own elaboration based on the Agricultural Census (IBGE).

For work productivity in agriculture, model 6 follows the same pattern. However, the marginal effect becomes negative for higher levels of ECI-Agro.

The results of this session corroborate the linear approach's finding in the previous one that economic complexity is positively related to productivity. However, they also show how the interpretation of economic complexity in agriculture, especially regarding its relations with diversification and productivity, has to be taken cautiously locally. More importantly, they illustrate how it is essential to consider that even though economic complexity is a driver for productivity in general, it does not necessarily imply that an increase in economic complexity is always immediately associated with improvements in local economic and social indicators. Moving from lower to higher complexity production, especially departing from lower levels of development, does not always guarantee, for instance, a more environmentally sustainable production or better working and living conditions for people who live in rural areas.

7. CONCLUSIONS

This paper has contributed to the literature on regional development by demonstrating, for the first time, that the economic complexity perspective can be successfully transposed and applied to disaggregated data on regional agricultural production, substantially broadening the analytical possibilities for this sector.

It showed that ECI-Agro accurately depicts the high level of productive heterogeneity across the country by showing a strong correlation between gross value added and machinery per worker. By constricting the agricultural product space, it was also possible to demonstrate how the production of goods of similar levels cluster together, following the country's north–south inequalities and socio-economic inequalities within and outside the agricultural sector.

The econometric exercises also revealed that ECI-Agro is positively correlated with our distinct indicators of productivity levels in the sector, which are controlled by several other variables. It is also positively correlated with the agricultural added value per capita growth rate. All regressions that included an interaction term showed a significant association between agricultural ECI and productivity growth.

Finally, the regressions also unveiled a non-linear relationship between ECI-Agro and productivity, suggesting an inverted 'U'-shaped curve as economic complexity increases from very low to higher levels. The introduction of density indicators also provided evidence that the nature of diversification within the agricultural sector may determine, locally, a different relationship between diversification and productive diversity, as larger scale goods replace diverse low-productivity goods.

In sum, the exercises showed that ECI-Agro can be an important tool for assessing economic heterogeneity. We found that agricultural product complexity is closely associated with structural and socio-economic variables such as education and the human development index. Hence, beyond being a test for the index, this exercise is useful for identifying critical socio-economic issues closely related to the nature of agricultural production, especially in poorer, more isolated areas.

The results shown in the paper have important policy implications. First, it provides an extremely useful toolkit to characterise agricultural production by using one multi-dimensional index closely correlated with an array of socio-economic variables while depicting a network of goods that tend to cluster. This is particularly valuable for a continental country with acute productive heterogeneity for which large amounts of data are available.

Hence, ECI-Agro can help to paint a broad picture, pointing the research to specific areas and types of production.

Second, identifying the clusters of goods and their regional variability provides a means to help identify diversification possibilities, considering specific geographical and environmental conditions. Given the positive association between the agricultural and non-agricultural ECIs, it also opens the possibility of investigating diversification opportunities outside the agricultural sector.

Third, once the usefulness of ECI-Agro has been shown, the principle of relatedness can be brought to bear at full force to investigate the possibilities of diversification across macro-sectors. This is an important bottleneck to investigating the possibilities of creating local and international value chains, as well as taking into account more sustainable production and occupations.

Finally, by investigating different diversification strategies, either related or unrelated, this perspective allows for identifying possible paths to development for poorer and isolated areas, which are characterised by very low productivity and abysmal social conditions. In this sense, ECI-Agro and the agricultural product space analysis are not only tools that paints a general picture of the country's agricultural production. They also bring a new hall of possibilities for regions that have no services and manufacturing knowledge bases from where to start productive diversification processes. These regions have been consistently left behind.

In other words, there is guidance to identify the who, where and how to diversify the agricultural sector, following the argument set out by Hidalgo (2023) for the industrial policy. However, the paper brings an important word of caution. Moving up the complexity ladder certainly means different paths for different regions, given the available levels of capabilities and other conditions related to agriculture. For instance, higher economic complexity in this sector may not always mean more sustainable production. On the contrary, it may well mean the suppression of the production for local markets as well as the devastation of important biomes by exporting monocultures. This perspective is especially relevant when considering the role that peripheral industry-lacking regions, such as the Amazon, play in environmental issues, where sustainable development is imperative. The methodology presented here provides an empirical framework for development policies in those regions that until now were missing in economic development literature.

The application also has limitations that suggest a fruitful research agenda. First, more work is needed using different datasets to assess the nonlinearities found in the connection between ECI-Agro and productivity. This finding calls for more production and region-specific studies. Secondly, the available Census data used in the exercises is aggregated by regions, and access to the raw microdata, which is not public, could provide better results. Also, more can be done to make Agricultural Census data compatible over the years to provide a longitudinal analysis of agricultural complexity between regions and across time. Finally, full integration between agricultural products and broader economic activity classifications is required in order to obtain a complete set of possible productive diversification policies for regions that go all the way from agriculture to more dynamic sectors.

ACKNOWLEDGEMENTS

We are deeply grateful to the journal editors for their diligent editorial work and swift handling of the manuscript. We also thank the anonymous referees for their insightful and constructive feedback. Any remaining errors are solely our responsibility.

AUTHOR CONTRIBUTIONS

Alexandre de Queiroz Stein: conceptualisation; data curation; formal analysis; methodology; visualisation; software; writing the original draft. Gustavo Britto: conceptualisation; supervision; writing – review and editing. Victor Medeiros: formal analysis; methodology; software; writing – review and editing.

DATA AVAILABILITY

The data used in this paper were generated by the Instituto Brasileiro de Geografia e Estatstica and can be accessed at https://sidra.ibge.gov.br

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

FUNDING

The research undertaken for this paper was financed in part by the Brazilian National Research Council (CNPq) [grant numbers 421124/2022-9 and 311052/2022-3], by the Coordination for the Improvement of Higher Education Personnel (CAPES) [grant numbers 88887.837602/2023-00; 001 and 88887.911150/2023-00; and 88887.828037/2023-00], and by the Minas Gerais State Research Support Foundation (FAPEMIG) [grant number 04746-22].

NOTES

1. For more about Brazilian geographical classifications, see IBGE (2017, 2018).
2. For more on the method of reflections, see Hidalgo & Hausmann (2009a); and for its methodological appendixes, see Hidalgo & Hausmann (2009b).
3. The simulations followed the same procedures proposed by Hidalgo & Hausmann (2009b).
4. This is distinct from Hidalgo et al. (2007), in which the definition of product space considers all connections between products, or, in other words, the edges weight

criteria is just $\rho_{pp'} > 0$. However, the authors use a threshold of $\rho_{pp'} \geq 0.5$ for the network representation.
5. For a more detailed representation of the agricultural product space, see Figure A2 in the supplemental data online. For lists of most and less complex products, see Table A1 online.
6. *Matopiba* is composed by areas of the northern and north-eastern states of Maranhão (MA), Tocantins (TO), Piauí (PI) and Bahia (BA), where it is possible to identify the rapid growth of soybean production in the last decade.
7. The preservation areas were discounted from the total areas in order to obtain an approximation for the usable area for agriculture.
8. The human development index for regions is just an approximation, once the data are available just for municipal level in 2013. For the immediate region level, we took the average of municipalities in each region.
9. The Municipal GDP Survey (*Produto Interno dos Municípios*) is an annual estimate of the municipal GDP made by IBGE, based on administrative records from the states. It presents the sectoral added value of agriculture, industries, services, public administration and the amount of taxes for all Brazilian municipalities.
10. We are not able to test for more extended time periods due to data availability.
11. The ECI for industry and services was calculated using employment for sectors, following the proposal of Freitas (2019) and Freitas et al. (2024). It seems clear from the descriptive statistics (see the supplemental data online) and model's coefficients that the relation between agricultural and non-agricultural ECI is a promising field of research in order to understand which specific economic activities in agriculture and other sectors present a complementary knowledge basis that can boost production in both sectors.
12. It is important to note here that the ECI of a region (or country) always has a perfect correlation with the average PCI of the goods it makes, which means that ECI is equal to the average of PCI multiplied or not by a multiplier factor.
13. For the relation between ECI-Agro and diversity, as well as ECI-Agro and densities, see Figures A4 and A5 in the supplemental data online.

ORCID

Alexandre de Queiroz Stein http://orcid.org/0000-0002-9978-6917
Gustavo Britto http://orcid.org/0000-0002-5285-3684
Victor Medeiros http://orcid.org/0000-0003-0451-508X

REFERENCES

Alabdulkareem, A., Frank, M. R., Sun, L., AlShebli, B., Hidalgo, C., & Rahwan, I. (2018). Unpacking the polarization of workplace skills. *Science Advances*, *4*(7), eaao6030. https://doi.org/10.1126/sciadv.aao6030

Balassa, B. (1965). Trade liberalisation and revealed comparative advantage 1. *The Manchester School*, *33*(2), 99–123. https://doi.org/10.1111/j.1467-9957.1965.tb00050.x

Balland, P.-A., Boschma, R., Crespo, J., & Rigby, D. L. (2018). Smart Specialisation policy in the European Union: Relatedness, knowledge complexity and regional diversification. *Regional Studies*, *53*(9), 1252–1268. https://doi.org/10.1080/00343404.2018.1437900

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. https://doi.org/10.1038/s41562-019-0803-3

Balland, P.-A., Broekel, T., Diodato, D., Giuliani, E., Hausmann, R., O'Clery, N., & Rigby, D. (2022). Reprint of the new paradigm of economic complexity. *Research Policy*, *51*(8), 104568. https://doi.org/10.1016/j.respol.2022.104568

Balland, P.-A., & Rigby, D. (2017). The geography of complex knowledge. *Economic Geography*, *93*(1), 1–23. https://doi.org/10.1080/00130095.2016.1205947

Bezemer, D., & Headey, D. (2008). Agriculture, development, and urban bias. *World Development*, *36*(8), 1342–1364. https://doi.org/10.1016/j.worlddev.2007.07.001

Boschma, R., Balland, P.-A., & Kogler, D. F. (2015). Relatedness and technological change in cities: The rise and fall of technological knowledge in us metropolitan areas from 1981 to 2010. *Industrial and Corporate Change*, *24*(1), 223–250. https://doi.org/10.1093/icc/dtu012

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*, 191–213.

Can, M., & Gozgor, G. (2017). The impact of economic complexity on carbon emissions: Evidence from France. *Environmental Science and Pollution Research*, *24*(19), 16364–16370. https://doi.org/10.1007/s11356-017-9219-7

Cardille, J. A., & Foley, J. A. (2003). Agricultural land-use change in Brazilian Amazonia between 1980 and 1995: Evidence from integrated satellite and census data. *Remote Sensing of Environment*, *87*(4), 551–562. https://doi.org/10.1016/j.rse.2002.09.001

Chávez, J. C., Mosqueda, M. T., & Gómez-Zaldívar, M. (2017). Economic complexity and regional growth performance: Evidence from the Mexican economy. *Review of Regional Studies*, *47*(2), 201–219.

Chinazzi, M., Gonçalves, B., Zhang, Q., & Vespignani, A. (2019). Mapping the physics research space: A machine learning approach. *EPJ Data Science*, *8*(1), 33. https://doi.org/10.1140/epjds/s13688-019-0210-z

Cimoli, M., Pereima, J. B., & Porcile, G. (2016). Introduction to the special issue SCED: Complexity and economic development. *Structural Change and Economic Dynamics*, *100*(38), 1–2. https://doi.org/10.1016/j.strueco.2016.04.004

Depaula, G. (2020). The distributional effect of climate change on agriculture: Evidence from a Ricardian quantile analysis of Brazilian census data. *Journal of Environmental Economics and Management*, *104*, 102378. https://doi.org/10.1016/j.jeem.2020.102378

Eberhardt, M., & Vollrath, D. (2018). The effect of agricultural technology on the speed of development. *World Development*, *109*, 483–496. https://doi.org/10.1016/j.worlddev.2016.03.017

Ellis, F. (2000). The determinants of rural livelihood diversification in developing countries. *Journal of Agricultural Economics*, *51*(2), 289–302. https://doi.org/10.1111/j.1477-9552.2000.tb01229.x

Fawaz, F., & Rahnama-Moghadamm, M. (2019). Spatial dependence of global income inequality: The role of economic complexity. *The International Trade Journal*, *33*(6), 542–554. https://doi.org/10.1080/08853908.2018.1535336

Felema, J., Raiher, A. P., & Ferreira, C. R. (2013). Agropecuária brasileira: Desempenho regional e determinantes de produtividade. *Revista de Economia e Sociologia Rural*, *51*(3), 555–573. https://doi.org/10.1590/S0103-20032013000300008

Felipe, J., Kumar, U., Abdon, A., & Bacate, M. (2012). Product complexity and economic development. *Structural Change and Economic Dynamics*, *23*(1), 36–68. https://doi.org/10.1016/j.strueco.2011.08.003

Ferreira, M. D. P., & Féres, J. G. (2020). Farm size and land use efficiency in the Brazilian Amazon. *Land Use Policy*, *99*, 104901. https://doi.org/10.1016/j.landusepol.2020.104901

Freitas, E., Britto, G., & Amaral, P. (2024). Related industries, economic complexity, and regional diversification: An application for Brazilian microregions. *Papers in Regional Science*, *103*(1), 100011. https://doi.org/10.1016/j.pirs.2024.100011

Freitas, E. E. (2019). *Indústrias relacionadas, complexidade econômica e diversificação regional: Uma aplicação para microrregiões brasileiras* (Doctoral dissertation, Universidade Federal de Minas Gerais, Repositório Instiucional da UFMG). http://hdl.handle.net/1843/37951

Fritz, B. S., & Manduca, R. A. (2021). The economic complexity of us metropolitan areas. *Regional Studies*, *55*(7), 1299–1310. https://doi.org/10.1080/00343404.2021.1884215

Gasques, J. G., Bonanno, A. M., Foster, L. C., & Cotula, L. (2010). Produtividade total dos fatores e transformações da agricultura brasileira: Análise dos dados dos censos agropecuários. In *A agricultura brasileira: Desempenho, desafios e perspectivas* (pp. 19–44). Instituto de Pesquisa Econômica Aplicada (IPEA).

German, L. A., Bonanno, A. M., Foster, L. C., & Cotula, L. (2020). 'Inclusive business' in agriculture: Evidence from the evolution of agricultural value chains. *World Development*, *134*, 105018. https://doi.org/10.1016/j.worlddev.2020.105018

Gollin, D., Lagakos, D., & Waugh, M. E. (2014). The agricultural productivity gap. *The Quarterly Journal of Economics*, *129*(2), 939–993. https://doi.org/10.1093/qje/qjt056

Gomez-Lievano, A., & Patterson-Lomba, O. (2021). Estimating the drivers of urban economic complexity and their connection to economic performance. *Royal Society Open Science*, *8*(9), 210670. https://doi.org/10.1098/rsos.210670

Hane-Weijman, E., Eriksson, R. H., & Rigby, D. (2022). How do occupational relatedness and complexity condition employment dynamics in periods of growth and recession? *Regional Studies*, *56*(7), 1176–1189. https://doi.org/10.1080/00343404.2021.1984420

Hartmann, D., Guevara, M. R., Jara-Figueroa, C., Aristarán, M., & Hidalgo, C. A. (2017). Linking economic complexity, institutions, and income inequality. *World Development*, *93*, 75–93. https://doi.org/10.1016/j.worlddev.2016.12.020

Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., & Simoes, A. (2014). *The atlas of economic complexity: Mapping paths to prosperity*. MIT Press.

Hidalgo, C. (2015). *Why information grows: The evolution of order, from atoms to economies*. Basic.

Hidalgo, C. A., Klinger, B., Barabási, A.-L., & Hausmann, R. (2007). The product space conditions the development of nations. *Science*, *317*(5837), 482–487. https://doi.org/10.1126/science.1144581

Hidalgo, C. A., Balland, P. A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., Glaeser, E., He, C., Kogler, D. F., Morrison, A., Nefke, F., Rigby, D., Stern, S., Zheng, S., & Zhu, S. (2018). The principle of relatedness. In Morales, Gershenson, C., Braha, D., Minai, A., & Bar-Yam, Y. (Eds.), *Unifying themes in complex systems IX: Proceedings of the Ninth International Conference on Complex Systems 9* (pp. 451–457). Springer. http://doi.org/10.1007/978-3-319-96661-8_46

Hidalgo, C. A. (2021). Economic complexity theory and applications. *Nature Reviews Physics*, *3*(2), 92–113. https://doi.org/10.1038/s42254-020-00275-1

Hidalgo, C. A. (2023). The policy implications of economic complexity. *Research Policy*, *52*(9), 104863. https://doi.org/10.1016/j.respol.2023.104863

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

Hidalgo, C. A., & Hausmann, R. (2009b). *Supporting material for: The building blocks of economic complexity* (Working Paper). Center for International Development and Harvard Kennedy School, Harvard University .

Instituto Brasileiro de Geografia e Estatística (IBGE). (2017). *Divisão regional do brasil em regiões geográficas imediatas e regiões geográficas intermediárias: 2017*.

Instituto Brasileiro de Geografia e Estatística (IBGE). (2018). *Regiões de influência das cidades: 2018*.

Instituto Brasileiro de Geografia e Estatística (IBGE). (2019). *Censo agropecuário 2017: Resultados definitivos*.

Janvry, A. d., & Sadoulet, E. (2020). Using agriculture for development: Supply- and demand-side approaches. *World Development*, *133*, 105003. https://doi.org/10.1016/j.worlddev.2020.105003

Junior, W. C. d. A., Rodrigues, M., & Silva, D. C. C. (2022). Does agricultural efficiency contribute to slowdown of deforestation in the Brazilian legal Amazon? *Journal for Nature Conservation*, *65*, 126092. https://doi.org/10.1016/j.jnc.2021.126092

Klein, H. S. (1990). The supply of mules to central Brazil: The Sorocaba market, 1825–1880. *Agricultural History*, *64*(4), 1–25.

Lall, S. (2000). The technological structure and performance of developing country manufactured exports, 1985–98. *Oxford Development Studies*, *28*(3), 337–369. https://doi.org/10.1080/713688318

Lapatinas, A., Litina, A., & Zanaj, S. (2021). The impact of economic complexity on the formation of environmental culture. *Sustainability*, *13*(2), 870. https://doi.org/10.3390/su13020870

Lewis, W. A. (1954), Economic development with unlimited supplies of labour. *Manchester School*, *22*, 139–191. https://doi.org/10.1111/j.1467-9957.1954.tb00021.x

Marconato, R., Larocca, A. P. C., & Quintanilha, J. A. (2012). Análise do uso de tecnologias em estabelecimentos agropecuários por meio dos índices de moran global e local. *Revista de Política Agrícola*, *21*(1), 5–21.

McMillan, M. S., & Rodrik, D. (2011). *Globalization, structural change and productivity growth*. National Bureau of Economic Research (NBER).

Montresor, S., & Quatraro, F. (2020). Green technologies and smart specialisation strategies: A European patent-based analysis of the intertwining of technological relatedness and key enabling technologies. *Regional Studies*, *54*(10), 1354–1365. https://doi.org/10.1080/00343404.2019.1648784

Morais, M. B., Swart, J., & Jordaan, J. A. (2021). Economic complexity and inequality: Does regional productive structure affect income inequality in Brazilian states? *Sustainability*, *13*(2), 1006. doi:10.3390/su13021006

Muneepeerakul, R., Lobo, J., Shutters, S. T., Goméz-Liévano, A., & Qubbaj, M. R. (2013). Urban economies and occupation space: Can they get 'there' from 'here'? *PLoS ONE*, *8*(9), e73676. https://doi.org/10.1371/journal.pone.0073676

Neagu, O., & Teodoru, M. C. (2019). The relationship between economic complexity, energy consumption structure and greenhouse gas emission: Heterogeneous panel evidence from the EU countries. *Sustainability*, *11*(2), 497. https://doi.org/10.3390/su11020497

Neffke, F., Henning, M., & Boschma, R. (2011). How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. *Economic Geography*, *87*(3), 237–265. https://doi.org/10.1111/j.1944-8287.2011.01121.x

Neves, M. d. C. R., Silva, F. d. F., Freitas, C. O. d., & Braga, M. J. (2021). The role of cooperatives in Brazilian agricultural production. *Agriculture*, *11*(10), 948. https://doi.org/10.3390/agriculture11100948

Niederle, P. A. (2018). A pluralist and pragmatist critique of food regime's genealogy: Varieties of social orders in Brazilian agriculture. *The Journal of Peasant Studies*, *45*(7), 1460–1483. https://doi.org/10.1080/03066150.2017.1313238

Oliveira, K. L. P. d., Andrade, B. B. d., Silva e Souza, G. d., & Castro, B. S. d. (2022). Endogeneity in stochastic production frontier with one and two-step models: An application with municipal data from Brazilian Agricultural Census. *Pesquisa Operacional*, *42*, e243504. https://doi.org/10.1590/0101-7438.2022.042.00243504

Pinto, A. (1970). Naturaleza e implicaciones de la 'heterogeneidad estructural' de la américa latina. *El trimestre económico*, *37*(145 (1)), 83–100.

Ploeg, J. D. V. d., Jingzhong, Y., & Schneider, S. (2012). Rural development through the construction of new, nested, markets: Comparative perspectives from China, Brazil and the European union. *The Journal of Peasant Studies*, *39*(1), 133–173. https://doi.org/10.1080/03066150.2011.652619

van der Ploeg, J. D., Ye, J., & Schneider, S. (2022). Reading markets politically: On the transformativity and relevance of peasant markets. *Journal of Peasant Studies*, *50*(5), 1852–1877. https://doi.org/10.1080/03066150.2021.2020258

Queiroz, A. R., Romero, J. P., & Freitas, E. E. (2023). Economic complexity and employment in the states of Brazil. *CEPAL REVIEW*, (139), 185–205. https://doi.org/10.18356/16820908-2023-139-9

Rada, N., Helfand, S., & Magalhães, M. (2019). Agricultural productivity growth in Brazil: Large and small farms excel. *Food Policy*, *84*, 176–185. https://doi.org/10.1016/j.foodpol.2018.03.014

Restrepo, D. E., Duque, J. C., & Church, R. (2022). Optimal region design to foster industrial diversification. *Regional Studies*, *56*(7), 1125–1141. https://doi.org/10.1080/00343404.2021.1986210

Reynolds, C., Agrawal, M., Lee, I., Zhan, C., Li, J., Taylor, P., Mares, T., Morison, J., Angelakis, N., & Roos, G. (2018). A sub-national economic complexity analysis of Australia's states and territories. *Regional Studies*, *52*(5), 715–726. https://doi.org/10.1080/00343404.2017.1283012

Romero, J. P., & Gramkow, C. (2021). Economic complexity and greenhouse gas emissions. *World Development*, *139*, 105317. https://doi.org/10.1016/j.worlddev.2020.105317

Ruzzante, S., Labarta, R., & Bilton, A. (2021). Adoption of agricultural technology in the developing world: A meta-analysis of the empirical literature. *World Development*, *146*, 105599. https://doi.org/10.1016/j.worlddev.2021.105599

Sambuichi, R. H. R., Galindo, E. P., Oliveira, M. A. C. D., & Pereira, R. M. (2014). A diversificação produtiva como forma de viabilizar o desenvolvimento sustentável da agricultura familiar no Brasil. In Monatério, L. M., Neri, M. C., & Soares, S. S. D. (Eds.), *Brasil em Desenvolvimento 2014: Estado planejamento e políticas públicas* (Vol. 2, pp. 61–84). Instituto de Pesquisa Econômica Aplicada (IPEA).

Sambuichi, R. H. R., Galindo, E. P., Pereira, R. M., Constantino, M., & Rabetti, M. S. (2016). *Diversidade da Produção nos Estabelecimentos da Agricultura Familiar no Brasil: Uma análise econométrica baseada no cadastro da Declaração de Aptidão ao Pronaf (DAP)* (Discussion No. 2202). Instituto de Pesquisa Econômica Aplicada (IPEA). https://hdl.handle.net/10419/144638

Santos, G. R. D., & Vieira-Filho, J. E. R. (2012). *Heterogeneidade produtiva na agricultura brasileira: Elementos estruturais e dinâmicos da trajetória produtiva recente. Texto para Discussão no. 1740*. Instituto de Pesquisa Econômica Aplicada (IPEA). https://repositorio.ipea.gov.br/handle/11058/1138

Sbardella, A., Pugliese, E., & Pietronero, L. (2017). Economic development and wage inequality: A complex system analysis. *PLoS ONE*, *12*(9), e0182774. https://doi.org/10.1371/journal.pone.0182774

Schneider, S., & Niederle, P. A. (2010). Resistance strategies and diversification of rural livelihoods: The construction of autonomy among Brazilian family farmers. *The Journal of Peasant Studies*, *37*(2), 379–405. https://doi.org/10.1080/03066151003595168

Souza, P. M. d., Fornazier, A., Da Silva, A. M., & Ponciano, N. J. (2018). Tecnologia na agricultura brasileira: Uma análise das desigualdades regionais para os segmentos familiar e não familiar. *Revista Econômica do Nordeste*, *49*(3), 147–169. https://doi.org/10.61673/ren.2018.812

Stojkoski, V., Koch, P., & Hidalgo, C. A. (2023). Multidimensional economic complexity and inclusive green growth. *Communications Earth & Environment*, *4*(1), 130. https://doi.org/10.1038/s43247-023-00770-0

Stojkoski, V., Utkovski, Z., & Kocarev, L. (2016). The impact of services on economic complexity: Service sophistication as route for economic growth. *PLoS ONE*, *11*(8), e0161633. https://doi.org/10.1371/journal.pone.0161633

Sweet, C., & Eterovic, D. (2019). Do patent rights matter? 40 years of innovation, complexity and productivity. *World Development*, *115*, 78–93. https://doi.org/10.1016/j.worlddev.2018.10.009

Sweet, C. M., & Maggio, D. S. E. (2015). Do stronger intellectual property rights increase innovation? *World Development*, *66*, 665–677. https://doi.org/10.1016/j.worlddev.2014.08.025

Tacchella, A., Mazzilli, D., & Pietronero, L. (2018). A dynamical systems approach to gross domestic product forecasting. *Nature Physics*, *14*(8), 861–865. https://doi.org/10.1038/s41567-018-0204-y

Timmer, C. P. (2002). Agriculture and economic development. In *Handbook of agricultural economics* (Vol. 2A, pp. 1487–1546). Springer. https://doi.org/10.1016/S1574-0072(02)10011-9

Traag, V. A., Waltman, L., & Eck, N. J. V. (2019). From Louvain to Leiden: Guaranteeing well-connected communities. *Scientific Reports*, *9*(1), 5233. https://doi.org/10.1038/s41598-019-41695-z

Vennet, B. V., Schneider, S., & Dessein, J. (2017). Different farming styles behind the homogenous soy production in southern Brazil. In Oliveira, G. L. T., & Hecht, S. B. (Eds.), *Soy, globalization, and environmental politics in South America* (pp. 146-168). Routledge. https://doi.org/10.4324/9781315099651

Vieira-Filho, J. E. R., & Fishlow, A. (2017). *Agricultura e indústria no Brasil: Inovação e competitividade*. Instituto de Pesquisa Econômica Aplicada (IPEA).

Zhu, S., Yu, C., & He, C. (2020). Export structures, income inequality and urban-rural divide in China. *Applied Geography*, *115*, 102150. https://doi.org/10.1016/j.apgeog.2020.102150