



The economic complexity of US metropolitan areas

Benedikt S. L. Fritz & Robert A. Manduca

To cite this article: Benedikt S. L. Fritz & Robert A. Manduca (2021) The economic complexity of US metropolitan areas, *Regional Studies*, 55:7, 1299-1310, DOI: [10.1080/00343404.2021.1884215](https://doi.org/10.1080/00343404.2021.1884215)

To link to this article: <https://doi.org/10.1080/00343404.2021.1884215>



View supplementary material [↗](#)



Published online: 15 Mar 2021.



Submit your article to this journal [↗](#)



Article views: 860



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 3 View citing articles [↗](#)



The economic complexity of US metropolitan areas

Benedikt S. L. Fritz^a  and Robert A. Manduca^b 

ABSTRACT

We calculate measures of economic complexity for US metropolitan areas for the period 1998–2015 based on employment data. We show that the concept translates well to the regional setting and to local and traded industries. Large cities and the Northeast have the highest complexity, while most traded industries are more complex than most local ones. In cross-section, metropolitan complexity is associated with higher incomes, though to a lesser extent recently than in the past. However, within-city increases in complexity from year to year are associated with income decreases. Our findings highlight the need for caution when interpreting the relationship between complexity and socioeconomic outcomes.

KEYWORDS

economic complexity; United States; urban economic development; income; economic geography

JEL R1, O1, O40

HISTORY Received 8 September 2019; in revised form 4 January 2021

INTRODUCTION

Measures of economic complexity, which characterize and analyse economies on the basis of their productive structures, are widely used to study national economies (Hidalgo & Hausmann, 2009; Tacchella et al., 2012). In recent years, there have been increasing attempts to adapt them to the study of subnational and metropolitan regions.

Initial versions of these metrics have been implemented at the subnational level in China, Australia, Mexico, Russia, New Zealand, Italy and the United States, based on different kinds of data, such as patents, exports, and employment (Balland & Rigby, 2017; Chávez et al., 2017; Davies & Maré, 2019; Gao & Zhou, 2018; Gomez-Lievano & Patterson-Lomba, 2019; Lyubimov et al., 2018; Mealy et al., 2019; Reynolds et al., 2018). These applications of regional complexity measures join a strong tradition of analysing the productive structure of sub-national economic units in the field of evolutionary economic geography (Frenken et al., 2007; Neffke, 2009; Neffke et al., 2011). Measures of economic complexity are also being used in the policy sphere as indicators of future growth potential (Balland et al., 2019; Escobari et al., 2019) and, for example, have become commonly used by Chile's main state funding agency (CORFO, 2018).

However, key questions remain about the applicability of methods designed for countries to the subnational

setting. Additionally, there are concerns as to whether these methods are extendable to service sectors, which dominate the economies of most developed countries (Buera & Kaboski, 2012; Stojkoski et al., 2016), and to local industries alongside the traded industries that they were developed for.


In this paper,¹ we extend the research on economic complexity by exploring three questions:


- Is the concept of economic complexity transferable to the regional context, and if so, how is it best implemented?
- What can we learn about the economic geography of the United States, and its evolution over time, through this measure?
- What is the predictive potential of such measures for regional growth?

We produce economic complexity measures for metropolitan areas in the United States between 1998 and 2015. Using employment data from the US Census County Business Patterns (CBP), we are able to include almost all types of economic activity in the country. We show that the spatial distribution of economic activity across US metro areas is suitable for economic complexity analysis, in line with previous results. Major cities with developed economies tend to have industries of all types,

CONTACT

^a (Corresponding author)  benedikt.fritz@economics.ox.ac.uk
Department of Economics, University of Oxford, Oxford, UK.

^b  rmanduca@umich.edu
Department of Sociology, University of Michigan, Ann Arbor, MI, USA.

 Supplemental data for this article can be accessed at <https://doi.org/10.1080/00343404.2021.1884215>

common as well as specialized: Los Angeles produces handbags as well as guided missiles. But less developed metros tend to have only common industries. This pattern of ‘nestedness’, which is well documented across countries (Bustos et al., 2012), is core to American economic geography. This finding holds for both local industries (e.g., those aimed at within-region demand, such as bakeries) and traded ones (e.g., those aimed at cross-region demand, such as guided missiles and handbags), so we argue that both should be included in complexity analysis.

We also show that the differences between the two most prominent economic complexity algorithms – the economic complexity index (ECI) and the fitness index (FI) (Albeaik et al., 2017a; Pietronero et al., 2017) – are relatively small in the context of the United States. More consequential is the method for determining the presence of an industry in a region. Previous work has typically considered industries present in a region only if they have a higher than average footprint there. We show that this approach risks systematically underestimating the capabilities of large and diverse cities – particularly in local industries, which are known to scale sublinearly with city size (Youn et al., 2016). To address this source of potential bias, we introduce a measure of presence based on a combination of the location quotient (LQ) and a simple threshold of whether total industry employment in a region exceeds 50 employees.

We find that the areas with the highest complexity are major cities, especially Los Angeles, New York and Chicago, while traded industries tend to rate higher on complexity than local ones. The Northeast of the United States emerges as the most complex part of the country on average, though the stronger distinction is between metropolitan and rural areas. Regions were highly stable in their economic complexity over our sample period: the Pearson correlation between ECI in 1998 and that in 2015 was 0.96. In cross-sectional regressions, cities with higher complexity have significantly higher per capita incomes, controlling for a number of social, economic and institutional characteristics. However, this relationship has weakened over time, largely because of income growth in low-complexity areas, particularly those with large resource extraction industries. This income growth has occurred without corresponding changes in the economic complexity of these regions, weakening the overall relationship.

The changing relationship between economic complexity and earnings highlights the contingent nature of connections between regional productive structure and income. While local incomes are shaped by local productive capacity, they are also strongly influenced by other factors, including political institutions and global macroeconomic context. Results on within-region changes in income and complexity, which also show a fluctuating relationship over time, emphasize the importance of caution when making claims about any fundamental relationship between productive structure and income.

The remainder of the paper is structured as follows. The next section gives a short review of previous research

on economic complexity and regional economic development. The following sections introduce the data and the complexity methodology and discuss several theoretical and methodological challenges associated with constructing a subnational version of complexity. After analysing the general transferability of economic complexity into the regional context, we conduct and discuss our regression analyses. The paper ends with a summary and implications for future research.

LITERATURE REVIEW

Studying productive structures: the capabilities approach

Economic complexity measures draw on what researchers have termed the ‘capabilities approach’ to understanding economic development (Hausmann et al., 2013; Hidalgo & Hausmann, 2009). In this approach, an economy is seen as a system of knowledge accumulation, and its prosperity depends on its ability to transform that knowledge into useful products (Hidalgo, 2015). Productive knowledge is thought to exist in discrete units, or capabilities, such as the capability to weld metal or the capability to spin thread. These capabilities are combined to produce the prime economic outputs of the system: its products and services (Hidalgo & Hausmann, 2009; Hidalgo et al., 2007).

The key claim of the capabilities approach is that while it is typically impossible to directly observe capabilities, they can be inferred from the presence of industries: if an economy is able to competitively produce handbags, say, then it must have access to whatever capabilities go into handbag production – from sewing and tanning to fashion design and factory organization. By studying the collection of products that are produced, the portfolio of capabilities that must exist in a given economy can be determined. Implications for development policy have been that countries can most easily achieve growth by diversifying into closely related industries/products and that it is hard to expand directly into less similar industries/products (Hidalgo & Hausmann, 2009). The same has been observed for regions (Balland et al., 2019; Ciccone et al., 2020; Daboín et al., 2019; Kogler et al., 2013; Mealy & Coyle, 2019; Neffke et al., 2011). For an overview of the branching out literature, see Hidalgo et al. (2018).

Empirically, when the product portfolios of different countries are constructed, they often tend to follow a distinct pattern of ‘nestedness’, in which advanced economies produce both common and uncommon products, while less developed economies generally produce only common products (Bustos et al., 2012). This empirical regularity deviates from that predicted by standard economic theories of trade under comparative advantage. Researchers condense the information of the full product space matrix into a one-dimensional indicator in order to be able to employ it in empirical investigations. The economic complexity index (ECI) (Hausmann et al., 2013; Hidalgo & Hausmann, 2009) and the fitness index (FI) (Tacchella

et al., 2012) are two such dimensionality reduction techniques.

Complexity economics has become widely used in policymaking (Balland et al., 2019; Escobari et al., 2019). Prominent examples are the government of Chile (CORFO, 2018), projects of the Harvard Growth Lab with governmental agencies of (among others) Saudi Arabia and Albania (The Growth Lab, 2020), and the inclusion of the FI in the World Bank data repository (World Bank, 2018).

Studying regional economies

There are compelling theoretical and empirical reasons to use metropolitan areas as a basic unit for studying economic geography. Metro areas capture individual commuting sheds and are designed such that jobs within a given metro are held by residents of that metro, and vice versa. This makes them coherent in a way that municipalities (which often have large net flows of commuters in or out) and states/provinces (which may include multiple self-contained commuting sheds) are not. From a theoretical perspective, metropolitan areas are the geographical entity at which agglomeration begins (Glaeser et al., 1992; Marshall, 1890), and at which new industries are created (Jacobs, 1969). They can be functionally defined as the geographies in which face to face meetings can be undertaken without friction (Storper & Venables, 2004).

Empirically, the variation in economic performance among metros within a country is often comparable with the difference between countries, even countries at very different levels of development. For instance, the gross domestic product (GDP) per capita of the San Francisco metropolitan statistical area (MSA) in 2017 was US \$90,000 – more than four times that of the McAllen Texas MSA (US Bureau of Economic Analysis, 2017). This is roughly the ratio between the GDP per capita of the United States as a whole and that of Peru (International Monetary Fund (IMF), 2017).

For these reasons, regional scientists have long sought to understand the economic performance of metropolitan regions and have used a wide swath of metrics and concepts to characterize regional economies (Isard & Reiner, 1966; Storper, 2011). Many of these, such as the capabilities approach, are focused on the productive structure of metro areas, particularly on their export industries. Examples include studies of regional ‘economic bases’ (Andrews, 1953; Heilburn, 1981), ‘growth poles’ (Parr, 1973; Perroux, 1955), industry clusters (Porter, 1998) and industrial variety (Frenken et al., 2007).

A growing literature, mostly within evolutionary economic geography, employs the capabilities approach at the regional scale (Cicerone et al., 2020; Daboïn et al., 2019; Essletzbichler, 2015; Farinha et al., 2019; Kogler et al., 2013; Mealy & Coyle, 2019; Muneeppeerakul et al., 2013; Neffke et al., 2011). Like work on the productive structures of nations, these studies seek to construct the ‘product space’ within a given country, and identify paths that subnational regions can take to develop their economies by branching into industries that are related to the ones

they already have. In Europe, this relatedness approach has been used to justify regional development policies that differ from place to place depending on location within the product space (Balland et al., 2019).

In the case of the United States, recent years have provided both a normative and a positive case for focusing on regional development (Storper, 2018). Since 1980, the long-term trend of regional convergence in income levels has stalled and begun to reverse (Amos, 2014; Manduca, 2019). This divergence has led to a great deal of interest in regional economic performance in the popular media, and to calls for policymakers to treat regional economic development as a core national policy concern (Badger et al., 2016; Leonhardt, 2018).

Previous research on regional economic complexity

There have been several previous attempts to apply the lens of economic complexity to regional or subnational productive structures. These fall into two major groups. Some studies have used economic complexity measures computed from international trade data, as in the original cross-national studies, and then calculated the complexity of regions by taking dollar-weighted averages of their industry portfolios. Reynolds et al. (2018) describe the productive structure of Australia’s states and find that minor differences in inter-state as well as rest-of-world exports matter greatly to relative complexity. Operti et al. (2018) apply a similar method to Brazilian states and show that the more economically developed states are fairly stable in their levels of complexity, while the less developed states show more fluctuation over time.

The second group consists of studies that, like the present paper, calculate regional complexity using the national rather than global economic network. Gao and Zhou (2018) calculate the ECI and FI for Chinese provinces based on the number of publicly traded firms. Despite the limited coverage of these data, which they acknowledge, they corroborate cross-country findings of a positive relationship between complexity and growth. Instead of using the number of firms per sector, Chávez et al. (2017) use employment across sectors to compute the ECI of Mexican states and provide evidence that the higher a state’s ECI is, the higher its GDP per capita. Davies and Maré (2019) use a continuous rather than discrete measure of industry presence for urban areas in New Zealand. They show that increases in complexity are associated with increases in employment.

In the United States, a handful of studies have calculated measures of economic complexity for subnational regions, but usually as an example of a larger theoretical point or a comparison to a preferred measure rather than an object of analytical interest in itself (Gomez-Lievano & Patterson-Lomba, 2019; Mealy et al., 2018, 2019; Sbardella et al., 2017). A notable exception is Balland and Rigby (2017), who offer one of the first measures of complexity for US metro areas, focusing specifically on the development of new technology. Using detailed patent data for US metropolitan areas between 1975 and 2010,

they find that high-complexity patenting tends to be geographically sticky and that the highest ranked places (in terms of patent complexity) are not exclusively the most populous nor those with the highest income per capita.

While the papers described here represent important advances in our understanding of subnational economic complexity, they have not fully engaged with major methodological questions that emerge when translating measures developed for international exports to the subnational context, such as what types of industries should be included in complexity metrics and how to determine when an industry is present. Additionally, with the exceptions of Balland and Rigby (2017) and Davies and Maré (2019), they have typically been conducted at a coarse level of either geography or industrial classification. As we show in Appendix F4 in the supplemental data online, many of the stylized facts underpinning complexity analysis weaken at coarse aggregation levels. Here, we build on these previous studies by using data with high resolution in terms of both space and industry category for US metropolitan areas. We begin by describing our data and addressing several major methodological questions that arise when moving from the national to the metropolitan level of analysis.

DATA

We use data from the US Census County Business Patterns (CBP), an annual employment count by county and North American Industry Classification System (NAICS) industry classification for the entire United States, for the period 1998–2015. The dataset covers approximately 84% of the civilian employment level for March 2015 (US Bureau of Labor Statistics, 2017).

We define industries by six-digit NAICS code, the most disaggregated level available. Appendix F4 in the supplemental data online shows that the assumptions underlying the ECI tend to break down when industries are aggregated to levels coarser than four-digit NAICS. For geographical aggregation, we use core-based statistical areas (CBSAs) as defined by the US Office of Management and Budget based on commuting flows (US Census Bureau, 2012). CBSAs consist of MSAs and micropolitan statistical areas, and are the most common way to define metropolitan areas in the United States. We include non-core counties as individual observations in the calculation of complexity but aggregate them to one observation per state in our regressions because of data limitations. By calculating complexity based on the complete geography of the United States, we ensure that our measures are as inclusive of the employment network in the United States as possible. While other geographical agglomerations are available, CBSAs are the most widely used and are based on commuting patterns. Further information on the data, the industry classification schemes, and the geographic aggregation can be found in Appendix B in the supplemental data online. The code behind our dataset and all our results can be found online.²

An advantage of the CBP data is that they are based on employment rather than exports. Researchers have noted the limitations of measuring overall economic competitiveness based on exports alone, since even at a similar level of development some economies may export more than others (Morrison et al., 2017). Additionally, the employment-based CBP data includes services – which account for 80% of employment in the United States (US Bureau of Labor Statistics, 2020) – as well as goods producing industries. While some papers on economic complexity have included services in their calculations (see especially Hausmann et al., 2013; Stojkoski et al., 2016), they have been absent from most analyses.

However, no data source comes without disadvantages. A conceptual limitation of employment data is that they may underweight the presence of highly efficient industries, which produce more output at a given level of employment. A further drawback of the NAICS industry classifications specifically is that many firms and establishments produce more than one type of product, yet they are classified in just one industry category, based on where the majority of their output is located. That said, the classification system is designed to group establishments into industries based on the similarity of their production processes (Office of Management and Budget, 2017), and only rarely do establishments produce in more than one four-digit NAICS industry group.

To see if economic complexity adds additional power to predictions of metropolitan income per capita beyond standard measures, we use a number of control variables. These can be clustered into economic controls, sociodemographic controls and institutional factors (Breau et al., 2014; Florida & Mellander, 2016). For the full list of sources and a further explanation for all variables employed in our analysis, see Appendix A in the supplemental data online.

METHODOLOGY

This paper studies the economic complexity of US metropolitan areas. The most frequently used indicator of complexity is the economic complexity index (ECI) (Hausmann et al., 2013; Hidalgo & Hausmann, 2009). A second and strongly related indicator is the fitness index (FI) (Tacchella et al., 2012). The formula and logic behind each of these indicators are described in Appendix C in the supplemental data online.

There has been a great deal of debate recently about the relative merits of the ECI and the FI (Albeaik et al., 2017a, 2017b; Gabrielli et al., 2017; Pietronero et al., 2017). However, for regions of the United States, they are extremely highly correlated. As shown in Figure F3 in Appendix F in the supplemental data online, in 2015 the correlation of ECI with the log of FI for US metro areas was 0.94. Thus, it appears that the primary difference between the ECI and the FI, at least as computed on these data, is one of scaling rather than a fundamental difference in concept. Because it produces a less skewed distribution of values and has fewer concerns about convergence

(Morrison et al., 2017), we choose to focus on ECI here, though we reiterate that both measures appear to be capturing the same underlying construct with remarkable consistency, and results using the FI are similar to those with the ECI.

The ECI is sometimes interpreted as measuring diversity. However, it is in fact mathematically orthogonal to diversity and instead of ranking places according to their degree of diversification rather works as a dimensionality reduction algorithm which defines a distance between nodes in a graph on the basis of their similarity (Mealy et al., 2019). It does this through reciprocal averaging – what Hidalgo and Hausmann (2009) call the ‘method of reflections’ – where a region’s ECI is the average of the product complexity indices (PCIs) of products it is specialized in.

Applying economic complexity to metropolitan areas

Although it was developed to study countries rather than metropolitan areas, we believe the ECI is well suited to the metropolitan context. Conceptually, the ECI was developed to study systems of self-contained economic units connected by trade. That definition applies to metropolitan areas within a country just as it does to countries within the world. Empirically, as we show in Figure 1, industries are distributed across US metros in the same nested pattern as products often are across countries.

Nonetheless, the changed context does raise certain conceptual and empirical questions. For example, the total lack of trade barriers between regions of a country may induce regional specialization that is due to economies of scale rather than capability limitations. There may also be products that are relatively simple to produce that are nonetheless only made in few regions because a

few places are sufficient to meet the demand of the whole country. This phenomenon may occur for more sophisticated industries as well. For instance, the 12 regions chosen as headquarters of Federal Reserve branches might not be the only ones endowed with the necessary capabilities for central banking, but nevertheless they suffice under the present system to supply the whole nation.

Whether this is analytically problematic is debatable: even if other regions could easily develop the capabilities to produce simple products – or even central banks – they do not currently have them. Still, we remain alert to this possibility, which would most likely manifest in seemingly simple industries being identified as complex because they only appear in a handful of metros.

Two further methodological questions that arise when translating economic complexity to the US metropolitan context are whether to include local industries as well as traded ones, and how to define when an industry is present in a region. We consider each of these in turn.

The role of local industries

Regional scientists often find it helpful to distinguish between local and traded industries. Local industries are those that meet the needs of people living in their region, while traded industries are primarily aimed at producing exports to other regions. Because most previous work on economic complexity is based on international export data, it has been limited to traded industries by necessity. We do not face this data constraint, but the question remains as to whether local industries logically fit with complexity analysis.

The concern about including local industries in the complexity score arises because they play a fundamentally different role in the regional economy from that of traded industries. Many theorists have noted the importance of

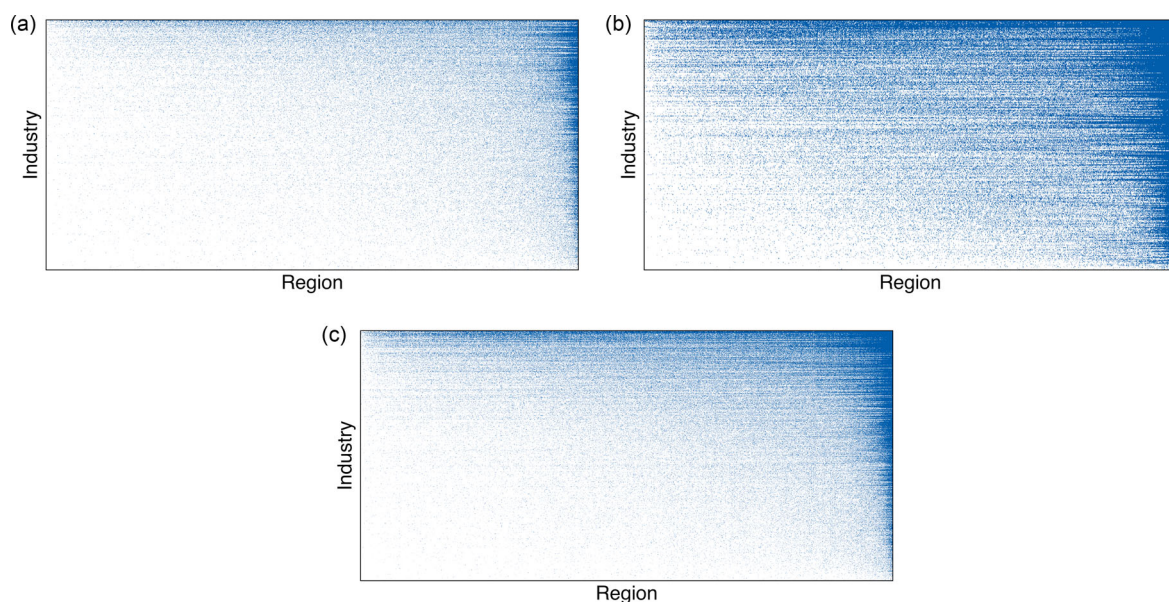


Figure 1. Industry presence ordered by ubiquity and diversity: (a) traded, (b) local and (c) all industries (CM_{r,i} input matrix six-digit NAICS Industries, 2015).

Sources: Authors' calculations; Delgado et al. (2016); County Business Patterns (CBP).

exports in bringing money into a region from the outside (Andrews, 1953; Parr, 1973). The money earned by these export industries in turn supports local industries, which form the bulk of employment (Moretti, 2010; Porter, 2003). For the purposes of measuring regional economic performance, it may thus make sense to focus only on the export industries believed to drive economic growth (cf. Gomez-Lievano & Patterson-Lomba, 2019).

However, there are both conceptual and empirical counterarguments to this reasoning. First, local industries can vary in complexity and ubiquity just like anything else. Consider the case of restaurants – perhaps the quintessential local service. Although restaurants exist within every metro area in the country, a big city such as New York has a much wider variety of restaurants than a small town. That variety indicates a much wider range of capabilities related to food service: the capabilities to use spices from multiple ethnic cuisines, for instance, or those required to run a Michelin-starred kitchen. Importantly, these capabilities contained in local industries can still contribute to the development of new traded industries. For example, the internet meal subscription company Blue Apron was founded by combining the expertise of a chef running a local catering business with two tech industry workers (Konrad, 2015). The full economic capacities of a place can be characterized more accurately when all its economic activity is being taken into account.

In addition, a large and diverse set of local services for consumption purposes has been proposed as a source of competitive advantage for cities: having a wide array of local restaurants, boutiques, entertainment venues, and art galleries is thought to attract members of the ‘creative class’ with high levels of human capital (Clark et al., 2002; Florida, 2002).

Empirically, looking at local industries collectively, it turns out that they have a similar spatial distribution to traded ones. Figure 1 shows the presence of traded, local and all industries across metros, ordered by ubiquity and diversity, where industries are classified as local or traded according to Delgado et al. (2016). The overall nested structure documented by Bustos et al. (2012) for exports across countries is present for local, traded and all industries across US metros. We quantify this nestedness pattern using the NODF (Nestedness metric based on Overlap and Decreasing Fill) score, a measure of nestedness used in ecological studies that varies from 0 (no nestedness) to 100 (perfect nestedness) (Almeida-Neto et al., 2008). The nestedness of industries within US metros is 49.87 for all industries together, 60.06 for local industries and 30.77 for traded industries. This compares with a value of 70.81 in the country-product network reported by Bustos et al. (2012). Note that the nestedness pattern in our data on US CBSAs stands in contrast to the lack of nestedness found by Mealy et al. (2019) in their analysis of occupations and US states, perhaps because of the finer resolution in our data (see Appendix F4 in the supplemental data online).

Given the importance of local industries and their comparable geographical distributions, we opt to use both traded and local industries in the main text. However,

we also give results for complexity values calculated solely based on traded as well as solely based on local industries in Appendix F2 in the supplemental data online. If anything, the cross-section effects in the regression of income per capita on ECI are stronger when ECI is computed based on local industries alone than when it is computed based on traded industries alone or all industries combined.

The measurement of industry presence

Economic complexity measures are constructed from a matrix of industry presence by geography. A key question is how to define when an industry is ‘present’ in a region. Previous research has typically used a binarized version of revealed comparative advantage (RCA), counting a given product as present in a given country if it is overrepresented in that country’s basket of exports relative to its share in the rest of the world. The equivalent approach when using employment data is to consider an industry present in a metro if its $LQ > 1$, indicating that workers in that industry form a larger percentage of local employment than they do for the nation as a whole.

This is a reasonable starting point, but it has some weaknesses. The most notable of these is that it tends to under-report the production of common goods in diversified regions. Consider a country with two regions, one that produces just apples and one that produces equal quantities of apples and computers. The latter is clearly the more complex economy because it has both the capabilities that go into making apples and those that go into making computers. However, it would have a LQ in apples of < 1 , because apples form a smaller portion of its economy than they do for the nation as a whole. Arithmetically, regions with more diverse economies will tend to have $LQs < 1$ for ubiquitous products simply because they are also producing uncommon products. We argue that this means a binarized version of the LQ will under-report the presence of common industries in diverse regions.

This issue is of particular concern given the inclusion of non-traded sectors in our analysis. Previous research has documented that most local industries scale sublinearly with city size (Youn et al., 2016). This means that local industries are likely to have lower LQs in large regions, simply because of how their presence scales with population. One example of this is the case of gas stations with convenience stores (NAICS code 447110) in New York City. In 2015, the New York MSA had 13,972 employees at these gas stations, clearly an indication of the capability to provide cars with fuel. Yet, as a share of total metropolitan employment, gas stations were underrepresented in New York, with $LQ = 0.27$. Using the standard approach of $LQ > 1$, New York would be identified as lacking the capabilities involved in running gas stations. We believe this conclusion would be inappropriate.

To address this limitation, in our primary analysis we use an alternative measure of industry presence in which we consider an industry to be present in a metro if $LQ > 1$ or if total employment is > 50 . This ‘composite matrix’ (CM) allows ubiquitous industries to be correctly

identified as present in diverse metro areas while still identifying the industries that smaller metros specialize in.

While we believe the CM matrix is an improvement on the standard LQ matrix, particularly for the analysis of local industries, it does have some drawbacks. It might for example blur the differences between the most complex and medium complexity regions by inferring full capabilities in an industry from the presence of a relatively small number of workers. A full comparison of our measure with the standard $LQ > 1$ measure and an alternative measure based on the presence of any employment in an industry whatsoever is provided in Appendix D in the supplemental data online. Future work in this area might also consider fully continuous alternatives to the ECI, such as those proposed by Davies and Maré (2019) and Gomez-Lievano and Patterson-Lomba (2019).

RESULTS

Complexity and the economic geography of the United States

Figure 2 maps the economic complexity of CBSAs and non-metropolitan counties in 2015. The most complex regions are the metropolitan areas of Los Angeles, New York, Chicago, Philadelphia and Boston. These are large cities with very diverse economies, producing all types of goods and services. The least complex regions tend to be rural counties, most notably those in the Great Plains. These areas' employment is centred on industries that are common in more diverse regions as well.

There are both high- and low-complexity metros in all parts of the country, including the Northeast, South, Midwest and West. However, there is a very tight relationship between population size and economic complexity, with ECI correlated with the log of metro population size at 0.85 in 2015. This finding stands in some contrast to previous results at the country level, where many relatively small countries rank quite high. This may be because there is less overall variation in the level of economic development within the United States than there is across

countries, so sheer size matters more than other determinants of productive capacity. Note that this result is not driven by our approach to measuring industry presence: if we use the typical approach of $LQ > 1$, instead, the correlation is 0.81. Nor is it driven by our choice to include both local and traded industries – correlations between the log of population and ECI computed with local and traded industries alone are 0.73 and 0.7, respectively. In terms of broader parts of the country, the Midwest Census Division has the lowest average complexity, followed by the South, West and Northeast.

There is remarkable stability in the economic complexity of metro areas over time. The 1998 and 2015 values for ECI are correlated at 0.96. In total, variation across time within metros accounts for just 1.7% of total variation in ECI across both space and time. This suggests that there was relatively little change in the productive structures of US metros during our sample period compared to the overall variation across metros that already existed in 1998. That said, there was fluctuation in both raw and rank terms from year to year, and some CBSAs did steadily rise or fall. The Villages, Florida, for example, saw remarkable development during this period, growing from an ECI of -0.22 (rank 921 out of 978 CBSAs) to one of 0.79 (rank 556). Moving in the other direction, Wheeling, West Virginia, saw its ECI drop from 1.27 (rank 205) in 1998 to 0.88 (rank 497) in 2015. The ranks over time are plotted in Figure E1 in Appendix E in the supplemental data online. The core takeaway from Figure E1 is that most changes occur at the middle ranks whereas the top and the bottom ranks are more consistent.

Zooming into industries

Table 1 presents the 10 most and least complex industries in 2015. Broadly, the most complex industries were found in finance and advanced manufacturing, while the least complex industries tended to relate to resource extraction, agriculture and logistics. The appearance of central banking, whose geography was set by the Federal Reserve Act of 1913, highlights how the processes that determine industry presence can vary. However, we do not believe this is a weakness because – as discussed above – only the 12 designated regions chosen as headquarters of the Federal Reserve branches have fully developed capabilities in central banking.

Of the 10 highest complexity industries, only two are local. These are both subsumed under the NAICS industry group 4851 'Urban Transit Systems'. This points to the underrepresentation of public transportation in many of the smaller regional units included in our analysis. Local and traded industries are roughly equally represented among the 10 least complex industries. On average, local industries had a complexity of -0.638 , whereas traded ones had an average of 0.286 (note that complexity values are normalized to have mean 0 and variance 1, with scores > 0 indicating greater than average complexity). This suggests that local industries on the whole use less specialized capabilities than traded ones. However, 65 local

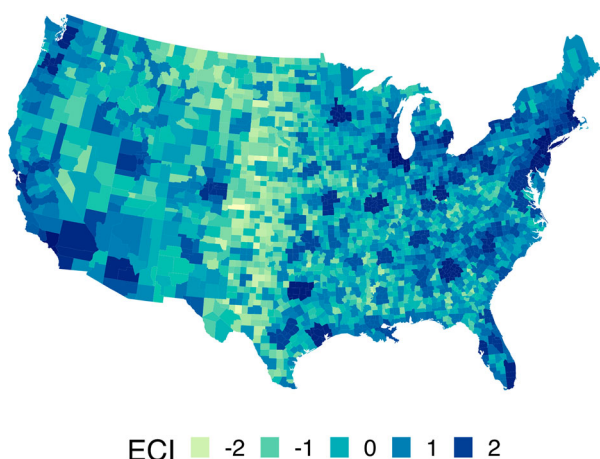


Figure 2. ECI of US core-based statistical areas ($CM_{r,i}$ input matrix, six-digit NAICS Industries, 2015).

Source: Authors' calculations; County Business Patterns (CBP).

Table 1. Highest and lowest complexity industries (ECI, $CM_{r,i}$ input matrix, six-digit NAICS industries, 2015).

	NAICS	Name	ICI	Type
1	523210	Securities and Commodity Exchanges	3.71	Traded
2	521110	Monetary Authorities – Central Bank	3.00	Traded
3	485112	Commuter Rail Systems	2.65	Local
4	485119	Other Urban Transit Systems	2.65	Local
5	522190	Depository Credit Intermediation	2.51	Traded
6	334517	Irradiation Apparatus Manufacturing	2.30	Traded
7	336415	Guided Missile and Space Vehicle Propulsion Unit ... Manufacturing	2.22	Traded
8	336414	Guided Missile and Space Vehicle Manufacturing	2.18	Traded
9	524130	Reinsurance Carriers	2.16	Traded
10	334112	Computer Storage Device Manufacturing	2.12	Traded

969	115111	Cotton Ginning	–2.20	Traded
970	447190	Other Gasoline Stations	–2.21	Local
971	424910	Farm Supplies Merchant Wholesalers	–2.21	Traded
972	115112	Soil Preparation, Planting, and Cultivating	–2.30	Traded
973	484220	Specialized Freight ... Trucking, Local	–2.31	Local
974	213111	Drilling Oil and Gas Wells	–2.46	Traded
975	813910	Business Associations	–2.51	Local
976	424510	Grain and Field Bean Merchant Wholesalers	–2.52	Traded
977	213112	Support Activities for Oil and Gas Operations	–2.76	Traded
978	211111	Crude Petroleum and Natural Gas Extraction	–2.78	Traded

industries (roughly one-fifth of the total) have higher than average complexity, including NAICS codes 561613, ‘Armored Car Services’; 621491, ‘HMO Medical Centers’; and 515120, ‘Television Broadcasting’. The remaining 238 local industries have lower than average complexity. In contrast, 452 of the 675 traded industries are more complex than average.

The changing relationship between economic complexity and economic performance

We now turn to the relationship between economic complexity and the economic performance of US metros over the period 1998–2015. We use per capita income as

calculated by the Bureau of Economic Analysis as our outcome measure and consider both the bivariate relationship between ECI and income and the relationship after controlling for a range of economic, social and institutional factors. Before doing so, we reiterate that our investigation and regressions are not primarily aiming at a causal analysis. Rather, we want to draw a richer picture of the economic network of the US metro areas and show in which respects, for example, high- and low-income per capita places differ. Any further claims can only be of a very limited nature due to possible issues of endogeneity and are complicated by our findings on the high intertemporal stability of ECI over time.

We begin by testing how higher levels of ECI are related to cross-sectional variation in regional economic outcomes. Figure 3(a) shows coefficients from yearly bivariate population weighted cross-sectional regressions of logged income per capita on ECI. ECI is always a positive and highly significant predictor of cross-sectional increases in income per capita. However, this relationship has noticeably weakened over time: around 2007 it started to decrease, such that the estimator for 2015 is substantially lower than that for 1998. In absolute terms, while in 1998 a one-unit difference in ECI between regions was associated with approximately a 17% difference in income per capita, the same difference in ECI in 2015 was only associated with a 12% difference in income per capita. Looking at the R^2 (see regression Table E1 in Appendix E in the supplemental data online), we also see that the explanatory power decreased markedly (from 40% to 17% between 1998 and 2015) and the correlation between the variables fell from 0.58 to 0.33 (Figure 3c, d).

Full regression output for 1998, 2007 and 2015, including results with 11 sociodemographic controls in the latter two years, is presented in Appendix E2 in the supplemental data online (control data are not available before 2007, and are only observed for the roughly one-third of regions that have sufficient population). The results with controls conform to the bivariate findings of an initially positive relationship that declines over time. In 2007, controlling for our set of economic, social and institutional variables (which reduces the number of regions we have data for from 932 to 284), ECI has a significant coefficient of 0.07. In 2015, this coefficient is reduced to 0.02 and has become insignificant.

Behind these changes

As discussed above, metropolitan ECI levels did not change dramatically during this period, not even in the years around 2007. That means that the weakening relationship between ECI and income must be a function not of changes to metropolitan productive structure but of changes to metropolitan income that were *not* directly tied to changes in industrial portfolio. Figure 3(c, d) explores this relationship, plotting per capita income and ECI by CBSA in 1998 and 2015, respectively. The solid lines show the fit of a quadratic regression function of per capita income on ECI.

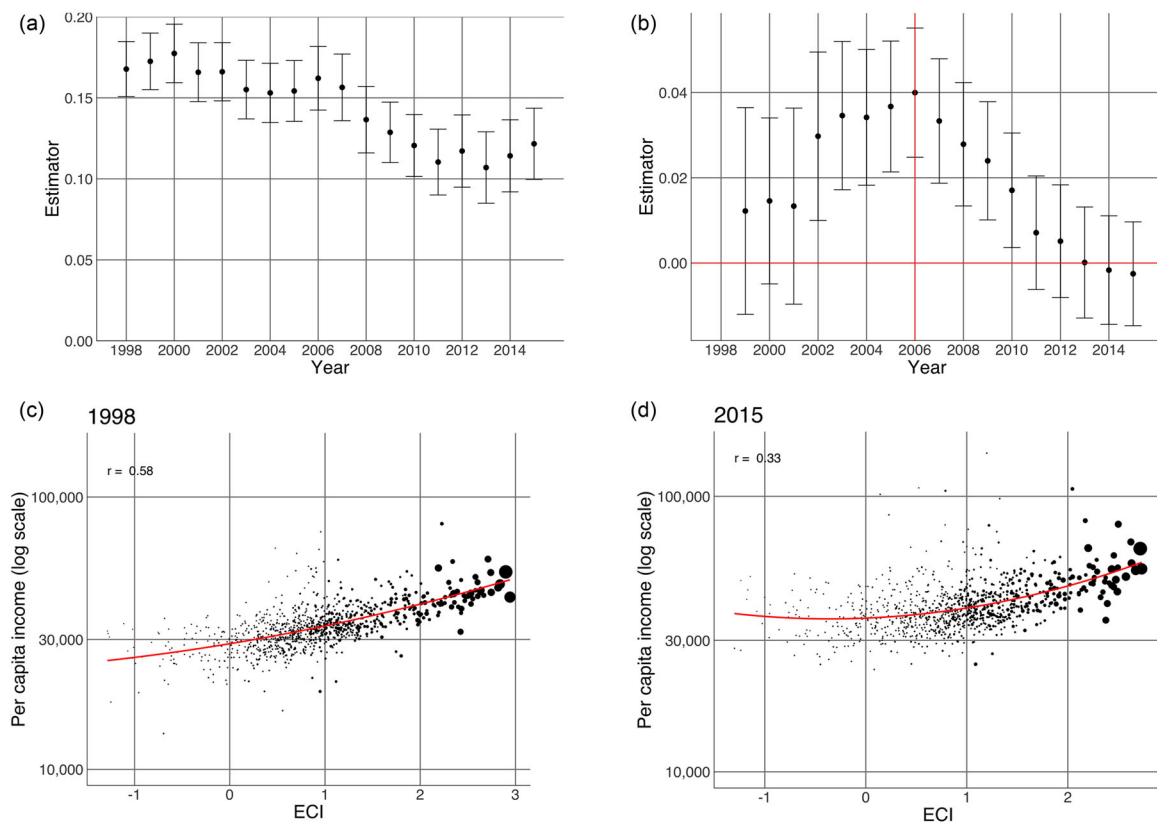


Figure 3. Different aspects of the changing relationship between income per capita and ECI (CM_{ri} input matrix, six-digit NAIS industries): (a) coefficients from yearly bivariate cross-section regression of income per capita on ECI; (b) coefficients from rolling fixed effects regression of income per capita on ECI adding one year at a time; (c) scatter plot of income per capita and ECI, 1998 (dot size is proportional to the 1998 population); and (d) scatter of income per capita and ECI, 2015 (dot size is proportional to the 2015 population).

Source: Authors' calculations; County Business Patterns (CBP).

In 1998, the relationship between ECI and per capita income was generally positive across the entire spectrum, a pattern similar to the relationship found across countries. The curvature of the quadratic fit is barely visible. On average the highest-complexity metros had the highest incomes. By 2015, however, the relationship had become slightly 'U' shaped, because the low-complexity regions had made bigger gains (in relative and absolute terms) in income per capita than the high-complexity ones. Since metro ECI values did not change much during this period, it is the growth of incomes in low-complexity metros that appears to have altered the overall relationship.

To investigate the drivers of income growth in low-complexity metros, we explore the industry makeup of those CBSAs and non-metropolitan counties that had ECI values below the mean (0) and per capita incomes above the mean (US\$40,402) in 2015. These 114 metros collectively had a population of 4.6 million in 2015, up from 4.2 million in 1998. On average, their per capita income grew from US\$34,796 in 1998 to US\$48,573 in 2015 – an increase of more than US\$13,700 during a time where the overall national per capita income grew by just US\$6837. Over half of them are found in the states of Texas, Maine, Nebraska, South Dakota, Iowa, Kansas, Oklahoma and Wyoming.

Employment in these counties was heavily overrepresented in the mining (NAICS 21; LQ = 7.4) sector. The top 10 industries in terms of LQ in these regions were entirely in mining and manufacturing, with Gold Ore Mining (NAICS 212221) and Potash, Soda, and Borate Mineral Mining (NAICS 212391) at the top.

Thus, it appears that the weakening cross-sectional relationship between metropolitan economic complexity and per capita incomes was the result of fast-growing incomes in regions whose economies were centred on resource extraction – a sector that saw a global boom in the 2000s (Jacks, 2019; Stuermer, 2018). This changing relationship between economic complexity and income is an important reminder that income levels are not simply an arithmetic reflection of the capacity to produce but are contingent on a variety of macroeconomic and institutional factors. These factors can shift quite suddenly, even when productive structures do not change.

Panel regression results

We now turn to comparing within-metro changes in ECI to within-metro changes in economic outcomes using panel regressions with robust standard errors, time fixed effects and time-constant dummies for every region (i.e.,

the standard demeaning approach for two-way fixed effects regressions).

The full regression output is presented in Table E2 in Appendix E in the supplemental data online. For the whole period, there is a non-significant, weakly negative relationship. However, when looking at the individual time period after 2007, we find that the partial effects are significantly negative. The effect is fully robust to the inclusion of our control variables (column 4). A one-unit deviation above the mean is associated with a decrease of 4% below the mean of the average region when we employ all our controls. The converse is true for the period before 2007: here, ECI is a significant positive predictor. Figure 3(b) plots the partial effect between ECI and income per capita as we expand our dataset by one year at a time. It shows a similar pattern: the partial effect becomes increasingly positive, peaking in 2006, and then gradually declines until becoming insignificant and negative in 2014. As with the cross-sectional results, the results of panel regressions suggest a changing relationship between ECI and income over time.

DISCUSSION

Using industry employment data, this paper has applied the ideas and methods of economic complexity to the context of US regions between 1998 and 2015. After scrutinizing the underlying assumptions of complexity in a regional context, and updating the methodology where appropriate, we show that economic complexity is well suited to analysing the productive structure of US regions. Specifically, the general nested pattern of the region–industry matrix that has been observed among countries also exists in US regions. This highlights the importance of regions as a focus of economic geography.

We have also shown that the inclusion of local sectors in complexity indicators is valid and strengthens their analytic power. Local industries are an important part of regional economic life and – as we have argued – vary widely in terms of quality and sophistication. Furthermore, we described how measuring industry presence using the LQ alone may underreport the industrial diversity of large and diverse regions, and introduced a composite measure ($CM_{r,i}$) of presence that circumvents many of these problems by taking into account (almost) every industry that is present in a region. In terms of indicators, we argue that differences between ECI and FI are less dramatic than they seem, at least in this context.

Our empirical analysis revealed that the largest US cities have the highest complexity and that metropolitan areas and the Northeast of the United States have a higher complexity than non-metropolitan areas and other US regions, respectively. Traded industries dominate the list of the most complex industries, but many local ones also have above-average complexity values. In general, higher ECI regions have a significantly higher income per capita than lower ECI regions. However, this relationship has become weaker and somewhat ‘U’-shaped over time. Fixed-effects regressions suggest a marked regime change in the within-metro relationship between income and

complexity: it peaked in 2007 and became negative thereafter.

The weakening relationship is not due to changes in economic complexity over time. On the contrary, metropolitan economic complexity has been remarkably stable. Rather, changes in the relationship between these variables are largely a result of changes in income per capita. Many low-complexity regions saw a bigger (relative and absolute) increase in income per capita than high-complexity regions did between 1998 and 2015. In the former group are regions whose economies were centred on resource extraction, suggesting that these findings may result from the commodities boom since the 2000s.

These results emphasize the need for care when studying the relationship between economic complexity – and perhaps productive structure more generally – and other socioeconomic characteristics. It is likely that also regions outside of the United States see a high degree of intertemporal stability in their regional specialization profile. For instance, Gao and Zhou (2018) calculate a Pearson correlation of 0.89 between the ECI of Chinese regions in 2000 and 2015, even though the Chinese economy as a whole grew much faster than the United States during this period. Similarly, Chávez et al. (2017) find a correlation of 0.96 for Mexican regions for the vectors of 1998 and 2013. When the ECI exhibits such stability, attributing changes in other social indicators to changes in ECI may be misleading.

As we have shown, the capabilities approach offers a powerful perspective on local as well as national economic development. Applying it successfully requires careful attention to its underlying assumptions. But done carefully, it has a great deal to teach about the fortunes of local economies.

ACKNOWLEDGEMENTS

The authors thank Jarko Fidrmuc, Andres Gomez-Lievano, Dominik Hartmann, César Hidalgo, Penny Mealy, Emanuele Pugliese, Akira Sasahara, Angelica Sbardella, Marcel Tyrell and the participants of the North American Regional Science 2017 Annual Conference in Vancouver, Canada, for helpful comments and ideas.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

FUNDING

R.M. was supported in this research by the Harvard Multidisciplinary Program in Inequality and Social Policy. B.F. was supported by the Cusanuswerk.

NOTES

1. For an earlier and outdated version of this paper, see Fritz and Manduca (2019).
2. For replication materials, see Fritz and Manduca (2021).

ORCID

Benedikt S. L. Fritz  <http://orcid.org/0000-0003-4048-8998>

Robert A. Manduca  <http://orcid.org/0000-0001-5924-3224>

REFERENCES

- Albeaik, S., Kaltenberg, M., Alsaleh, M., & Hidalgo, C. (2017a). 729 New measures of economic complexity (Addendum to improving the economic complexity index). *Arxiv Working Paper* arXiv: 1708.04107, 1–15.
- Albeaik, S., Kaltenberg, M., Alsaleh, M., & Hidalgo, C. A. (2017b). Improving the economic complexity index. *Arxiv Working Paper* arXiv:1707.05826, 1–21.
- Almeida-Neto, M., Guimarães, P. R., Loyola, R. D., & Ulrich, W. (2008). A consistent metric for nestedness analysis in ecological systems: Reconciling concept and measurement. *Oikos*, 117(8), 1227–1239. <https://doi.org/10.1111/j.0030-1299.2008.16644.x>
- Amos, O. M. (2014). Evidence of increasing regional income variation in the United States: 1969–2006. *Modern Economy*, 5(5), 520–532. <https://doi.org/10.4236/me.2014.55049>
- Andrews, R. B. (1953). Mechanics of the urban economic base: Historical development of the base concept. *Land Economics*, 29(2), 161. <https://doi.org/10.2307/3144408>
- Badger, E., Bui, Q., & Pearce, A. (2016, November 11). The election highlighted a growing rural–urban split. *The New York Times*. <https://www.nytimes.com/2016/11/12/upshot/this-election-highlighted-a-growing-rural-urban-split.html>
- Balland, P.-A., Boschma, R., Crespo, J., & Rigby, D. L. (2019). Smart specialization 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., & Rigby, D. L. (2017). The geography of complex knowledge. *Economic Geography*, 93(1), 1–23. <https://doi.org/10.1080/00130095.2016.1205947>
- Breau, S., Kogler, D. F., & Bolton, K. C. (2014). On the relationship between innovation and wage inequality: New evidence from Canadian cities. *Economic Geography*, 90(4), 351–373. <https://doi.org/10.1111/ecge.12056>
- Buera, F. J., & Kaboski, J. P. (2012). The rise of the service economy. *American Economic Review*, 102(6), 2540–2569. <https://doi.org/10.1257/aer.102.6.2540>
- Bustos, S., Gomez, C., Hausmann, R., & Hidalgo, C. A. (2012). The dynamics of nestedness predicts the evolution of industrial Ecosystems. *PLoS ONE*, 7(11), e49393. <https://doi.org/10.1371/journal.pone.0049393>
- Chávez, J. C., Mosqueda, M. T., & Gómez-Zaldívar, M. (2017). Economic complexity and regional growth performance: Evidence from the Mexican economy. *The Review of Regional Studies*, 47(2), 201–219.
- Cicerone, G., McCann, P., & Venhorst, V. A. (2020). Promoting regional growth and innovation: Relatedness, revealed comparative advantage and the product space. *Journal of Economic Geography*, 20(1), 293–316. <https://doi.org/10.1093/jeg/lbz001>
- Clark, T. N., Lloyd, R., Wong, K. K., & Jain, P. (2002). Amenities drive urban growth. *Journal of Urban Affairs*, 24(5), 493–515. <https://doi.org/10.1111/1467-9906.00134>
- CORFO. (2018). *Informe final de Evaluación Programas Gubernamentales (EPG)*. Technical report. http://www.dipres.cl/597/articles-177360_informe_final.pdf
- Daboín, C., Escobari, M., Hernández, G., & Morales-Arilla, J. (2019). *Economic complexity and technological relatedness: Findings for American cities* (Technical Paper). Brookings Institution. <https://www.brookings.edu/wp-content/uploads/2019/05/Technical-Paper.pdf>
- Davies, B., & Maré, D. C. (2019). *Relatedness, complexity and local growth* (Discussion Paper No. 12223. IZA). <https://www.iza.org/publications/dp/12223/relatedness-complexity-and-local-growth>
- Delgado, M., Porter, M. E., & Stern, S. (2016). Defining clusters of related industries. *Journal of Economic Geography*, 16(1), 1–38. <https://doi.org/10.1093/jeg/lbv017>
- Escobari, M., Seyal, I., Morales-Arilla, J., & Shearer, C. (2019). *Growing cities that work for all*. Brookings Institution.
- Essletzbichler, J. (2015). Relatedness, industrial branching and technological cohesion in US metropolitan areas. *Regional Studies*, 49(5), 752–766. <https://doi.org/10.1080/00343404.2013.806793>
- Farinha, T., Balland, P. A., Morrison, A., & Boschma, R. (2019). What drives the geography of jobs in the US? Unpacking relatedness. *Industry and Innovation*, 26(9), 988–1022. <https://doi.org/10.1080/13662716.2019.1591940>
- Florida, R. (2002). *The rise of the creative class – Revisited*. Perseus.
- Florida, R., & Mellander, C. (2016). The geography of inequality: Difference and determinants of wage and income inequality across US metros. *Regional Studies*, 50(1), 79–92. <https://doi.org/10.1080/00343404.2014.884275>
- Frenken, K., Van Oort, F., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies*, 41(5), 685–697. <https://doi.org/10.1080/00343400601120296>
- Fritz, B., & Manduca, R. (2019). The economic complexity of US metropolitan areas. *Arxiv Working Paper* arXiv:1901.08112.
- Fritz, B., & Manduca, R. (2021). Supplemental materials for paper: The economic complexity of US metropolitan areas. <https://doi.org/10.17605/OSF.IO/Y9RGZ>
- Gabrielli, A., Cristelli, M., Mazzilli, D., Tacchella, A., Zaccaria, A., & Pietronero, L. (2017). Why we like the ECI+ algorithm. *Arxiv Working Paper* arXiv:1708.01161, 1–6.
- Gao, J., & Zhou, T. (2018). Quantifying China's regional economic complexity. *Physica A: Statistical Mechanics and its Applications*, 492, 1591–1603. <https://doi.org/10.1016/j.physa.2017.11.084>
- Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities. *Journal of Political Economy*, 100(6), 1126–1152. <https://doi.org/10.1086/261856>
- Gomez-Lievano, A., & Patterson-Lomba, O. (2019). The drivers of urban economic complexity and their connection to urban economic performance. *Arxiv Working Paper* arXiv:1812.02842v2.
- Hausmann, R., Hidalgo, C., Bustos, S., Coscia, M., Simoes, A., & Yildirim, M. (2013). *The atlas of economic complexity: Mapping paths to prosperity*. MIT Press.
- Heilburn, J. (1981). The urban economic base and economic policy. In *Urban economics and public policy* (pp. 153–169). St. Martin's.
- Hidalgo, C. (2015). *Why information grows: The evolution of order, from atoms to economies*. Basic Books.
- Hidalgo, C., Balland, P.-A., Boschma, R., Delgado, M., Feldman, M. P., Frenken, K., Glaeser, E. L., He, C., Kogler, D. F., Morrison, A., Neffke, F. M. H., Rigby, D. L., Stern, S., Zheng, S., & Zhu, S. (2018). The principle of relatedness. In A. Morales, C. Gershenson, D. Braha, A. Minai, & Y. Bar-Yam (Eds.), *Unifying themes in complex systems IX. ICCS 2018, July 22–27, 2018. Springer Proceedings in Complexity* (pp. 451–457). Springer.
- Hidalgo, C., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences, USA*, 106(26), 10570–10575. <https://doi.org/10.1073/pnas.0900943106>
- Hidalgo, C., 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>

- International Monetary Fund (IMF). (2017). *World economic data-base, technical report*. IMF.
- Isard, W., & Reiner, T. A. (1966). Regional science: Retrospect and prospect. *Papers of the Regional Science Association*, 16(1), 1–16. <https://doi.org/10.1007/BF01888932>
- Jacks, D. S. (2019). From boom to bust: A typology of real commodity prices in the long run. *Cliometrica*, 13(2), 201–220. <https://doi.org/10.1007/s11698-018-0173-5>
- Jacobs, J. (1969). *The economy of cities*. Vintage.
- Kogler, D. F., Rigby, D. L., & Tucker, I. (2013). Mapping knowledge space and technological relatedness in US cities. *European Planning Studies*, 21(9), 1374–1391. <https://doi.org/10.1080/09654313.2012.755832>
- Konrad, A. (2015, October 14). Blue Apron's got big plans for dinner – But so do its hungry rivals. *Forbes*. <https://www.forbes.com/sites/alexkonrad/2015/10/14/inside-blue-apron-and-the-meal-kit-rush/>
- Leonhardt, D. (2018, November 9). America's small-town crisis. *The New York Times*. <https://www.nytimes.com/2018/11/09/opinion/democrats-midterms-rural-ohio-midwest.html>
- Lyubimov, I. L., Lysyuk, M. V., & Gvozdeva, M. A. (2018). Atlas of economic complexity, Russian regional pages. *Voprosy Ekonomiki*, 6.
- Manduca, R. A. (2019). The contribution of national income inequality to regional economic divergence. *Social Forces*, 98(2), 622–648. <https://doi.org/10.1093/sf/soz013>
- Marshall, A. (1890). *Principles of economics*. Macmillan & Co.
- Mealy, P., & Coyle, D. (2019). To them that hath: Economic complexity and local industrial strategy in the UK. *SSRN Electronic Journal*. <http://dx.doi.org/10.2139/ssrn.3491153>
- Mealy, P., Farmer, J. D., & Hausmann, R. (2018). Determining the differences that matter: Development and divergence in US states over 1850–2010. *SSRN Electronic Journal*. <http://dx.doi.org/10.2139/ssrn.3235193>
- Mealy, P., Farmer, J. D., & Teytelboym, A. (2019). Interpreting economic complexity. *Science Advances*, 14(2), 1–9. <https://doi.org/10.1126/sciadv.aau1705>
- Moretti, E. (2010). Local multipliers. *American Economic Review*, 100(2), 373–377. <https://doi.org/10.1257/aer.100.2.373>
- Morrison, G., Buldyrev, S. V., Imbruno, M., Doria Arrieta, O. A., Rungi, A., Riccaboni, M., & Pammolli, F. (2017). On economic complexity and the fitness of nations. *Scientific Reports*, 7(1), 15332. <https://doi.org/10.1038/s41598-017-14603-6>
- Muneepeerakul, R., Lobo, J., Shutters, S. T., Gomez-Lievano, A., & Qubbaj, M. R. (2013). Urban economies and occupation space: Can they get 'there' from 'here'? *PLoS ONE*, 8(9), 1–8. <https://doi.org/10.1371/journal.pone.0073676>
- Neffke, F. (2009). *Productive places: The influence of technological change and relatedness on agglomeration externalities* (Doctoral dissertation). <http://dspace.library.uu.nl/handle/1874/34317>
- Neffke, F. M. H., 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>
- Office of Management and Budget. (2017). *North American Industry Classification System: United States*. Executive Office of the President.
- Operti, F. G., Pugliese, E., Andrade, J. S., Pietronero, L., & Gabrielli, A. (2018). Dynamics in the fitness–income plane: Brazilian states vs world countries. *PLoS ONE*, 13(6), e0197616. <https://doi.org/10.1371/journal.pone.0197616>
- Parr, J. B. (1973). Growth Poles, regional development, and central place theory. *Papers of the Regional Science Association*, 31(1), 173–212. <https://doi.org/10.1007/BF01943249>
- Perroux, F. (1955). Note sur la notion de poles croissance. *Economic Appliquee*, 7(1–2), 307–320.
- Pietronero, L., Cristelli, M., Gabrielli, A., Mazzilli, D., Pugliese, E., Tacchella, A., & Zaccaria, A. (2017). Economic complexity: 'Buttarla in caciara' vs a constructive approach. *Arxiv Working Paper* arXiv:1709.05272v1.
- Porter, M. E. (1998). Clusters and the new economies of competition. *Harvard Business Review*, 77–90. <https://hbr.org/1998/11/clusters-and-the-new-economies-of-competition>
- Porter, M. E. (2003). The economic performance of regions. *Regional Studies*, 37(6–7), 549–578. <https://doi.org/10.1080/0034340032000108688>
- 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>
- Sbardella, A., Pugliese, E., & Pietronero, L. (2017). Economic development and wage inequality: A complex system analysis. *PLoS ONE*, 12(9), 1–26. <https://doi.org/10.1371/journal.pone.0182774>
- 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>
- Storper, M. (1997). *The regional world: Territorial development in a global economy*. Guilford.
- Storper, M. (2011). Why do regions develop and change? The challenge for geography and economics. *Journal of Economic Geography*, 11(2), 333–346. <https://doi.org/10.1093/jeg/lbq033>
- Storper, M. (2018). Separate worlds? Explaining the current wave of regional economic polarization. *Journal of Economic Geography*, 18(2), 247–270. <https://doi.org/10.1093/jeg/lby011>
- Stuermer, M. (2018). 150 years of boom and bust: What drives mineral commodity prices? *Macroeconomic Dynamics*, 22(3), 702–717. <https://doi.org/10.1017/S136510051600050X>
- Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., & Pietronero, L. (2012). A new metrics for countries' fitness and products' complexity. *Scientific Reports*, 2(1). <https://doi.org/10.1038/srep00723>
- The Growth Lab. (2020). *Applied research*. <https://growthlab.cid.harvard.edu/applied-research>
- US Bureau of Economic Analysis. (2017). *Regional economic accounts, technical report*. US Bureau of Economic Analysis.
- US Bureau of Labor Statistics. (2017). *Civilian employment level*. US Bureau of Labor Statistics.
- US Bureau of Labor Statistics. (2020). *Employment by major industry sector*. US Bureau of Labor Statistics. <https://www.bls.gov/emp/tables/employment-by-major-industry-sector.htm>
- US Census Bureau. (2012). *2010 Census summary file – Technical documentation*. US Census Bureau.
- World Bank. (2018). *Economic fitness data*. World Bank.
- Youn, H., Bettencourt, L. M. A., Lobo, J., Strumsky, D., Samaniego, H., & West, G. B. (2016). Scaling and universality in urban economic diversification. *Journal of The Royal Society Interface*, 13(114), 1–7. doi:10.1098/rsif.2015.0937