Economic Complexity and Regional Development: Critical Essays

Margarida Bandeira Morais

Department of Geography and Environment

London School of Economics

A thesis submitted to the Department of Geography and Environment
of the London School of Economics for the Degree of Doctor of Philosophy

London, August 2023

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it). The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without my prior written consent. I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of approximately 48,000 words, excluding all appendices and the bibliography.

Margarida Bandeira Morais

Margarida Bandeira Morais, 08/08/2023

Statement of co-authored work

Chapter 3 of this thesis, “Economic complexity, exports and natural resources in the Gulf Cooperation Council”, is co-authored with Simona Iammarino and Neil Lee. My contribution amounts to at least 60% of the total work. A version of this chapter was published as a working paper in the Geography and Environment Discussion Paper Series (Paper No. 41, May 2023) with the same title.

Acknowledgements

It has been an exciting but incredibly unpredictable four years and I would not have been able to do this PhD without the support of many people.

I want to thank my supervisors, Simona Iammarino and Neil Lee, who offered me relentless support even in the midst of lockdowns, home schooling and personal challenges. Simona has been a huge source of inspiration and encouragement, sharing her knowledge of many years in the field, as well as invaluable advice. Neil was the first to take a bet on me, from accepting me to the MSc Local Economic Development to agreeing to supervise me and having me as a first-time teacher on GY103. His support was indispensable, and he shared the best piece of practical advice to any PhD student – make sure you know exactly what you should be doing each morning.

Many others at LSE contributed to my work and time here. I want to thank Michael Storper for his feedback and the in-depth discussion during my PhD upgrade meeting. I am also grateful to Felipe Carozzi for his advice and time to discuss my work, and to Olmo Silva and other academic staff members and peers for their feedback and suggestions during the Economic Geography work-in-progress seminars. I must also thank Tom Jones, Maxime van Poeteren and Pete Mills for their support over the years, from help with paperwork to their commitment to improving the conditions for PhD students in the department.

I was lucky to discuss my research with people beyond the LSE, both during conferences and in one-to-one meetings. I am particularly grateful to César Hidalgo for taking the time to discuss my research and for sharing data from the Observatory of Economic Complexity, to Tom Broekel for discussions on my work, and also to Teresa Farinha for discussions about economic complexity and Portuguese data. I am also hugely grateful to Athanasia Kalaitzi for reading the third chapter of this thesis, meeting with me and sharing detailed and invaluable feedback on our work.

I must also acknowledge FCT, the Portuguese Foundation for Science and Technology for funding my PhD (under scholarship SFRH/BD/144835/2019), and the Portuguese

Ministry of Labour, Solidarity and Social Security, and more specifically the staff at GEP (Gabinete de Estratégia e Planeamento), for granting me access to secure data.

My journey towards a PhD predates LSE, so I want to thank my supervisors at Utrecht University – Julia Swart was the first person to notice how much I enjoyed research and to suggest a PhD might be for me; Jacob Jordaan first told me about the Geography & Environment department and I am so glad I followed his advice to look into it.

One of the best parts of doing a PhD has been meeting people who are passionate about research and the work they do, from all parts of the world. I am glad for the many lunches, drinks at the George and long conversations. I want to thank Bea and Melissa, who have been there along all steps of the way and with whom I have shared most of the complaining over these years. I am also thankful to Fizzah, Gabi, Luca, Martina, Pedro and Tea who have all, at different times, made my time at LSE better! I am grateful to Lars and Carolin for their guidance and insights at the start of the PhD and teaching, and to Andreas for sharing this template.

I want to thank my friends in London, particularly Eleanor, and my friends and family in Portugal, who keep me grounded and help me distance myself from work, and friends who are spread a bit all over the world, especially Katherine who, against all odds, I have managed to see a few times during the PhD and with whom I reflect the most on life.

Most importantly, I am eternally grateful to those who are always there for me. To my parents, Emília and Carlos, for their encouragement and unconditional support, as well as to my sister, Catarina, a constant source of inspiration, and my soon-to-be brother-in-law, Conor – thank you for the many times we spent together over these four years and for supporting me, during both fun and stressful times. To Aidan, thank you for the last 11 years and for your support and encouragement; thank you also for all the food, the tennis practice and swimming sessions, the fun and exciting times, and the mundane ones, for putting up with tears and panic once in a while, but most of all, for making me believe in myself.

Abstract

Economic complexity indicators, which aim to infer the knowledge and capabilities within countries and regions, have become increasingly prominent in the economic geography literature, as well as on policy recommendations across several different geographical and socioeconomic contexts. Such applications often lack a critical understanding of what economic complexity encompasses and its applicability to different contexts. It is therefore crucial to understand what economic complexity captures, and where and when it should be applied, as well as to have a more grounded understanding of its conceptual shortcomings. This thesis contributes towards closing this gap in the literature and provides a critical evaluation of the economic complexity concept and measure for regional economic development. Structured along four chapters, the thesis provides, firstly, a critical survey of the economic complexity literature and explores the export-based complexity index at the country-level – this is the basis for the next chapters, which address key empirical issues. Moving to the sub-national level and using employment data, the second chapter explores occupation complexity across Portuguese municipalities – here, we analyse the applicability of the economic complexity index to occupation and industry data and explore the policy lessons and implications for regional growth and development. The third chapter focuses on export-based economic complexity and natural resource dependence, where we investigate what the index captures in countries with a very high dependence on oil and whether it is associated with economic growth, with a focus on the Gulf Cooperation Council countries. Lastly, the fourth chapter provides an exploratory empirical analysis, with a focus on the links between economic complexity, education and economic growth across countries, followed by a concluding reflection on the thesis findings. Overall, this thesis provides evidence that economic complexity indicators capture distinctive aspects across different geographical levels and types of data, unlike what some of the existing literature implies. Our findings call for more caution in empirical work, particularly for policy-making purposes – the definition of economic complexity across applications to different types of data and to unique socioeconomic contexts should be carefully considered, and any economic development policy based on this top-down indicator crucially requires consideration of local conditions.

Table of Contents

List of Figures

List of Tables

Introduction

For many decades, scholars have tried to understand and explain economic development across countries and regions, pointing to the importance of division of labour and trade, technological change and innovation, human capital and knowledge, economic and industrial structure, among other factors. Over the years, with improvements in computational capacity and more sophisticated tools, new indicators have emerged aimed at helping us better understand the process of economic growth and development over time.

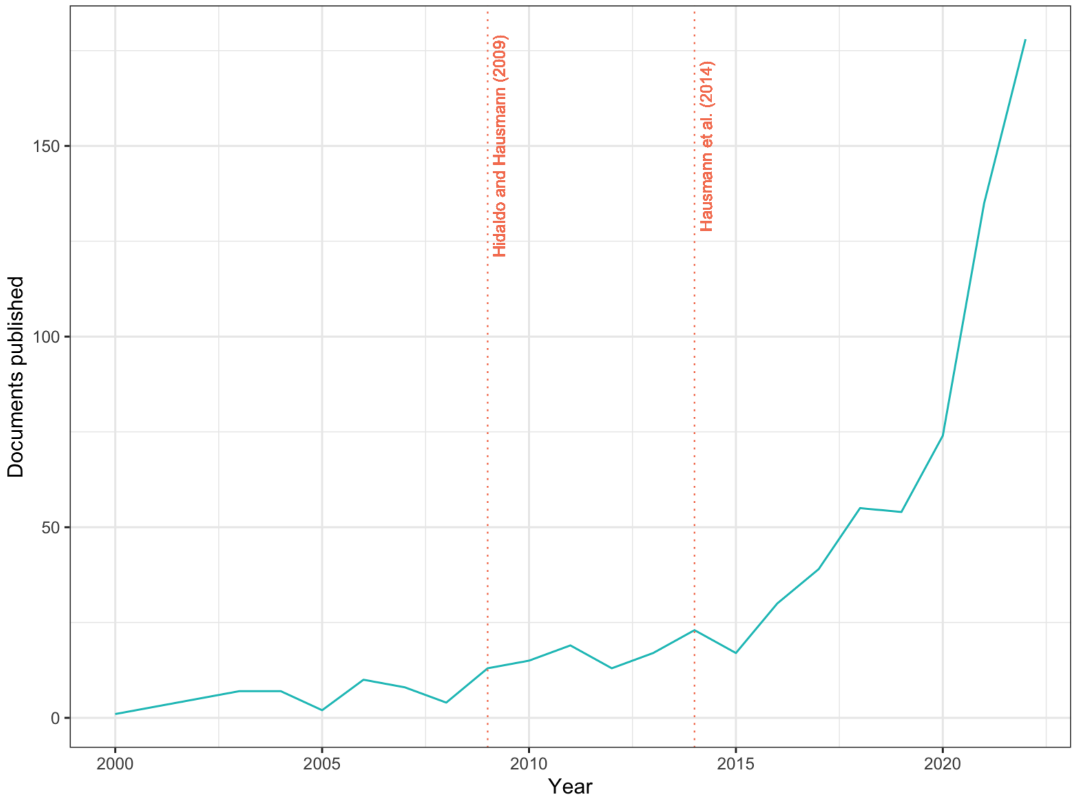
The Economic Complexity Index (ECI) was first introduced over a decade ago by Hidalgo and Hausmann (2009), as a new way of measuring the knowledge available within a country, inferred from the products it exports competitively. Since its introduction, economic complexity has become increasingly prominent in research investigating the development of countries and regions, and has been applied across several geographic and development contexts. Moreover, the method has been applied to different types of data well beyond exports, including patents and employment.

To provide evidence of this, Figure 1 shows the evolution of work published containing the term "economic complexity" in the title, abstract or keywords from 2000 to 2022, showing an exponential increase in recent years.[1] The keywords that most come up in these papers are "economic development" and "economic growth", followed by words related to the environment such as "carbon dioxide", "sustainable development" and "renewable energy", particularly as researchers moved towards looking at how economic complexity is linked with pollution and other environmental outcomes in recent contributions.

Despite this increased interest, many empirical applications lack a critical understanding of what economic complexity measures and its applicability to different contexts. It is therefore crucial to assess what economic complexity captures, and where and when it should be applied, as well as to have a more grounded understanding of its conceptual shortcomings. Structured along four main chapters, this thesis contributes towards closing this gap in the literature and provides a critical evaluation of the economic complexity concept and measure, with a focus on regional economic development.

[1]Own elaboration using data from Scopus.

Figure 1: Publications with “economic complexity” in title, abstract or keywords over time



Chapter overview

There are four closely related chapters in this thesis. They each depart from the broad motivation presented here and address different aspects about the Economic Complexity Index that are not fully understood. The chapters investigate, firstly, the theoretical foundations at national and sub-national levels of research and the main critical drawbacks of the economic complexity literature; secondly, what the ECI captures when applied to a single country at the sub-national level, and whether occupation data can address some of the drawbacks of other types of data used in complexity measures; thirdly, whether the ECI works in countries that are highly dependent on natural resources, in particular oil; fourthly, whether economic complexity captures something beyond education levels when it comes to explaining economic growth across countries. To do so, Chapters 1, 3 and 4 rely on the same country-level dataset, focusing on complexity measures built from world export data, while Chapter 2 relies on employment data, focusing on the sub-national level for the case of Portugal. These chapters are followed by a final conclusion reflecting on the evidence provided in the thesis and highlighting key areas for future research.

Chapter 1 – Economic complexity: A critical survey

The first chapter provides a critical survey of the economic complexity literature and explores the export-based complexity index for countries. Despite increasing popularity, research on a theoretical level on what the economic complexity measure captures, and how or where it should be applied is still lacking. This chapter takes a step towards filling this gap. First, we situate economic complexity within the theoretical literature, both at the national and regional levels. Second, we introduce the methodology and examine the empirical literature to date. Third, we conduct exploratory data analysis, focusing on economic complexity across countries, calculated from export data from the Observatory of Economic Complexity, along with several additional variables to understand what this measure is capturing. Finally, we identify the current drawbacks in economic complexity, as well as its contribution to the wider literature and our understanding of regional development. We argue that, while the notion of economic complexity provides new and advanced methods, it still suffers from important drawbacks in three broad areas, relating to conceptualisation and theory, data and methodology, and empirical research. This chapter also provides some grounding to the following chapters, which address key empirical issues.

Chapter 2 – Economic complexity & employment in Portuguese municipalities

Moving to the sub-national level and using employment data, the second chapter explores occupation complexity across Portuguese municipalities – here, we analyse the applicability of the economic complexity index to occupation and industry data and explore the policy lessons and implications for regional growth and development. While the ECI has been applied to several types of data (including exports, patents and industries), we argue that the use of occupations is likely to help overcome some of the drawbacks of current applications at the sub-national level. However, more work is needed in conceptualising and interpreting occupation-based ECI measures. To move in this direction, we apply the methodology to a network connecting municipalities to occupations, derive occupational diversity and ubiquity, and get a ranking and level of complexity. Focusing on 308 Portuguese municipalities, from 1985 to 2019, we investigate how complexity evolves over time and look at the association between complexity measures and regional outcomes such as employment growth, through panel data analysis. Following this, we explore how occupation-based complexity relates to the task-based approach to labour markets, as well as to complexity measures based on employment industries. The empirical analysis shows a negative association between both economic complexity measures and employment growth, though any association disappears once we control for initial employment levels, suggesting this link may be simply reflecting convergence mechanisms. While discussing these results, we suggest that occupation and industry-based ECI measures in this context may be reflecting specialisation patterns too strongly, without necessarily reflecting the underlying ‘capabilities’ within a municipality.

Chapter 3 – Economic complexity, exports & natural resources in the GCC

The third chapter focuses on export-based economic complexity and natural resource dependence, and we investigate what the index captures in countries with a very high dependence on oil and whether it is associated with economic growth, with a focus on the Gulf Cooperation Council (GCC). The applicability of the ECI method and concept across different contexts has remained unquestioned in existing literature. In particular, we argue that the unique characteristics of natural resource dependent countries are largely disregarded. Using the ECI for 179 countries from 1995 to 2019, we focus on the case of the Gulf Cooperation Council countries, generally considered high-income economies heavily reliant on oil and natural gas exports. While we find that the link between the ECI and subsequent economic growth observed across countries holds for the GCC and other oil-dependent countries, our analysis exposes important ways in which the ECI is affected by the high dependence on oil and its price volatility. Contrary to existing literature, we found no association between economic complexity and economic growth within countries over time. Our analysis calls for more caution when relying on economic complexity measures for policy-making, and highlights the need for additional and more granular analysis of different contexts, particularly those heavily reliant on natural resources.

Chapter 4 – Economic complexity, education & economic growth

Lastly, the fourth chapter is more exploratory than the previous ones, providing a concluding empirical analysis focused on the links between economic complexity, education and economic growth across countries. Education has long been recognised as a key driver of economic growth and development across countries. At the same time, the proponents of the ECI argue that economic complexity can go beyond long-standing education variables in explaining future economic growth, as it captures productive knowledge within economies, rather than how much of the same knowledge individuals acquired in school (Hausmann et al., 2014b, p. 36). To test this hypothesis, this chapter investigates whether initial economic complexity levels can explain future economic growth better than education variables, such as years of education and schooling attainment. Our preliminary results show that education measures have a stronger predictive power than economic complexity in explaining both short- and long-term economic growth across countries. Importantly, this suggests that we should not overlook simpler variables such as education or other key determinants of economic development, particularly for policy purposes.

Overall, this thesis provides evidence that economic complexity indicators capture distinctive aspects across different geographical levels and types of data, unlike what some of the existing literature implies. Our findings call for more caution in ECI applications, particularly for policy-making purposes – the definition of the ECI across applications to different types of data and to unique socioeconomic contexts should be carefully considered, and any economic development policy based on this top-down indicator crucially requires consideration of local conditions.

Chapter 1

Economic complexity: A critical survey

1.1 Introduction

In the last decade, a group of scholars introduced the Economic Complexity Index. According to this concept and methodology, the knowledge a country has, inferred from the products it makes, can predict its future economic growth (Hidalgo & Hausmann, 2009). As described by Hausmann et al. (2014b), countries with higher economic complexity levels can bring larger volumes of knowledge together, across broader networks of people to generate a larger number of more knowledge-intensive goods. In contrast, countries with lower economic complexity levels have more limited knowledge bases and produce fewer and simpler goods.

Since its introduction, economic complexity has become increasingly prominent in research investigating the development of countries and regions, and has been applied to several contexts, from broad cross-country analyses to sub-national EU regions' smart specialisation strategies. This is sometimes done with limited consideration of socioeconomic or development contexts.

Despite increased popularity among social scientists, several aspects related to the ECI have yet to be fully grasped. Specifically, we need to understand whether it is in fact a novel approach to understanding economic growth and, if so, whether the methods deliver what we want to measure. Critiques so far consist of technical papers (e.g., Mealy et al., 2018b; Tacchella et al., 2012), focused exclusively on the methodology, while assessment of the usefulness and legitimacy of this measure for social sciences is still lacking.

To take a step in this direction, this chapter does three things. First, it situates economic complexity alongside theories of economic development, both in national and regional contexts. Here, we question the originality of economic complexity as a novel explanation of economic growth at both geographical levels. Second, it examines the empirical literature to date, at cross-country, national and sub-national levels. Third, it presents an exploratory analysis using data from the Observatory of Economic Complexity. With the two latter points, the chapter challenges the widespread and uncritical use of this measure to understand economic growth and capabilities.

We argue that, while the notion of economic complexity provides a new and advanced method, which allows us to identify countries' and regions' competences through the lens of the activities they are able to perform competitively, it still suffers from important drawbacks in three broad areas, relating to theory, data and methodology, and empirical applications.

This chapter makes two contributions. On a theoretical level, it brings much needed scrutiny to an increasingly popular measure. On a practical level, the chapter identifies the current drawbacks to set research priorities towards understanding whether and when economic complexity is useful to understand economic growth and development.

The rest of the chapter is organised as follows. Section 1.2 introduces the intuitive idea behind economic complexity and reviews the relevant existing theory, at both national and regional levels, to place economic complexity in context. Section 1.3 outlines the methodology, followed by empirical applications to date in Section 1.4. Section 1.5 presents the exploratory data analysis, using country-level data. Section 1.6 provides a discussion on the current challenges and the contribution of economic complexity to the literature, followed by a conclusion and thesis roadmap in Section 1.7.

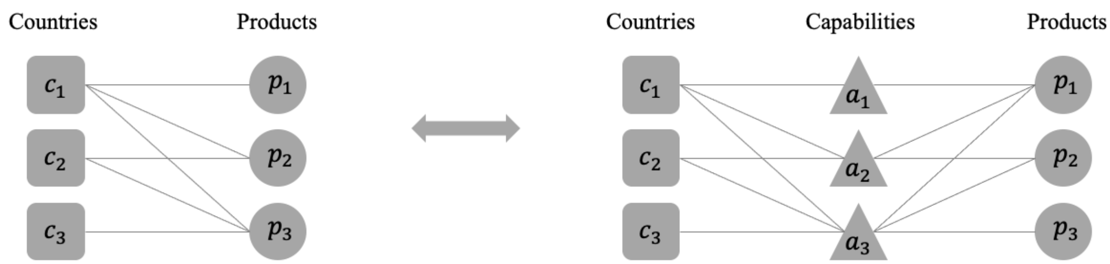
1.2 Understanding economic complexity

1.2.1 What is economic complexity?

The concept of economic complexity and the associated ECI were first introduced by Hidalgo and Hausmann (2009) as a tool for measuring the knowledge in a society. They suggested that it is possible to infer the amount of 'capabilities' or 'knowledge' in a country by looking at the products it makes, through a network-based method borrowed from the natural sciences. While methodologically advanced, this approach, and the terms used by its proponents to describe it, remain somewhat abstract. This section introduces the intuition and ideas behind the ECI, as portrayed in the literature, to guide the theoretical grounding.

The authors start from the two-mode network connecting countries to the products they export competitively, and assume that this two-mode network is part of a larger, unobserved, three-mode network connecting countries to the capabilities they possess and products to the capabilities they require, as illustrated in Figure 1.1. The intuition is that if one country is able to make a product, it must have the capabilities that are needed to produce it. Even though we cannot identify those capabilities (i.e., the network on the right panel), we can see which countries can make what products (i.e., the network on the left panel).

Figure 1.1: Schematic representation of the two-mode and three-mode networks



Source: Hidalgo and Hausmann (2009), own drawing.

In order to define capabilities (used interchangeably with knowledge), Hausmann et al. (2014b) distinguish between codified knowledge, which they argue can be more easily transmitted, and tacit knowledge, which cannot be easily communicated or transferred. The authors argue that tacit knowledge and countries' ability to combine and share it are at the heart of differences in economic development, as it is hard for other countries to replicate – an idea that is far from new. Furthermore, in order to perform certain activities, individuals need to hold pieces of knowledge that are coherent among each other. These "modularised chunks of embedded knowledge" are what the authors coin 'capabilities', and can be at the individual level, grouped within an organisation, or across networks of organisations (Hausmann et al., 2014b, p.16).

Intuitively, if a country is able to export a specific pharmaceutical product competitively, it is assumed that it can combine several pieces, or 'capabilities', including for example, chemical engineering, laboratories and legal capacity that allow them to do so. These capabilities all need to be present, they complement each other, and they are likely to involve both codified and tacit knowledge components.

As described by Hausmann et al. (2014b), the economic complexity of countries is an outcome-based measure, expressed in terms of their productive structures. From the two-mode network, two variables are calculated – *diversity* and *ubiquity*. Diversity refers to the number of products that a country exports competitively, which the authors argue can be a crude proxy for the variety of capabilities available in a country (Hausmann et al., 2014b). Intuitively, this means that if a country is able to make a larger number of products (i.e., is more diversified), it has a higher number of capabilities. Ubiquity refers to the number of countries that are able to export a product competitively, and can provide an indication of the variety of capabilities required by a product (Hausmann et al., 2014b). In practical terms, if a product is exported by a large amount of countries (i.e., it is more ubiquitous), it should require fewer and more widely available capabilities.

Since each of these variables alone is an imperfect measure of complexity, diversity and ubiquity need to be combined. Thus, the ECI is constructed by using diversity to adjust the information carried by ubiquity, and ubiquity to adjust the information carried by diversity, repeating this iteratively (Hausmann et al., 2014b). As an illustration of why these two variables are corrected using each other, we can think about the example of diamonds – they are produced by few countries (i.e., have low ubiquity) because they are rare, thus it is a matter of geography or luck. The intuition behind this iterative mechanism is that if diamonds were complex, countries producing them should also be able to make other complex products; however, this is not observed (Hausmann et al., 2014b).

Overall, countries with higher economic complexity levels are those that bring many pieces of knowledge together, across large networks of people and generate a diverse mix of knowledge-intensive products. In contrast, countries with low economic complexity levels tend to hold fewer pieces of knowledge, have more limited networks of people and make fewer and simpler goods (Hausmann et al., 2014b).

The implications of having higher or lower economic complexity levels are two-fold. First, Hidalgo and Hausmann (2009) showed that the ECI correlates with income per capita across countries and can predict future economic growth. Second, due to path dependence, current economic complexity levels shape the possibilities and incentives for future development of capabilities. As described by Hausmann and Hidalgo (2010), the returns to accumulation of new capabilities is higher for high complexity countries, that already possess a large number of capabilities, than for low complexity countries. This occurs because countries with lower complexity have fewer capabilities, are able to produce fewer goods, and will have little incentives to develop additional capabilities, since these are unlikely to complement, or derive from, existing ones or to be useful for the production of new goods (Hausmann & Hidalgo, 2010).

1.2.2 Theoretical grounding I: National level

The ideas behind the economic complexity index, as originally introduced, relate first and foremost to country-level theories of economic development. The proponents draw on Adam Smith's division of labour and specialisation and the idea that "development is associated with an increase in the number of individual activities and with the complexity that emerges from the interactions between them" (Hidalgo & Hausmann, 2009, p.1). Furthermore, the authors relate the ECI to trade theory and comparative advantage models, highlighting the importance of the structure of the product space and the path dependence it creates (Hausmann & Klinger, 2007). More recently, Balland et al. (2022) also linked economic complexity with knowledge and technology, in particular the importance of new combinations of existing ideas in developing inventions, previously emphasised by Weitzman (1998) and Fleming and Sorenson (2001). Nevertheless, the idea at the heart of economic complexity is far from new and relates to existing theory well beyond the one its proponents drew upon. As a result, this sub-section focuses on key theoretical contributions at the national level and how they relate to economic complexity.

Classical and dynamic legacies

As described by Hidalgo and Hausmann (2009), Adam Smith emphasised the role of the division of labour, increasing productivity and demand. Smith argued that economic processes are shaped by continued development that stems from the introduction of new technologies, and emphasised the key role of the economics of knowledge in this process (Smith, 1776; Antonelli, 2008a).

Antonelli (2008a) described Marshall as a key contributor to the 'dynamic legacies', following Adam Smith. In particular, Marshall elaborated on Smith's legacy to further introduce the complexity of structural change, characterised by the interaction between specialisation and technological change and leading to a growing heterogeneity of firms in a context of increased variety and complementarity (Marshall, 1890; Antonelli, 2008a). Moreover, Marshall highlighted the collective character of technological knowledge, which makes co-location crucial, and where various agents contribute complementary bits of knowledge. Finally, he introduced the idea of knowledge externalities, which he argued play a central role for firms, and are a key input for the generation of new knowledge – thus, identifying the knowledge generated by firms both as an output and as an input in economic processes (Marshall, 1890; Antonelli, 2008a).

Consistently with Marshall's interpretation of competitive processes, which emphasised the importance of variety and selection, Schumpeter made crucial seminal contributions to the economics of innovation and technological change, further adding to the dynamic legacies (Antonelli, 2008a). Schumpeter (1942) articulated the role of large corporations as the engines for the introduction of innovations, and emphasised that technological change is not only endogenous to economic processes, but an intrinsic element of capitalism. Moreover, writing nearly eight decades ago, Schumpeter (1942, p. 83) conveyed the importance of goods and production for growth in capitalist societies:

> *"The fundamental impulse that sets and keeps the capitalist engine in motion comes from the new consumers' goods, the new methods of production or transportation, the new markets, the new forms of industrial organization that capitalist enterprise creates."*

Thus, the idea that the products a country makes, as well as the underlying ability they possess in terms of production methods, industrial organisation and capacity to innovate determine its economic growth fate goes back to Schumpeter and others.

Further highlighted in this literature, and relevant to the mechanisms behind economic complexity, is the role of the firm and the micro-foundations of innovation, which were missing from neoclassical theory, as Antonelli (2009) argues. Rather than simply adjusting prices and quantities, as was traditionally modelled in neoclassical growth models (e.g., Solow, 1957), Schumpeter highlighted the key role that firms play in innovation as they react creatively to changing or unexpected conditions, thus making innovation fully endogenous to the economic system (Schumpeter, 1947; Antonelli, 2008a). Nelson and Winter (1982) further highlighted the role of innovation as a crucial source of profits and growth for firms, usually achieved by adapting the resources and knowledge within the firm, while also drawing from complementary skills and resources of other firms, often located in proximity.

The literature evolved towards understanding the links between innovation, firm size and market structure, and drawing a distinction between Schumpeterian Mark I pattern of innovation, characterised by 'creative destruction', in which firm entry and entrepreneurship play and major role, and Schumpeterian Mark II of 'creative accumulation', implying a more important role of large established firms and barriers to entry (Breschi et al., 2000). This is closely linked with the notion of technological regimes, dating back to Nelson and Winter (1982) and Winter (1984), describing the technological environment in which firms operate. More specifically, a technological regime can be thought of as a combination of crucial properties of technologies, such as the opportunity and appropriability conditions, the degree of cumulativeness of technological knowledge, and the nature of the knowledge base, especially in terms of transmission (Malerba & Orsenigo, 1993; Audretsch, 1997; Malerba & Orsenigo, 1997).

This importance of the wider environment is also emphasised in the localised technological change literature that developed following the legacies of Smith, Marshall and Schumpeter among others. In addition to the importance of innovation as one of the forms of 'reaction' by firms facing unexpected events, which can lead to successful technological change, the extent to which it happens depends on whether underlying conditions allow for it (Nelson & Winter, 1973, 1982). This strand of literature also highlights the importance of learning, in which firms are viewed as key agents, and which is inherently localised, not only in terms of geography, but also technical space and factor intensity (Atkinson & Stiglitz, 1969). Moreover, rather than seeing one single firm as commanding all knowledge, technological knowledge became increasingly viewed as dispersed and fragmented into various complementary, yet specific, applications and contexts (Hayek, 1937, 1945; Antonelli, 2008b).

Economics of innovation, evolutionary theory and complexity

These legacies and contributions related to the economics of knowledge and innovation are very closely interlinked with evolutionary economics, as are many of the phenomena described, such as the importance of firms, technological change and the importance of variety and complementarity (Metcalfe, 1994). Moreover, in turn, there are important overlaps in conceptual ideas between the evolutionary economics approach and economic complexity. For example, the ideas of path dependence and cumulative technology in evolutionary thinking – i.e., that today's technological advances build from and improve upon the technology that was available at the start of the period, and that what can be done in the future is dependent on what is available and built in the present moment (David, 1985; Dosi & Nelson, 1994) – are fundamental to the way economic complexity is conceived.

As described by Hidalgo (2021), the concept of economic complexity derives from complex systems thinking and relates, in particular, to Weaver's (1948) early idea of 'organised complexity' focused on "vast systems for which the identity of the elements involved and their patterns of interaction could not be ignored" (Hidalgo, 2021, p.2). With the emergence of more advanced data and methods, matrices or networks started being used to capture this understanding and to measure the presence of multiple factors simultaneously by using dimensionality reduction techniques that preserve the identity and inter-dependencies of the factors involved. Despite the name, the ECI is not directly related to the field of complexity economics, which is concerned with viewing the entire economy as a 'complex system' and modelling several different aspects (e.g., people, firms and markets) simultaneously. Nevertheless, these research areas have some commonalities in the sense that they both derive from natural science methods, and can be closely linked with earlier work on the economics of innovation, which is attributed all the way back to evolutionary economic thinking (Antonelli, 2009; Foster & Metcalfe, 2009).

In fact, the ECI network-based approach is not the first application of a natural science method to the broad field of innovation and technology. Marshall can be considered the first economist to look to biology as a source of inspiration for a dynamic approach to economics (Foster, 1993; Antonelli, 2008a). Moreover, Fleming and Sorenson (2001) draw on evolutionary biology and on complex adaptive systems theory and methods to develop a framework of technological evolution as a recombination of new and existing technology components to understand how the likelihood of useful inventions can be maximised, while relying on patent data. Nevertheless, it remains unclear whether drawing on technical measures for social science analyses is always a good and reliable practice.

Technological and other capabilities

A third important broad avenue of research that can provide important insights to the concept of economic complexity relates to capabilities. The rationale behind 'capabilities' put forward by Hausmann et al. (2014b) is not unprecedented, nor are the theoretical ideas that accompany it, as they relate closely to those of several papers from the 1990s. In fact, Bell and Pavitt's (1995) notion of technology consists of complex 'bundles' of information, involving both codified and tacit knowledge, as well as physical capital – a description that does not fall far from the one seen in the economic complexity literature.

Previous authors, however, explored and developed the concept much further, providing relevant insights. Technological capabilities were defined by Kim (1997, p.4) as "the ability to make effective use of technological knowledge in efforts to assimilate, use, adapt and change existing technologies [...] to create new technologies and to develop new products and processes." As noted by Fagerberg and Srholec (2017), this description is close to the concept of 'absorptive capacity' introduced by Cohen and Levinthal (1990), who highlighted the importance of firms' capabilities to access, absorb and use knowledge for technological change and economic development.

This literature makes a crucial distinction regarding types of technological capabilities, namely between production capacity – the ability to operate technology and produce goods and services – and technological or innovation capability – the ability to create configurations of products and technologies and to implement improvements, necessary for technological change (Lall, 1992; Bell & Pavitt, 1995; Bell, 2009), as well as to reap the profits from those technologies (von Tunzelmann & Wang, 2007). This distinction between production and innovation capabilities is consistent with theories that highlight differences between developed and developing countries in their access to different amounts of information and the ability to accumulate further information through time, as is the case in the model developed by Acemoglu and Zilibotti (1999), leading to different economic relations and institutions. This distinction is not made in economic complexity but is particularly important when considering countries at different stages of development. This raises the question of whether exports can accurately reflect underlying capabilities across countries. A country that is able to export a product competitively today does not necessarily possess all the capabilities it needs to be able to generate technological change and move to other, more complex, products in the near future – particularly in today's globalisation context of intricate production networks and value chains. Thus, the opportunities to develop new products and to increase economic complexity may differ across countries that are seemingly the same in terms of economic complexity levels today.

Lall (1992) suggests the development of capabilities rests on the outcome of complex interactions between incentive structures, human capital, technological effort and institutional characteristics, and that governments play an indispensable role. Moreover, Lall (1992) argues that national capabilities do not consist simply of the sum of thousands of individual firm-level capabilities, but rather there are important externalities and linkages between these, leading to different national outcomes, and it ultimately means that countries differ in their ability to employ and further improve technology, which is mirrored in their productivity, growth or trade performance.

Beyond technological and production capabilities, Abramovitz (1986, 1994a, 1994b) suggested that differences in countries' economic success can also be explained by differing levels of 'social capabilities', a term that encompasses aspects such as general education levels, organisational competence, stable and effective government, functioning financial institutions and markets with capacity to mobilise capital, and the spread of trust among people (Abramovitz, 1986, 1994a, 1994b; Fagerberg & Srholec, 2008).

National innovation systems and institutional settings

The existence of externalities and linkages as well as the importance of the wider societal and institutional context are reflected on the literature on national innovation systems (Freeman, 1987; Freeman & Lundvall, 1988; Lundvall, 1992; Nelson, 1993; Edquist, 1997).

While definitions vary, national innovation systems refer to the institutional settings within which the creation, storage and transfer of knowledge and technological opportunities occur (see, for example, Fagerberg and Srholec, 2008).

Ultimately, there may be considerable variation in the gains derived from producing certain goods across countries, particularly for those at different development levels. Hausmann and Rodrik (2003) highlight the importance of 'learning what one is good at producing', which was overlooked in neoclassical theory, where production functions were assumed to be common knowledge. This is not an accurate assumption in developing countries as, even when there is full exposure to foreign technology and production techniques, there is the need to transfer it to new economic and institutional environments, which is an uncertain process. Thus, the importance of institutions for the long term development of countries has been widely recognised and cannot be overlooked (North, 1990, 1991; Rodrik et al., 2004).

Trade and international competitiveness

Economic complexity can also be approached in light of the literature on international competitiveness and trade performance, dating back to the 1980s. This is a wide-ranging literature that involves different levels of analysis, including countries, industries, firms and products. Technology gap theories of trade emphasise the importance of technological knowledge in explaining international trade patterns (Lundberg, 1988). Furthermore, Fagerberg (1988) suggests that beyond the ability to compete in technology and price, countries also compete in their capacity to deliver, and highlights the fundamental role played by investment, creation of productive capacity and technological diffusion.

As an outcome-based measure, economic complexity is likely to capture Fagerberg's ideas, as it is assessing not simply whether countries have the technology available or the most attractive prices, but rather whether they are able to export competitively a vast number of sophisticated products. The nature of international competitiveness as a relative term, given the focus on how well a country does relative to other countries (Fagerberg, 1996), is also inherent to the ECI, which is based on revealed comparative advantage in specific products across countries.

Indicators of technological intensity are often used as competitiveness proxies, including Research and Development (R&D) expenditure (Pavitt, 1984), employment of scientists and engineers, and the number of patents (Patel & Pavitt, 1987). As suggested by Buckley et al. (1988), these should be complemented by an indicator of the outcome of the technology process, as it is the outcomes, rather than the inputs or spending that matter for firms and industries. Furthermore, Buckley et al. (1988, p. 196) argue:

> *"Rather more sophisticated are measures which take into account the (changing) composition of exports and imports, more specifically the concern that market share in 'sophisticated' products is declining and unsophisticated ones is increasing. The argument then is that sophisticated products are technologically intensive and that the loss of technology intensive market share has detrimental social implications including declining employment and increasingly unskilled job provision."*

This idea that some products are more sophisticated than others, requiring higher technology intensity – which is likely not achievable across all countries – and that the mix of products results in different outcomes for economic growth, as well as other qualitative aspects of development, such as employment levels and quality, is central to economic complexity.

Moreover, the interest in the link between income per capita and the mix of exports goes further back to the idea of ladders of development by Chenery (1960), and to Leamer (1985) a few decades later. More recently, there were in fact attempts to measure product sophistication through exports. Lall et al. (2006) developed a classification of products based on the characteristics of the exporter rather than from parent industry data on factor content or R&D activity, as had been done previously. They called it export 'sophistication' – the higher the average income of its exporter, the more sophisticated a product is, with the rationale that products exported by richer countries have characteristics that allow high wage producers to compete globally, at least in the absence of trade interventions (Lall et al., 2006). Since then, Sutton and Trefler (2016) also developed a theoretical model to link countries' income with changes in their export mix and, turning to the empirical side, they find that the range of income levels of significant exporters of most products is very wide (this is because in equilibrium, there are a range of producer qualities and country income levels that are viable in a given industry; as quality rises, the country moves into the production of higher-ranked products and its equilibrium wage and income per capita rises). As will be further discussed, this is an important point that methodological critiques of the ECI have raised.

Lastly, a recent theoretical model by Dam and Frenken (2022) links capabilities, variety and complexity with economic development stages – they argue that countries tend to abandon the least complex activities as they develop (for instance, due to increases in the minimum wage) and that, consistently with a phenomenon known as 'the hump', variety of products first increases and then decreases with economic development. The authors argue that this is inconsistent with economic complexity analyses, which would predict ever-increasing variety as more capabilities are acquired over time. Thus, their model imposes a constraint on the range of the complexity of products that a country produces whereby at a certain level of development countries will drop their least complex products, implying that increases in the economic complexity of a country will accelerate as a function of new capabilities developed (Dam & Frenken, 2022). While this provides a first step in thinking more clearly about the theoretical links between economic complexity at different development levels, questions remain about the channels through which economic complexity leads to economic growth, and ultimately development, across countries.

1.2.3 Theoretical grounding II: Economic complexity and regions

While the original economic complexity literature followed a top-down view of development, looking at countries and pooling together all development contexts, this concept is relevant for and has been applied to regional contexts. This subsection explores economic geography and regional development theory, and the ways in which economic complexity fits in this literature.

Regional worlds and evolutionary economic geography

Regions are recognised as important geographical units of analysis, particularly following the work by Storper (1993, 1997). They are characterised by place-specific features that evolve over time, including human capital, firms, institutions and innovation capacity (Iammarino et al., 2019). These lead to different outcomes, not just in terms of income per capita and wages, but also with regards to qualitative aspects of development, such as employment opportunities and public service provision. Moreover, regional economic development involves enduring systems and gradual change (Maskell & Malmberg, 1999), thus altering a region's fate is not a straightforward process. Since the 1970s, regional inequality increased in both developed and developing countries (Iammarino et al., 2019), a trend that propelled researchers to try to understand why some regions within a country succeed while others lag behind.

Many of the concepts and ideas that researchers turned to in order to explain regional development mirrored those introduced at the national level. The field of evolutionary economic geography (EEG) emerged, first adapted from evolutionary economics to the regional world by Boschma and Lambooy (1999), and focused on the role of knowledge and innovation in regional development paths. As defined by Boschma and Frenken (2011, p. 295), EEG "explains the spatial evolution of firms, industries, networks, cities and regions from elementary processes of the entry, growth, decline and exit of firms, and their locational behaviour." Concepts such as path dependence, increasing returns and variety became increasingly seen as relevant and useful in understanding the processes of localised collective learning, problems of negative lock-in and the development of agglomeration economies and spatial emergence of industries experienced in regions (Boschma & Lambooy, 1999; Lambooy & Boschma, 2001; Martin & Sunley, 2006; Boschma et al., 2017).

The idea that geography and location are key factors in explaining innovation and technological change became a key insight and stylised fact from the 1990s onward (Jaffe et al., 1993; Audretsch, 1997; Audretsch & Feldman, 2004). Because of the tacit nature of much of the knowledge embedded in regions – which does not travel easily and is rooted in place – geographical location is crucial and continues to play a critical role in the development of technology, in particular more complex and valuable one, and tacit knowledge is a crucial source of competitive advantage for firms and regions, and in itself leads to the concentration of activity (Kogut & Zander, 1992; Lawson & Lorenz, 1999; Gertler, 2003; Balland et al., 2019). Collective learning and innovation are seen as core regional assets underpinning geographical concentration, specialisation and technological performance (Storper, 1993, 1997), and the capacity of regions for continuous learning and product innovation is at the heart of sustained competitive advantages (Lawson & Lorenz, 1999; Morgan, 2007).

It is recognised that technology moves along different trajectories across space and is conditioned by local processes of search and learning, ultimately leading to differing technological change trajectories across different regions (Rigby & Essletzbichler, 1997). Firms' processes of building competitiveness is further influenced by localised capabilities, which are regional characteristics, such as infrastructure, access to natural resources, specific institutional endowment and the knowledge and skills available (Maskell & Malmberg, 1999). This results in a diverging pattern, with some prosperous regions that succeed because they possess the right knowledge base, employ organisational methods that efficiently translate that knowledge into marketable products, while other regions lag behind due to less competitive technologies and inferior organisational methods (Rigby & Essletzbichler, 1997; Döring & Schnellenbach, 2006). As argued by Iammarino et al. (2019), a hierarchy of regions is becoming increasingly evident in terms of knowledge creation and non-routine activities.

Agglomeration economies

From the external economies that Marshall (1890) introduced, the term agglomeration economies emerged (Weber, 1929), along with the idea that agglomeration generates geographically bounded positive externalities and drives economic growth in large cities and regions (Glaeser et al., 1992; Glaeser, 1999). As described by Frenken et al. (2007), there are four main sources of agglomeration economies, namely increasing returns to scale (from serving large markets), localisation or 'Marshallian' economies (external economies derived from other firms within the same sector), urbanisation economies (derived from city size and density per se), and Jacobs externalities (stemming from the presence of a variety of sectors).

Agglomeration externalities are more prevalent in industries where new knowledge plays a greater role, and there is a higher propensity of innovative, technology- and knowledge-intensive activities to cluster geographically (Audretsch & Feldman, 1996; V. Henderson, 1999). In line with this, we expect knowledge externalities to be more prevalent in industries or products associated with a higher complexity level and for these industries to be more geographically concentrated, a hypothesis that was confirmed by Balland et al. (2020). Moreover, there are regional discrepancies in the efficiency of using knowledge spillovers (Döring & Schnellenbach, 2006), thus we expect regions with higher economic complexity levels to be more well-equipped to maximise the benefits derived from these externalities.

Local capabilities and regional diversification

Importantly, researchers are interested in how the process of diversification unfolds in regions, and there are several studies showing that existing local capabilities condition the range of new activities that regions are likely or able to develop (Neffke et al., 2011; Boschma et al., 2014; Boschma, 2017). The concepts of related variety and relatedness are considered important drivers of diversification and other regional outcomes, such as employment and labour productivity growth (Frenken et al., 2007; Boschma & Capone, 2016; Content & Frenken, 2016; Boschma, 2017). This literature posits that related activities demand similar capabilities which, as in economic complexity, are broadly defined and usually inferred from outcome-based empirical observations (Boschma, 2017). Thus, as for the case of country-level research, the term capabilities is not new; Maskell and Malmberg (1999) suggested that local capabilities are a combination of a region's infrastructure, natural resources, institutions, knowledge and skills.

More specifically, Frenken et al. (2007) initially distinguished between 'related' and 'unrelated' variety and suggested that related variety provides more opportunities for knowledge spillovers and learning, due to smaller cognitive distance between sectors and stronger Jacobs externalities, whereas unrelated diversity is more akin to a portfolio strategy, whereby regional income is more protected from sector-specific shocks. A central dilemma lied in defining how 'proximity' or 'relatedness' between activities and sectors should be understood (Kemeny & Storper, 2015). The 'related variety' entropy measure relied on industrial classifications and assigned proximity based on the distribution of sectors at different digit levels (Frenken et al., 2007). Conceptually, this presents several drawbacks as, firstly, the results are sensitive to the choice of aggregation level (Kemeny & Storper, 2015), secondly, industry classifications tend to favour industrial sectors over service ones (Frenken et al., 2007), and, thirdly, they present a delayed view of the industrial and technological landscape (Bishop et al., 2018), disproportionately overlooking the most dynamic and novel sectors, which we most want to capture. Moreover, classifications are based on taxonomies developed by statistical offices without relying on economic criteria that provides a clear rationale as to why certain industries should be considered related to each other or not.

Given these drawbacks, researchers turned to co-occurrence analysis in order to assess relatedness across industries, based on whether or not two industries are found together within the same region (Neffke et al., 2011). The underlying idea is that if two industries tend to co-locate in the same region, they are likely to share or require similar knowledge, technology, skills or other location-specific characteristics. The methods used are outcome-based and thus abstract in terms of pinpointing the specific complementarities that drive relatedness (Hidalgo et al., 2018). Relatedness has been widely applied to industries, technologies, products and occupations, at different geographical levels, including regions and cities. Researchers are interested in identifying how the activities found in a region determine or condition what the region is able to do in the future. There is a robust relationship between the probability that a location will enter a new industry (Neffke et al., 2011; Zhu et al., 2017), technology (Kogler et al., 2013), research area (Guevara et al., 2016), product (Hidalgo et al., 2007) or occupation (Mealy & Coyle, 2019), and the number of related activities that are already present in that location.

In this context, economic complexity can provide an additional lens through which to analyse regional paths – as argued by Mealy et al. (2018b), by creating a ranking of regions, it can shed light on the differences between the core regions racing ahead and the peripheral ones further behind in the ranking. As evidenced in this section, the focus on the importance of knowledge and capabilities has been present for decades in the regional development literature. Nevertheless, as will be further described, economic complexity provides a way of quantifying differences across regions and of assessing the relative performance of regions. As a result, some regional context applications use relatedness and economic complexity jointly for the identification of diversification strategies (e.g., Balland et al., 2019 for smart specialisation in European regions). As Davies and Maré (2020) argue, these frameworks suggest that economic complexity can complement analyses based on relatedness because it helps understand not just whether regions are moving into related areas of activity, but also provide some idea of the relative attractiveness of the direction – for instance, expanding into related activities with lower complexity levels can lead to 'lock in' and a lack of growth because local capabilities are not expanding.

Regional innovation systems and institutions

While agglomeration externalities have been recognised as crucial for regional development, other regional characteristics cannot be overlooked. The notion of regional system of innovation departed from the concept of national systems of innovation referred to in the previous subsection, and the view that a top-down approach needs to be integrated and complemented with a bottom-up perspective of development that tackles internal dynamics and socioeconomic structures that are embedded in regions, often requiring longer-term historical perspectives in order to be fully understood (Asheim, 1996; Asheim & Isaksen, 2002; Iammarino, 2005). Regional systems of innovation are defined as 'localised networks of actors and institutions, across both public and private sectors, whose activities and interactions generate, import, modify and diffuse new technologies within and outside the region' (Howells, 1999; Iammarino, 2005). Nevertheless, as argued by Cantwell and Iammarino (2003), proper regional systems of innovation are only present in a few places, with most regions lacking systemic interactions and knowledge flows between firms and other actors that are strong or integrated enough for the presence of system of innovation to be identified.

Moreover, as is the case at the national level, institutions – both formal (e.g., rule of law) and informal (such as routines, norms and values) – are increasingly regarded as crucial determinants of regional development, despite often being overlooked, due to being more subjective, controversial and inherently difficult to change (Rodríguez-Pose, 2013). Institutions determine the learning capacity of regions (Morgan, 2007), and the presence of untraded interdependencies highlighted by Storper (1997) further emphasises the importance of shared conventions embedded in regions through the positive externalities generated by local institutions (Rodríguez-Pose, 2013).

Inter-regional linkages and global production

Beyond the level of knowledge, capabilities and other regional characteristics, however, much has to be said about the external surroundings and neighbouring regions. While the literature on agglomeration economies viewed regions as self-contained and often overlooked the importance of linkages across regions, there are important ways in which regions can tap into external knowledge, which can be crucial to avoid lock-in (Asheim & Isaksen, 2002) and allows for new and related variety (Boschma & Iammarino, 2009).

More recently, Balland and Boschma (2021) investigated whether inter-regional linkages can affect the development of new activities and diversification in regions, while considering the role of relatedness. They found that being connected to regions with complementary capabilities significantly increases the probability that a region develops new technological specialisations in the European context. Furthermore, they showed that this is particularly relevant for peripheral regions which, although less diversified, see their capacity to diversify increase significantly when they connect to regions with complementary capabilities.

Given this evidence, inter-regional linkages are likely to play an important role in the context of economic complexity. Economic complexity may well be capturing different regions' capacity to tap into external markets, production capabilities and flows of people, capital and technology. Nevertheless, any discussion in this regard, or on the impact this has on economic complexity measurements is largely missing in the literature.

Taking a step further, the role of globalisation and the increased integration experienced over the past decades, as well as the major transformations they brought about, should not be overlooked. The importance of multinational corporations in technology generation (rather than simply in knowledge transfer) has long been recognised in the literature, going all the way back to Cantwell (1989). Their activity is believed to be self-reinforcing, given the nature of knowledge creation as a cumulative and localised process, thus leading to further regional concentration (Cantwell & Iammarino, 2003). Furthermore, the interaction between the local and global processes of knowledge creation can further reinforce this, leading to further disparities within countries (Cantwell & Iammarino, 2003).

Importantly, the global production network (GPN) literature takes a relational view of regions as interconnected worlds of innovation and production, emphasising the existence of exogenous sources of regional change (J. Henderson et al., 2002; Parrilli et al., 2013), and it is primarily anchored in economic geography (Coe & Yeung, 2019). The GPN framework is closely linked with the global commodity chain and global value chain (GVC) approaches – at their core, these different frameworks aim to capture the "the nexus of interconnected functions, operations and transactions through which a specific product or service is produced, distributed and consumed" (Coe et al., 2008, p.272). GPNs try to go further in two ways – first, by incorporating all kinds of network configurations, rather than just linear relationships and, second, by attempting to encompass all relevant sets of actors and relationships, rather than focusing more narrowly on inter-firm transactions (Coe et al., 2008). More recent developments in this framework and literature have moved towards attempting causal explanations for the links between global configurations and uneven regional development (Coe & Yeung, 2015; Coe & Yeung, 2019). While this field emerged in parallel with evolutionary economic geography, recent work by Yeung (2021) tried to bridge GPNs and diversification efforts, particularly by exploring the links between 'strategic coupling' in the GPN literature and related variety. More specifically, strategic coupling of regional firms and complementary actors in GPN, which is seen as a territorially-embedded mechanism that drives regional development, can provide an important lens to our understanding of diversification into related and unrelated activities (Yeung, 2021), thus further pointing to this literature's relevance for economic complexity.

The economic complexity literature has paid limited attention to multinationals, interregional linkages, global value chains and intermediate trade, particularly at a theoretical level. Some recent research has started to consider these issues empirically, for instance by focusing on value-added exports (e.g., Koch, 2021 for the case of countries), and by assessing whether GVC participation and relatedness density are associated with higher economic complexity across regions and whether GVC participation benefits higher complexity regions more than it does lower complexity regions (Colozza et al., 2021). Moreover, low complexity regions only appear to benefit from GVC participation if they have high levels relatedness density, which the authors use as a proxy for regional capabilities (Colozza et al., 2021).

These research areas bring further importance to the question of whether or not all regions that are able to competitively export a product today hold the same know-how regarding how that product is made and have the same ability to further improve or to combine the knowledge involved in producing it with other knowledge within the region to keep innovating and growing. Moreover, the role played by different regions in global production and location decisions of firms and their different functions (e.g., manufacturing, research and development, branding) have an important effect on what is captured by economic complexity. In fact, Boschma (2022) has recently put forward several opportunities for cross-fertilisation between these literature areas, including a suggestion that economic complexity may also help in the understanding of upgrading processes within the work on GPNs, among other ways.

1.3 Measuring economic complexity

1.3.1 The method of reflections

The original network-based methodology for the ECI, known as the method of reflections (MR), was first introduced by Hidalgo and Hausmann (2009) and further described in Hausmann et al. (2014b). The calculation starts from a two-mode network that links countries to the products they export competitively – defined as those products in which a country has Revealed Comparative Advantage (RCA)[1] greater than the threshold value of one.

Formally, where X_{cp} represents the exports of product p by country c, the RCA that country c has in product p is expressed as:

$$RCA_{cp} = \frac{X_{cp}}{\sum_c X_{cp}} / \frac{\sum_p X_{cp}}{\sum_{c,p} X_{cp}} \tag{1.1}$$

This measure is then used to construct the network that connects each country to the products it exports, represented by the adjacency matrix M_{cp}:

$$M_{cp} = \left\{ \begin{array}{ll} 1, & RCA_{cp} \geq 1 \\ 0, & RCA_{cp} < 1 \end{array} \right\} \tag{1.2}$$

This matrix, in which rows represent different countries and columns represent different products, summarises which countries export which products competitively and is used as the basis to construct the ECI.

From this matrix, the authors derive *diversity* and *ubiquity*, simply by summing over its rows and columns, respectively. Thus, *diversity* is the number of products that a country exports competitively, while *ubiquity* is the number of countries that are able to export a product competitively (Hausmann et al., 2014b). Formally, they are defined as:

$$Diversity = k_{c,0} = \sum_p M_{cp} \tag{1.3}$$

[1]The authors use Balassa's (1965) definition of Revealed Comparative Advantage.

$$Ubiquity = k_{p,0} = \sum_{c} M_{cp} \tag{1.4}$$

In order to generate a measure of the number of capabilities available in a country or required by a product, the information carried by each of these variables needs to be corrected using the other, thus correcting diversity with ubiquity and vice versa. For countries, this involves averaging the *ubiquity* of the products exported by a country, then averaging the *diversity* of the countries that make those products, and so forth in an iterative manner.[2] For products, it involves calculating the average diversity of the countries that export them, followed by the average ubiquity of the other products that those countries export, and so forth (Hausmann et al., 2014b).

This exercise can be expressed in the following way:

$$k_{c,N} = \frac{1}{k_{c,0}} \sum_{p} M_{cp} * k_{p,N-1} \tag{1.5}$$

$$k_{p,N} = \frac{1}{k_{p,0}} \sum_{c} M_{cp} * k_{c,N-1} \tag{1.6}$$

Inserting equation (6) into equation (5) and re-arranging yields:

$$\begin{aligned} k_{c,N} &= \frac{1}{k_{c,0}} \sum_{p} M_{cp} * \frac{1}{k_{p,0}} \sum_{c'} M_{c'p} * k_{c',N-2} \\ &= \sum_{c'} k_{c',N-2} * \sum \frac{M_{cp} M_{c'p}}{k_{c,0} k_{p,0}} \end{aligned} \tag{1.7}$$

This can be re-written as:

$$k_{c,N} = \sum_{c'} \widetilde{M}_{cc'} k_{c',N-2} \tag{1.8}$$

Where:

[2] The two measures are iterated until the addition of one more iteration does not alter the order of the ranking of countries. As the number of iterations increases, it becomes harder to define the exact meaning of each, but as shown by Hidalgo and Hausmann (2009), the ECI becomes more highly correlated and a better predictor of income.

$$\widetilde{M}_{cc'} = \sum_{p} \frac{M_{cp} M_{c'p}}{k_{c,0} k_{p,0}} \tag{1.9}$$

As described by Hausmann et al. (2014b), this is satisfied when $k_{c,N} = k_{c,N-2} = 1$, which corresponds to the eigenvector of $\widetilde{M}_{cc'}$ that is associated with the largest eigenvalue. Since this eigenvector is a vector of ones, it does not reveal useful information. Thus, the authors look for the eigenvector associated with the second largest eigenvalue, which is the eigenvector that captures the largest amount of variance in the system and is the measure of economic complexity. Thus, the ECI is defined as:

$$ECI = \frac{\vec{K} - < \vec{K} >}{stdev(\vec{K})} \tag{1.10}$$

where $\vec{K}$ is the eigenvector of $\widetilde{M}_{cc'}$ associated with the second largest eigenvalue, $<>$ represents an average and *stdev* the standard deviation.

The authors coin this the method of reflections as, due to the symmetry of the two-mode network, it produces a symmetric set of variables for each of the two types of nodes in the network – countries and products (Hidalgo & Hausmann, 2009). Therefore, the Product Complexity Index (PCI) can be obtained analogously by exchanging the index for countries (c) with that for products (p) in the previous mathematical definitions, obtaining:

$$PCI = \frac{\vec{Q} - < \vec{Q} >}{stdev(\vec{Q})} \tag{1.11}$$

where $\vec{Q}$ is the eigenvector of $\widetilde{M}_{cc'}$ associated with the second largest eigenvalue.

1.3.2 Alternative interpretations and reformulations

Since the introduction of the method of reflections by Hidalgo and Hausmann (2009), other authors have examined this methodology and proposed modified interpretations and reformulations. Kemp-Benedict (2014) and Mealy et al. (2018b, 2019) challenge the interpretation of the eigenvector of the matrix and argue that the ECI is orthogonal to diversity, rather than a 'corrected diversity' measure, and thus captures features that are distinct from diversity.[3] Furthermore, Mealy et al. (2018b, 2019) put forward a new interpretation, whereby the ECI represents a ranking that places countries or regions with similar exports

[3] Diversity in this context refers to the number of products, or activities more generally, in which a country has revealed comparative advantage.

closer together in the ordering and those with distinct exports further apart. The authors argue that, as a result, the ECI is able to shed light on what separates richer and more complex regions from those with lower income and complexity levels, and thus is more useful in sub-national contexts than simpler diversity measures Mealy et al. (2018b).

Separately, a group of researchers proposed a reformulation of the methodology and introduced the fitness-complexity method (Tacchella et al., 2012; Cristelli et al., 2013; Tacchella et al., 2013). The authors plot the country-product matrix and show that, by listing countries in an increasing order of specialisation and products in decreasing order of diffusion, a triangular shape is obtained, indicating that countries tend to make all the products they are able to, given their technological and development levels. This matrix structure implies that the information that a product is made by a diversified country conveys little information regarding the complexity level of that product, as diversified countries export almost all products; conversely, if a less advanced country is able to export a product, it is very likely that this product requires only the low level of technological sophistication of that country (Cristelli et al., 2013). As a result, the authors propose a non-linear approach that guarantees that the only possibility for a product to have a high complexity level is to be produced exclusively by technologically advanced countries.

Morrison et al. (2017) analyse the fitness-complexity method further, from a methodology perspective, through simulations and real data on exports and patents, and show that the method is inherently unstable to minimal perturbations in the network. In sparse networks, this instability becomes particularly severe, and the method assigns disproportionately high complexity levels to niche products exported competitively by very few countries (Morrison et al., 2017). This is driven by the non-linear nature of this method (thus, not applicable to the original ECI methodology), and to its definition that assigns a lower complexity to products produced by many producers, and ultimately limits complex products to the most advanced countries; the iterative process further amplifies this, ultimately meaning that the fitness-complexity of countries that produce a limited number of unique products converges to zero as the number of iterations increases (Morrison et al., 2017). More recently, Sciarra et al. (2020) combine the ECI and Fitness methodologies through a multidimensional framework and linear algebra tools to develop yet another metric, the GENeralised Economic comPlexitY index (GENEPY), which they argue combines the strengths of both methods.

Another important point raised by Morrison et al. (2017) relates to trade in intermediate products and value added – in particular, if a country imports intermediate inputs, that has an impact on competitiveness. Koch (2021) addresses precisely this point, by investigating whether complexity measures based on value-added exports (rather than gross exports) can better explain the link with economic growth. They construct a value-added exports fitness complexity metric, and show that the rankings obtained differ from those with the existing ECI and fitness-complexity methods; furthermore, their metric appears to have a stronger association with economic growth than existing ones (Koch, 2021).

Despite these differing interpretations and reformulations of the ECI, it has been shown that the resulting metrics are highly correlated with each other (Albeaik et al., 2017). For the purpose of this review chapter and the remainder of the thesis, we will carry on using the methodology originally introduced by Hidalgo and Hausmann (2009). A deeper investigation of the mathematical methods is beyond the scope of this thesis – while scrutiny is very important, there are several authors making contributions on this front; rather, we focus on the conceptual and theoretical issues that have yet to be addressed.

1.4 Empirical applications

Given the nature of the ECI as an outcome-based measure and mainly an empirical exercise, most of the literature consists of empirical applications at several geographical levels and across different development contexts. This section explores this literature, distinguishing between cross-country, country level and sub-national applications. We refer to Hidalgo (2021) for a comprehensive review of empirical applications to date, including both economic complexity and relatedness metrics.

1.4.1 Cross-country applications

The Atlas of Economic Complexity by Hausmann et al. (2014b) and the accompanying website[4] contain trade data and ECI calculations for 250 countries and territories.[5] This data and methodology have been used in cross-country econometric analyses that emphasise the strong association between economic complexity and economic growth across the globe.

Hidalgo and Hausmann (2009) provide the original empirical application. Firstly, they show that the ECI measures are correlated with income per capita and that the correlation

[4]http://atlas.cid.harvard.edu/

[5]The raw trade data on goods originates from the United Nations Statistical Division (COMTRADE) and is available in two trade classification systems, the Harmonized System (HS) 1992 and Standard International Trade Classification (SITC) revision 2 at different digit detail levels in both cases. HS data covers approximately 5000 goods, for the period 1995-2017. SITC data covers approximately 700 goods, for the period 1961-2017.

becomes stronger as the number of iterations – i.e., the number of times that ubiquity and diversity are averaged – increases. Secondly, they demonstrate that the ECI is associated with future economic growth by regressing income per capita growth on a country's initial income and ECI levels. These results hold for growth periods of different lengths (20-, 10- and 5-year periods) and are robust to the inclusion of country dummy variables. Thirdly, they show that the ECI outperforms the Herfindahl-Hirschman Index as well as entropy measures (e.g., related variety) in explaining economic growth. Finally, they highlight the path dependent nature of the ECI, which indicates that a country's current productive structure will influence the new products a country is able to export in the future, and argue that this is "consistent with the existence of an unobservable capability space that evolves gradually, because the ability of a country to produce a new product is limited to combinations of the capabilities it initially possesses plus any new capabilities it will accumulate" (Hidalgo & Hausmann, 2009, p. 10575).

Hausmann et al. (2014b) perform a similar analysis and regress the annualised growth rate of Gross Domestic Product (GDP) per capita on initial ECI and two control variables – the initial level of GDP per capita in each period and the increase in natural resource exports (as a share of initial GDP).[6] Furthermore, the authors show that, while the ECI is constructed using export data, its association with future growth is not driven by export growth, export concentration or population size. After controlling for increase in exports (in both goods and services, as a share of initial GDP), openness (exports to GDP), initial export concentration (Herfindahl-Hirschman index) and population in separate specifications, the coefficient on initial ECI remains strong and statistically significant.

Relying on their own complexity calculations using the method of reflections and data from Harmonised System (HS) 6-digit level classifications for 5132 products and 176 countries, Felipe et al. (2012) investigate trends in complexity and country development levels and compare the ECI with other measures of technological capability.[7] They find that, as expected, high-income countries are the biggest exporters of more complex products, whereas low-income countries largely export lower complexity products, and that the export share of higher complexity products increase with income, while the export share of the lower complexity products decrease with income (Felipe et al., 2012). Additionally, they find strong correlations between the method of reflections and other measures of technological capabilities.

[6]For the periods 1978-1988, 1988-1998 and 1998-2008.

[7]These consist of the indicator of technological capabilities by Archibugi and Coco (2004), the technology achievement index by Desai et al. (2002), the technology effort and industrial performance indices by Lall and Albaladejo (2002) and the science and technology capabilities index by Wagner et al. (2001).

Other cross-country analyses investigate the link between economic complexity and income inequality within countries, using data from the Atlas of Economic Complexity. Hartmann et al. (2017) show that across 79 countries, those with a higher ECI (i.e., exporting more complex products) have lower levels of income inequality, measured by the Gini coefficient, than countries with lower ECI levels. This negative correlation between economic complexity and income inequality is robust to controlling for income levels, institutions, export concentration and human capital (Hartmann et al., 2017). Hartmann et al. (2016) compare the productive structures of countries in Latin America and Caribbean with those of China and other high-performing Asian economies. The authors show that there is an increasing gap in the productive capabilities of Latin American countries and that the mix of products they export are associated with higher income inequality than those exported by China and the other Asian economies considered.[8] These papers argue that the ECI can also shed light on countries' income distributions, and that productive structures may help or hinder countries' efforts to reduce inequality.

Economic complexity has also been linked with environmental concerns. Researchers have linked the ECI with different environmental outcomes, showing a positive association with indicators on environmental performance, but a negative association between economic complexity and air quality (Boleti et al., 2021). Others have tested out the hypothesis of an environmental Kuznets curve, whereby emissions initially grow and then decline with economic growth – Chu (2021) find an inverted U-shaped relationship between economic complexity and carbon dioxide emissions, for a panel of 118 countries. In contrast, Romero and Gramkow (2021) do not find support in favour of an environmental Kuznets curve, but rather they find a negative and statistically significant association between economic complexity and greenhouse gas emissions per capita. The authors also go further and develop a Product Emission Intensity Index (PEII), which provides and weighted average of the emissions levels of countries that export a product competitively, thus assessing which products are associated with higher or lower emissions intensity (Romero & Gramkow, 2021). They find that more complex products tend to be associated with lower greenhouse gas emissions intensity.

Moreover, researchers developed indices aimed at capturing countries' positions in terms of their current productive capabilities or the extent to which they are stranded with productive structures relating to highly polluting products, with a focus on transition to greener activities. Namely, the Green Complexity Index (GCI) – aimed at capturing

[8] Their analysis is based on the Product Gini Index (PGI) introduced in Hartmann et al. (2017), which is a product-level measure of the income inequality that is expected for countries exporting a given product.

"the extent to which countries are able to competitively export green, technologically sophisticated products" (Mealy & Teytelboym, 2020, p.2), and the Brown Complexity Index – aimed at assessing the "dependence on brown activities which provide fewer and fewer opportunities to the economy as the green transition progresses" (Andres et al., 2023, p.3). These depart from the ECI methodology and logic, but focus on different networks, with the aim of understanding to what extent countries are more or less well-equipped to navigate the transition towards greener or less polluting activities.

As the vast majority of these papers relies on export data, they do not consider services and focus strictly on products. There have been efforts to include services in the ECI measures available on the Atlas of Economic Complexity, but this presents challenges as the level of granularity of data for services is much lower than that for products. Stojkoski et al. (2016) were the first ones to address this issue by relying on the World Bank Trade in Services database and aggregating product data further, so that products and services are comparable and can be meaningfully included in the same network as the basis for calculation of the ECI. Based on 22 broad categories (10 capturing goods and 12 capturing services) for 130 countries, they find that their aggregated ECI measure is a statistically significant predictor of long-term economic growth, despite significant differences in rankings vis-a-vis the traditional ECI, particularly for countries where service sectors play an important role (Stojkoski et al., 2016). Since then, Mishra et al. (2020) and Patelli et al. (2022) relied on Balance of Payment Statistics by the IMF to integrate services into Fitness and ECI calculations as well. Drawing comparisons between the relative complexity of products and services, Mishra et al. (2020) find that services, particularly modern ones, are among the most complex activities. Neither of the papers looks at outcomes such as income growth, but we might start seeing more research using the integrated complexity database provided by Patelli et al. (2022).

The broader limitations of using trade data are recognised in the literature and have been recently addressed by Stojkoski et al. (2023), who built economic complexity measures across countries using patents (technology complexity) and research publications (research complexity), in addition to the trade-based ECI (trade complexity). They compare these three measures in terms of their predictive power for economic growth, income inequality and greenhouse gas emissions across countries and find that technology and trade complexity together can explain all three outcomes across countries (even when combined), while research complexity only had a positive association with greenhouse gas emissions. The authors also validate their results by instrumenting economic complexity with the average complexity levels of a country's three most structurally similar non-neighbouring countries, in an attempt to address endogeneity concerns relating to the possibility that local aspects could be driving both complexity levels and other outcomes, which further supports their findings. Due to data limitations, the paper looks at a more limited sample of countries and fewer years than previous key contributions (Stojkoski et al., 2023); moreover, as will be further explored in the discussion, there are important drawbacks in relying in these two alternative sources of data across countries, even though they are more widely used in sub-national contexts.

Lastly, some recent papers have looked at the determinants of economic complexity, investigating what factors may explain why some countries achieve higher ECI levels than others. In particular, researchers have looked at intellectual property rights (Sweet & Maggio, 2015), taxation (Lapatinas et al., 2019), demographics and cultural diversity (Bahar et al., 2020), technology proxied through internet access (Lapatinas, 2019), and inward FDI (Antonietti & Franco, 2021). These papers focus on export data from the Atlas of Economic Complexity and find each of these factors to offer some explanation to ECI levels across countries. Nevertheless, the underlying mechanisms with regards to the ECI at the cross-country level remain unclear.

1.4.2 Country-level applications

In addition to macro analyses drawing comparisons across a large range of countries, this data and method have also been applied to country-level reports, particularly in lower income countries such as Uganda (Hausmann et al., 2014a), Panama (Hausmann et al., 2017) and Rwanda (Hausmann & Chauvin, 2015). These country reports all follow a similar structure – they start by looking at the growth trajectory experienced by each country and the major challenges faced, followed by an analysis of structural transformation, economic complexity and diversity levels in each country (still comparing to the rest of the world), and the identification of existing binding constraints (e.g., limited access to finance, lack of skills, inadequate infrastructure, weak institutions). Finally, they propose policy actions, which involve export diversification into products that are close to the country's current productive structure and that present opportunities in global and regional markets. While these policy implications are not unprecedented and could have been reached through existing analytical methods, the ECI provides an additional tool for understanding a country's productive structure relative to the rest of the world.

Relatedly, Hartmann et al. (2019) look at the case of Paraguay and develop an analytical framework to identify smart strategies for economic diversification and inclusive growth that identifies the most feasible and desirable products for Paraguay to start producing.

The authors do this by measuring the expected level of income, economic complexity, technology and income inequality associated with each product[9], and present a scoreboard that allows for a consideration of different diversification strategies (Hartmann et al., 2019).

Country-level applications are not limited to lower income countries. Zaccaria et al. (2016) look at the case of the Netherlands, with the aim of illustrating how economic complexity can be a useful tool to analyse a country's competitiveness. Using the fitness complexity method by Tacchella et al. (2012), they examine the top exported products, the complexity levels of different sectors and how these have changed over time, and compare the Netherlands to different countries. From this, they examine trends in different sectors and identify the most promising ones. O'Clery (2015) do a similar exercise for the case of Ireland, highlighting different key sectoral clusters in the country.

1.4.3 Sub-national applications

The economic complexity concept and method was quickly adopted in research at the sub-national level. Some researchers followed the original application and relied on exports to measure complexity across regions. Among the first ones, Poncet and Starosta de Waldemar (2013) examined the link between initial complexity level and subsequent GDP per capita growth across Chinese cities, using export data and the method of reflections and found a positive correlation between initial ECI and subsequent income growth, robust to controlling for initial income and traditional economic growth determinants, such as human capital, openness and Foreign Direct Investment (FDI). They conducted several robustness checks, including changing the threshold levels of RCA calculations, excluding the top decile of exporting cities and conducting an alternative system-GMM estimation. In a similar nature, Zhu et al. (2020) also calculated complexity across Chinese regions with export data and analysed the link with income inequality, focusing not just on exports' complexity levels, but also on the complexity level of the destination countries. Export data has also been used by Reynolds et al. (2018) for the case of Australia, with a focus on understanding the relative competitiveness of different territories, rather than explaining outcomes across them.

Beyond applications of the method of reflections to a two-mode network regions or cities to the products they export, researchers started applying this methodology to alternative types of data, such as patents and employment. These papers make up an important share of the research done at the sub-national level.

[9]See footnote 8.

Measures based on networks linking sub-national regions and patents have been referred to as knowledge or technological complexity. The first to develop these applications were Balland and Rigby (2017), who draw on a two-mode network of cities and patents, which they coin the city-tech knowledge network, and calculate the Knowledge Complexity Index (KCI) for US metropolitan areas, with the aim of capturing their technological structure. The authors found wide variations in knowledge complexity levels across metropolitan areas, with only a few specialising in the most complex technologies. Furthermore, their analysis shows that, in line with theoretical expectations and past research, not all knowledge is spatially sticky and that the extent to which it moves across space depends on the complexity level, with lower complexity knowledge travelling more easily across metropolitan areas, thus providing a precarious source of competitive advantage (Balland & Rigby, 2017).

Since then, patent data has been widely used in applications to European regions (Antonelli et al., 2017; Balland et al., 2019; Antonelli et al., 2020; Mewes & Broekel, 2022; Pinheiro et al., 2022; Pintar & Scherngell, 2022). These papers tend to explore different outcomes – such as regional economic growth (Mewes & Broekel, 2022; Pintar & Scherngell, 2022), innovation, and generation and exploitation of new technologies (Antonelli et al., 2017, 2020) – or be coupled with relatedness to build smart specialisation strategies and understand diversification trajectories across different regions (Balland et al., 2019; Pinheiro et al., 2022). In addition to these broad contexts of US cities and European regions, there is also research using patent data to explore regions within individual countries, such as the work by (Whittle, 2019) for the case of Ireland.

Beyond this, research exploring individual countries at the sub-national regional level tend to use employment data. One of the reasons for this is that employment data covers a lot of economic activity (including services) and tends to be available at quite granular levels within individual countries, but is not always easily comparable across nations. Key contributions have explored several countries, including Norway (Broekel et al., 2021), Sweden (Hane-Weijman et al., 2022), France (Lo Turco & Maggioni, 2020), New Zealand (Davies & Maré, 2020), the United Kingdom (Mealy et al., 2018b; Mealy & Coyle, 2022), and the United States (Mealy et al., 2018b).

Across these papers, as we might expect, there is a strong correlation between the ECI and regional income per capita, with higher ECI levels in cities or urban areas and lower levels in rural areas. Importantly, Fritz and Manduca (2021) measured economic complexity of US metropolitan areas using employment data and, while they argue that the spatial distribution of economic activity across US metropolitan areas is suitable for measuring economic complexity, their findings on the links between complexity and income changes are mixed – across metropolitan areas, higher complexity is associated with higher economic growth, while within metropolitan areas over time they find a negative association – and thus they call for caution when referring to the links between economic complexity and different outcomes at the sub-national level (Fritz & Manduca, 2021). It is also important to note that by looking at the case of the United States (US), they are drawing on a larger network than applications to the United Kingdom (UK) and other smaller countries, and the suitability for smaller countries has not been discussed in the literature.

Bishop et al. (2018) and Bishop and Mateos-Garcia (2019) also analyse the case of the UK and go a step further in investigating the mechanisms behind the relationship between the ECI and income per capita. In both papers, the authors combine industrial activity and business website data for the UK and employ Natural Language Processing to identify what drives the emergence of new ideas across LADs. They find that locations with higher ECI scores have a stronger share of companies active in emerging technologies, even when controlling for industrial composition (Bishop et al., 2018). Bishop and Mateos-Garcia (2019) also compare the ECI and fitness complexity measures[10] and find strong correlations between them, reaching broadly similar results across the different measures. Economic complexity is higher in urban travel-to-work areas and involves knowledge-intensive, creative and digital sectors (Bishop & Mateos-Garcia, 2019).

Finally, Balland et al. (2020) moved away from the method of reflections and defined complex activities as those that require more profound division of knowledge and labour. They suggest that what makes an activity complex is the need for a large network of people with complementary (advanced) knowledge, rather than simply requiring an individual that is more skilled than is required for another activity (Balland et al., 2020). Thus, they use data on scientific publications, patents and employment, proxying complexity by the age of the knowledge combined in patents, the average size of the team involved in a scientific publication, and the average number of years of education of the employees working in an occupation or industry, respectively.[11] They analyse 353 metropolitan areas in the US and show that more complex economic activities concentrate disproportionately in large cities (Balland et al., 2020). However, their analysis is simply descriptive, and they do not pinpoint the mechanisms linking complex activities and spatial concentration.

[10] Calculated using the fitness complexity method by Tacchella et al. (2012).

[11] The age of the knowledge combined in patents is measured by the average years of emergence of the sub-classes in which a patent makes a knowledge claim (Balland et al., 2020).

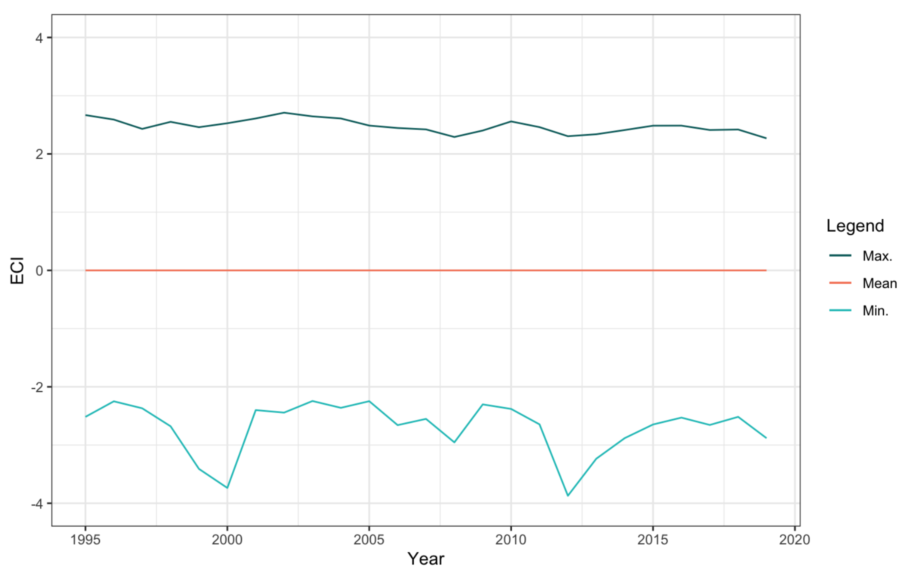
1.5 Exploring the data – Export-based ECI across countries

This section presents an exploratory analysis of economic complexity, focusing on the original export-based ECI for countries. The aim is to understand what is captured by the ECI in its original conceptualisation and to identify key trends.

Our economic complexity measure relies on export data downloaded from the Observatory of Economic Complexity (OEC), based on the BACI international trade database at the product level, and we use the HS-1992 four-digit level classification.[12] The yearly ECI (and PCI) measure is calculated from a network with 1241 products and 179 countries, for the period from 1995 to 2019.[13] It is based on the Method of Reflections introduced by Hidalgo and Hausmann (2009), following the equations and method described in section 1.3 of this chapter, and an RCA threshold of 1.[14]

Figure 1.2 shows average, maximum and minimum ECI values each year from 1995 to 2017. The ECI is centred around zero for the entire time period (due to the way it is constructed). The maximum values remain relatively stable at around 2.5 over the time period, whereas the minimum values seem much more volatile, ranging from just under -2 to below -3.

Figure 1.2: Maximum, minimum and average ECI values, from 1995 to 2019



[12]We use the four-digit level classification as it provides enough granularity and country reporting at this level is more reliable than at the six-digit level.

[13]We departed from all the countries with export data and selected those that had population over 200,000 in the year 2000, giving us a total of 179 countries (of which ten have some early years of export data missing). More details on the countries covered are provided in Chapters 3 and 4.

[14]The measure was built using the economiccomplexity R package and the default of 20 iterations.

We also draw on several variables that are often used to draw comparisons across countries. These include economic variables, such as GDP, trade and employment-related measures; investment and innovation data, including R&D measures and Foreign Direct Investment (FDI); human capital data, including education enrolment and the Human Development Index (HDI); poverty and inequality indicators; urban population and living conditions data; and, finally, several institutional and business environment indicators.

The vast majority of the additional variables was downloaded from the World Bank Open Data, including the institutional variables, which come from the World Bank's Worldwide Governance Indicators database.[15] The HDI originates from the United Nation's Human Development Reports.[16] Table 1.A.1 provides definitions and sources for the variables used, while Table 1.A.2 presents summary statistics (see Appendix).

Economic complexity across countries

We start by investigating the ECI data by itself in more detail, to investigate how complexity levels vary across different countries. Figure 1.3 shows the average ECI values for the 1995-2019 period for each country, ordered according to ECI level and split in two panels, each covering positive or negative average ECI values. These plots show the contrast between continents, with most countries in Europe showing positive ECI values, while a lot of countries in Africa present very low or negative ECI values.

In general, the countries at the top of the first panel, which have the highest positive ECI values, tend to be technologically-advanced and high-income countries. In contrast, the countries at the bottom of the second panel, presenting negative values, are countries that have been war-struck and/or are heavily reliant on natural resources. While this follows rough expectations, some inconsistencies are evident. For example, São Tomé and Príncipe scores above higher-income countries in Europe and Latin America that we might expect to see ranked higher.

To complement this, and explore actual ECI values across different years rather than averages, Table 1.1 shows the top and bottom five countries according to economic complexity for the years 1995, 2007 and 2019. Across each of these years, the top five countries remained relatively stable and represent advanced democracies with high income per capita, with the exception of Timor-Leste which ranks as fifth highest in 1997, counter-intuitively given its income level and broader socioeconomic context. The bottom five countries have among the lowest levels of income per capita, and some are war-torn places. In the first

[15] Available at: https://data.worldbank.org/
[16] Available at: http://hdr.undp.org/en/data

Figure 1.3: Average ECI from 1995 to 2019, panel 1 (positive values)

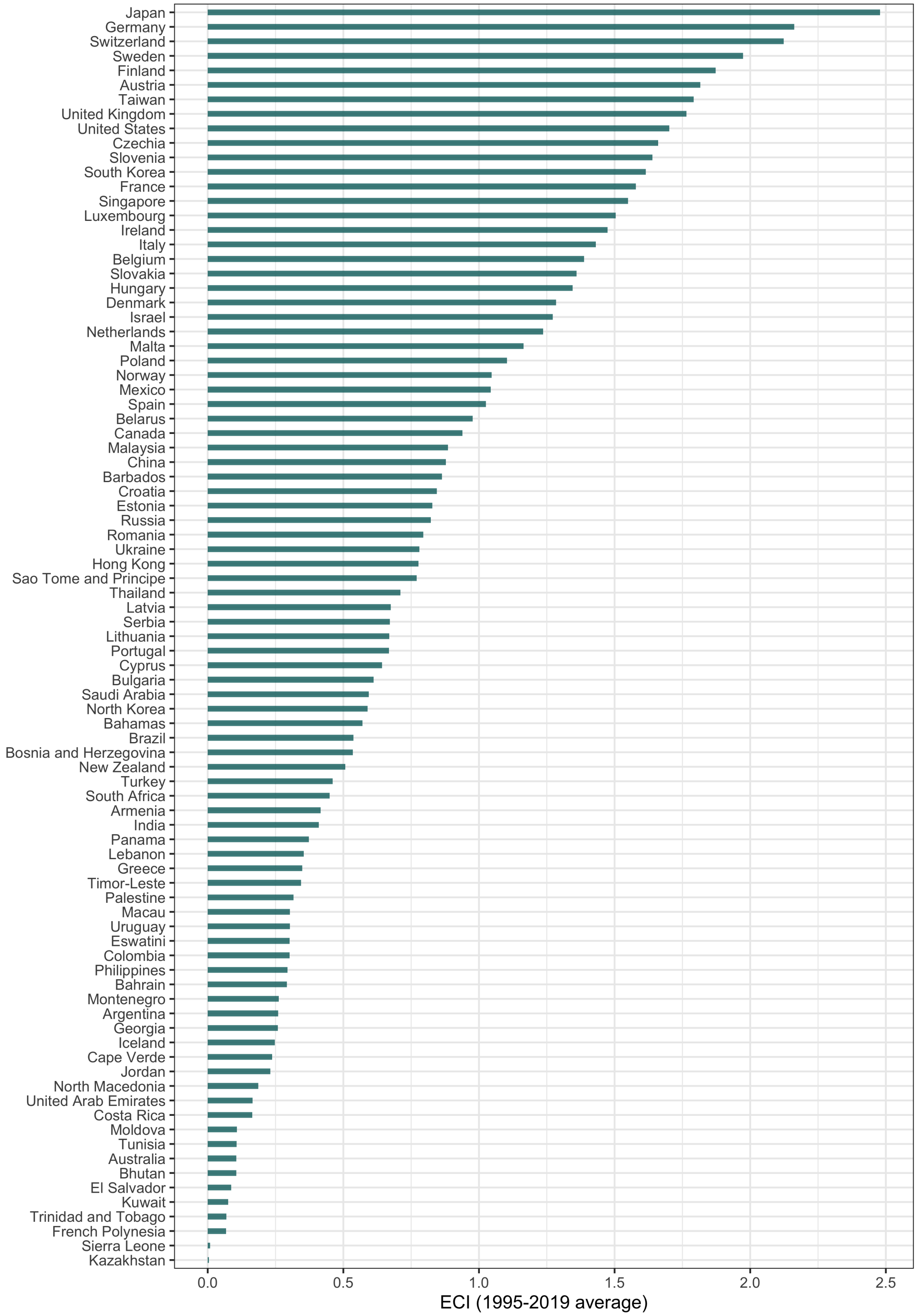
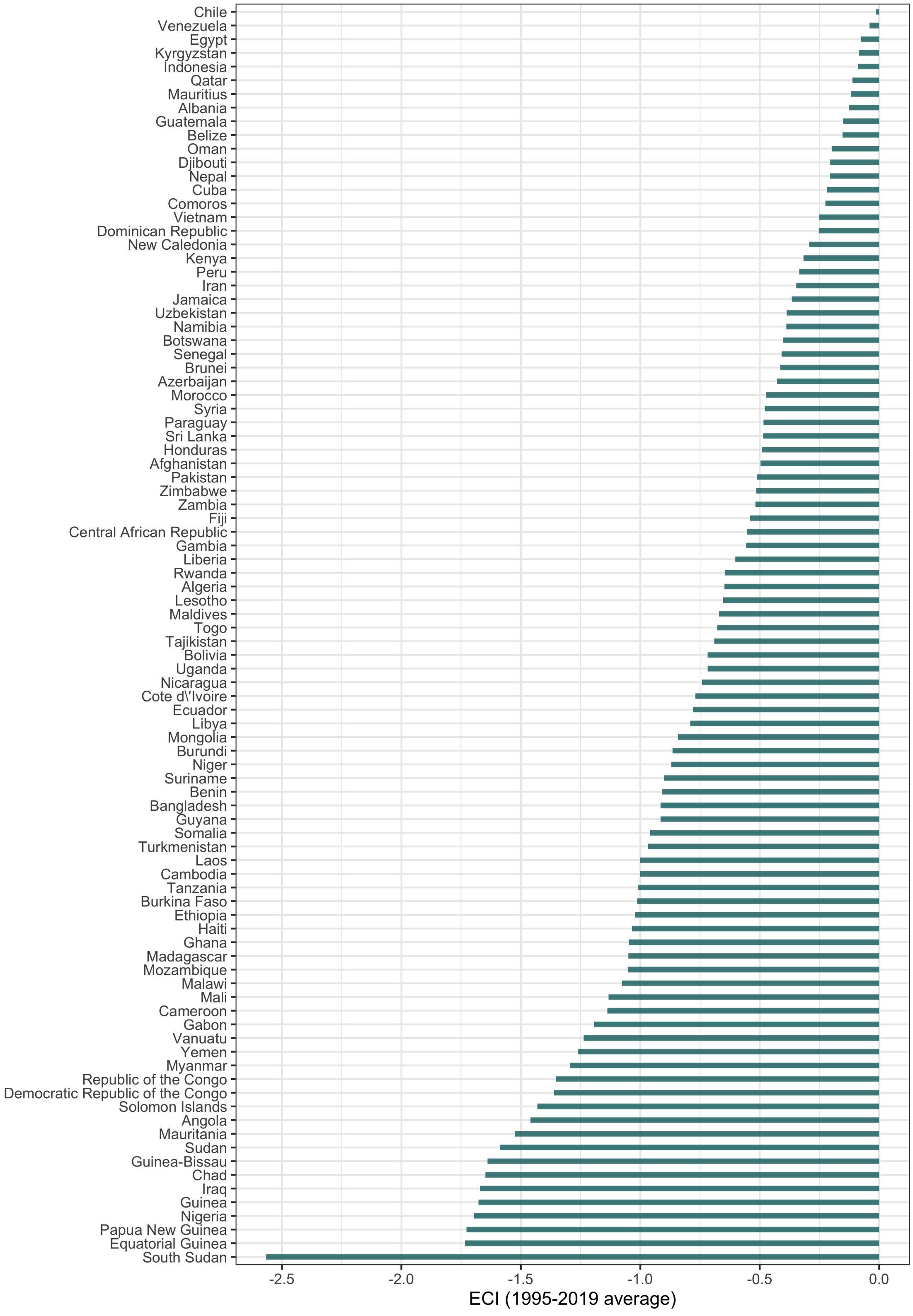


Figure 1.3 (continued): Average ECI from 1995 to 2019, panel 2 (negative values)



panel of Figure 1.3, we can see that the average ECI level in Timor-Leste is below 0.5 (albeit not among the lowest), pointing to sharp swings in ECI values across the period. Chapters 3 and 4 will explore in more detail changes in ECI over time.

Table 1.1: Top and bottom 5 countries by economic complexity in 1995, 2007 and 2019

Top 5 countries by economic complexity each year				***Bottom 5 countries by economic complexity each year***			
1995				**1995**			
ECI Rank	**Country**	**ECI**	**GDPpc**	**ECI Rank**	**Country**	**ECI**	**GDPpc**
1	Japan	2.667	34868	165	Mauritania	-1.961	4674
2	Germany	2.345	39366	166	Nigeria	-1.993	2902
3	Sweden	2.242	34234	167	Angola	-2.328	4140
4	Switzerland	2.235	54079	168	Equatorial Guinea	-2.349	1781
5	Timor-Leste	2.157	-	169	Papua New Guinea	-2.516	3272
2007				**2007**			
ECI Rank	**Country**	**ECI**	**GDPpc**	**ECI Rank**	**Country**	**ECI**	**GDPpc**
1	Japan	2.421	39281	174	Mauritania	-1.724	4979
2	Germany	2.239	47101	175	Iraq	-1.907	7695
3	Sweden	2.081	48557	176	Sudan	-2.036	4834
4	Switzerland	2.021	65642	177	Papua New Guinea	-2.048	3119
5	Finland	2.010	48664	178	Chad	-2.551	1570
2019				**2019**			
ECI Rank	**Country**	**ECI**	**GDPpc**	**ECI Rank**	**Country**	**ECI**	**GDPpc**
1	Japan	2.268	42022	175	Equatorial Guinea	-2.080	18503
2	Taiwan	2.105	-	176	Guinea	-2.112	2567
3	South Korea	2.084	42805	177	South Sudan	-2.329	-
4	Switzerland	1.986	70944	178	Guinea-Bissau	-2.527	1939
5	Germany	1.908	53930	179	Chad	-2.882	1580

Notes: Ranks are based on the total number of countries with ECI data available for that specific year. GDPpc refers to GDP per capita, PPP (constant 2017 international $).

Economic complexity, income per capita and other variables

The next step is to investigate the ECI data alongside GDP per capita. Figure 1.4 shows scatter plots of the correlation between ECI and the natural log of GDP per capita across countries, for 1995 and 2019, the start and end of our period of analysis. In both cases, we observe a positive relationship between the ECI and income per capita, that is more accentuated in the final year. These plots also allow us to see where countries from different regions fall in this correlation. European and Central Asian countries are more represented in the top-right quadrant, with both high ECI and high GDP per capita figures, whereas Sub-Saharan African countries are more widely represented in the bottom-left quadrant, with the lowest ECI and GDP per capita values. In contrast, countries in the East Asia and Pacific region are more evenly represented across the plot, and in terms of ECI and GDP per capita levels.

Beyond GDP per capita, the ECI may also be correlated with other variables of interest related to the economy, education, institutions, development, among other areas. Figure 1.5 shows the correlations between the ECI and several variables for averagevalues 1995-2019. Aggregate measures such as the Human Capital Index (HCI) and HDI show up at the top, with very strong positive correlations. In addition, variables related to education, institutional quality and R&D also show strong positive correlations with economic complexity, which are higher than the correlation with GDP per capita.

Among the strongest negative correlations in Figure 1.5 are variables related to the importance of natural resources and agriculture, as well as poverty and two of the inequality indicators. Several variables lie between the two extremes, showing only moderate or no correlation with economic complexity – for instance, the population variables, unemployment, FDI, and trade-related variables such as imports, merchandise trade as a percentage of GDP and, to a lesser extent, exports.

Economic complexity, country size and composition effects

An important possibility to consider is that the ECI simply reflects certain aspects of countries' size or economic structure, with the method disproportionately favouring countries that export more extensively than others (e.g., due to geographic characteristics), or are more reliant on a specific macro-sector. To go a step further in investigating this possibility, we present scatter plots between the ECI and other key variables to identify any patterns.

We start by looking at whether the ECI is closely related to population size. For instance, it could be that larger countries have different diversification or specialisation strategies, leading to higher economic complexity. To this end, Figure 1.6 presents scatter plots of the ECI and the natural logarithm of population for 1995 and 2019. In both cases, as expected from Figure 1.5, there is no apparent correlation between the ECI and population size.

Following this, Figure 1.7 presents scatter plots of the ECI and exports of goods and services (as a percentage of GDP) in 1995 and 2019. In 1995 there is no correlation between the ECI and exports, while in 2019 there is a positive correlation, though it might be, to some extent, driven by outliers. Thus, the ECI does not appear to be heavily affected by the importance that exports of goods and services play in a country's economy.

Lastly, Figure 1.8 presents scatter plots of the ECI and the share of employment across industry, services and agriculture in 1995 and 2019 (region colours follow previous figures).

This allows us to investigate the possibility of composition effects in the ECI — for instance, the extent to which countries with a larger manufacturing sector tend to be 'favoured' in ECI estimations, leading to higher complexity levels. Across the two periods, both the share of employment in industry and in services show positive correlations with the ECI – in the case of employment in industry, the correlation is weaker in 2019 than at the start of the period, while for employment in services it is more accentuated in the final year. In contrast, there is a negative correlation between the share of employment in agriculture and the ECI, as expected, as countries more reliant on agriculture tend to export simpler goods, which are more widely accessible to other countries and therefore more ubiquitous. Interestingly, the correlation between the share of employment in services (which assesses the relative importance of the service sector in a country) and the ECI in 2019 is stronger than that for employment in industry, despite ECI calculations not taking into account services. This may be driven simply by the fact that higher income countries have both relatively large service sectors and high economic complexity.

Figure 1.4: Scatter plots of ECI and GDP per capita across countries, 1995 and 2019

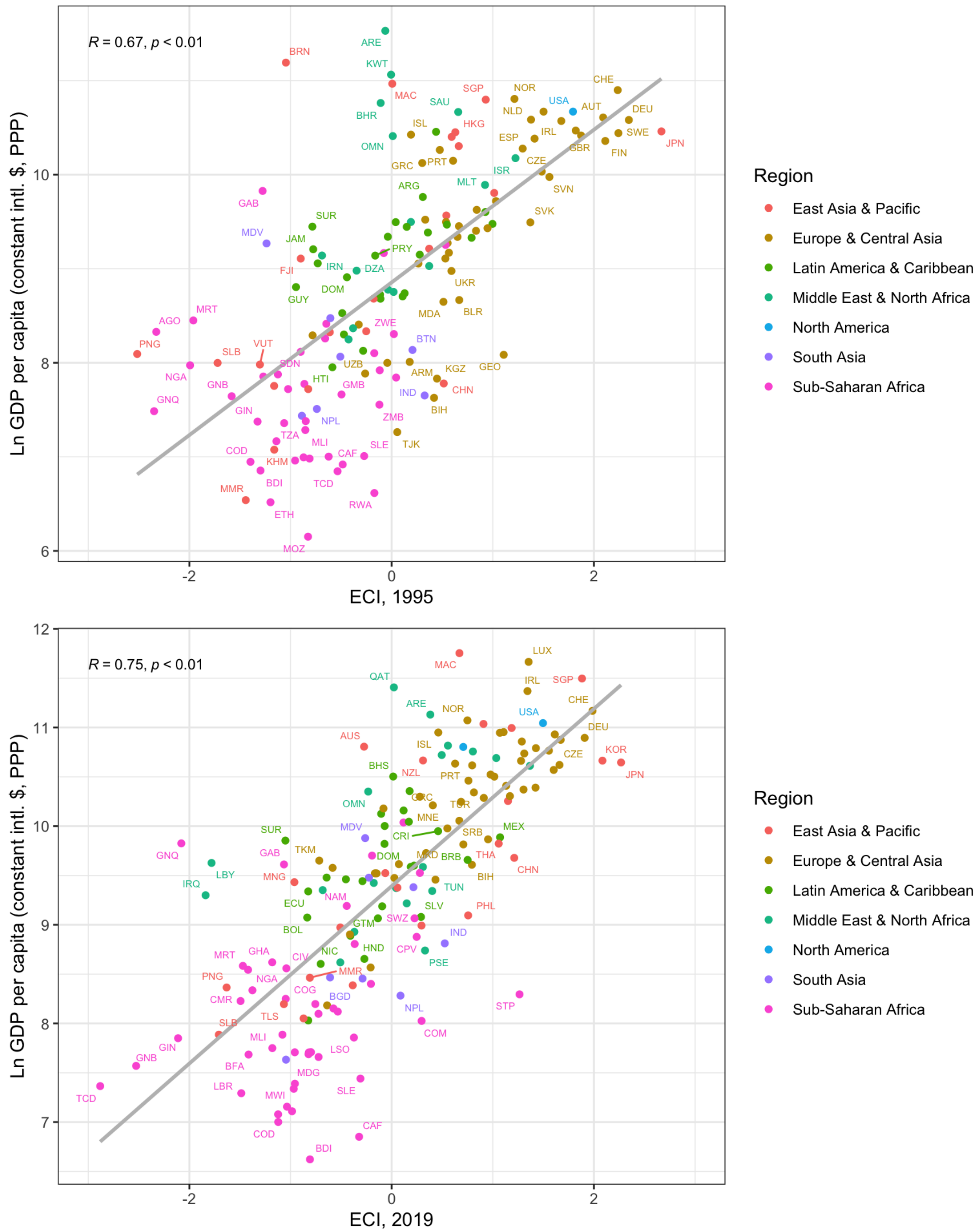


Figure 1.5: Correlation between ECI and other variables, 1995 to 2019 average

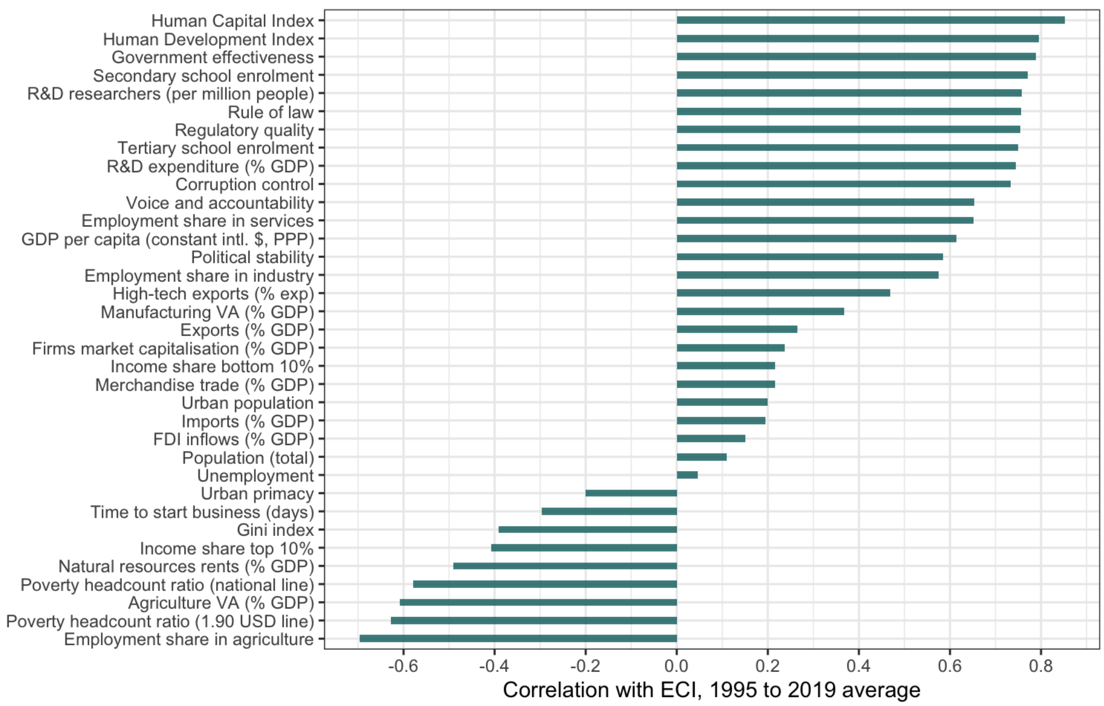


Figure 1.6: Scatter plots of ECI and population size across countries, 1995 and 2019

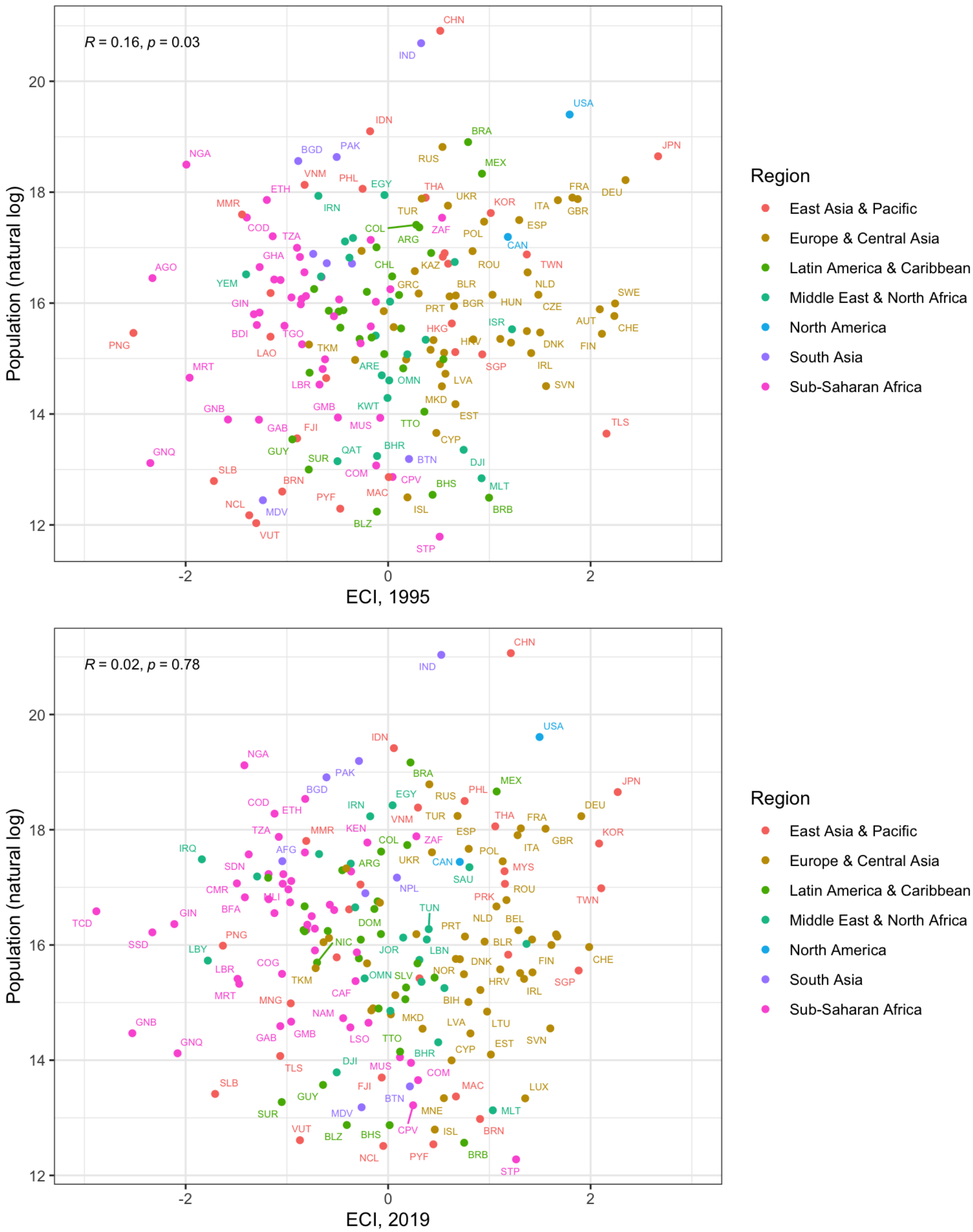


Figure 1.7: Scatter plots of ECI and exports across countries, 1995 and 2019

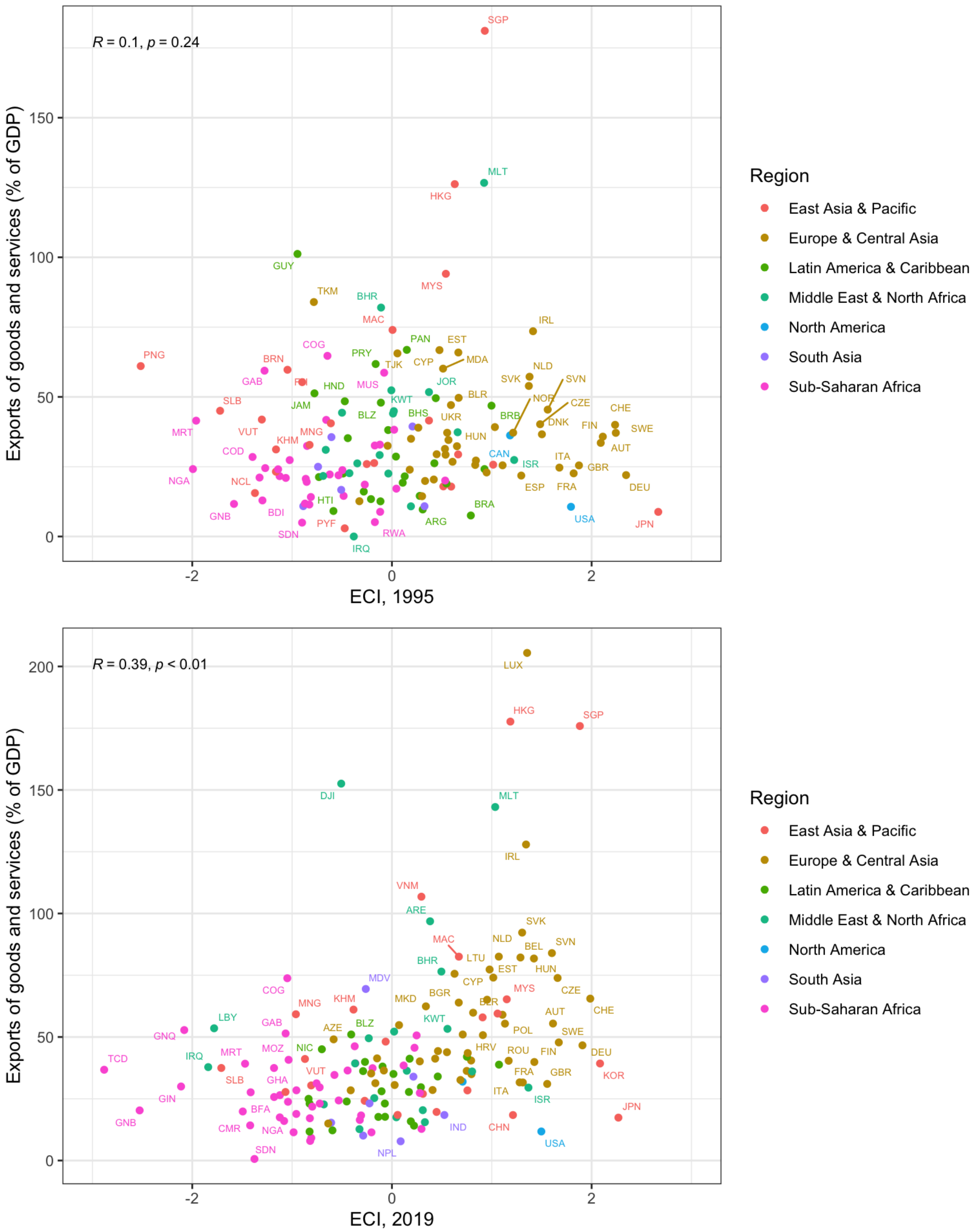
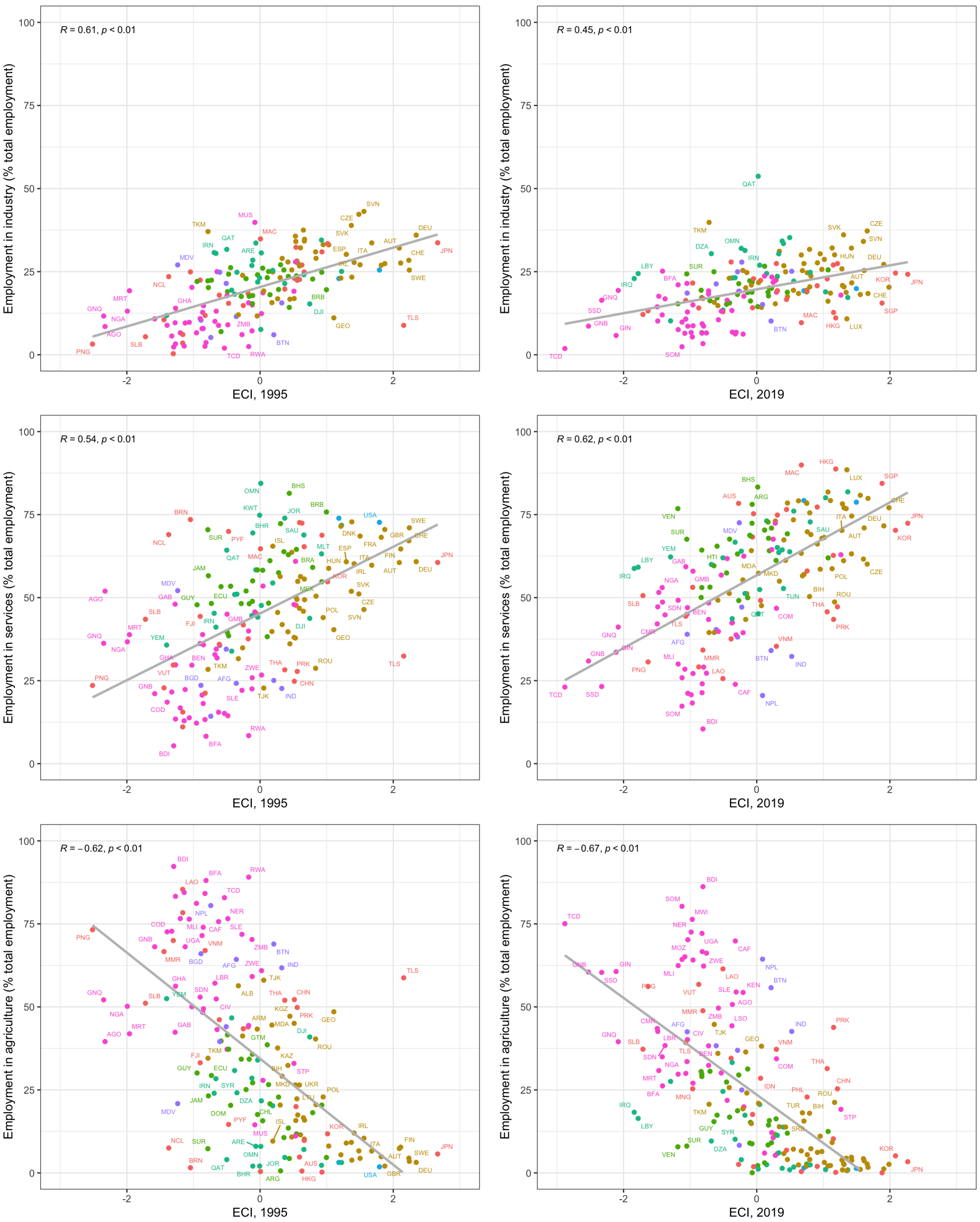


Figure 1.8: Scatter plots of ECI and employment across macro-sectors, 1995 and 2019



1.6 Discussion

While it has been well over a decade since the ECI was first introduced, there are several drawbacks that persist, regarding its conceptualisation, data, methodology and empirical applications. At the same time, there are concerns over the novelty and usefulness of this concept – in particular, whether it helps us unveil anything we did not already know regarding regional economic development and reach meaningful and actionable policy implications. This section addresses, on a first instance, the current drawbacks in the economic complexity literature, followed by its contribution to existing research.

1.6.1 Current drawbacks of economic complexity

This subsection outlines and discusses the drawbacks in the economic complexity literature, organised along three main areas: theory, data and methodology, and empirical applications.

Theory and conceptualisation

There are pressing issues in economic complexity regarding the lack of clear and exact definitions of the term and the disagreements between different interpretations and mathematical methods proposed. While it appears that the different methods generate measures that are highly correlated with each other and lead to similar conclusions, the lack of an unequivocal definition and interpretation presents risks as the term can easily be used to portray whatever a researcher or user may want, leading to misguided analyses and policy conclusions. This is particularly pressing given that the Atlas of Economic Complexity and similar platforms are being promoted as tools that can be widely used everywhere across the world.

Moreover, the literature overlooks important theoretical considerations that are likely to impact, on the one hand, what economic complexity captures and, on the other hand, how we interpret it. The first relates to blindly pooling countries across dramatically different income and development levels, often with very limited consideration of socioeconomic contexts – a top-down approach that risks overlooking important singularities. The second relates to the lack of discussion regarding the role of intermediate trade and global value chains that characterise today's production processes (and how this is reflected on the data used and the final measure calculated).

Beyond this, several theoretical ramifications of this concept have not been analysed extensively – for example, the use of terms such as 'knowledge' and 'capabilities' without clearly relating it to existing definitions of the terms surrounding them. There is still a somewhat superficial understanding of what exactly the ECI is capturing and what is meant by these terms in this context, as well as what their implications may be for countries – beyond the often-cited empirical correlation with future economic growth.

When moving from national to regional context, these concerns are even more justified, as the concept was initially developed for countries. Nevertheless, theoretical considerations at the sub-national level have been more extensive, particularly in applications relying on patent data and the idea of 'technological complexity'. Moreover, these applications tend to use networks covering locations that are more comparable – for instance, metropolitan areas within the US, regions within Europe, or even comparing regions simply to the rest of the country to which they belong, allowing for more meaningful comparisons than when considering countries across the globe with vastly different historical, geographical and socioeconomic contexts.

Data and methodology

Beyond the issues concerning theory, there are important practical-level drawbacks of the ECI method. The first drawback relates to the type of data used. The original proponents justified their focus on exports due to the wide availability of comparable data across world countries. Nevertheless, this has three major limitations – firstly, it does not account for domestic market production; secondly, it disregards services, as they do not go through customs and thus are not recorded in the same way as products[17]; thirdly, it does not capture the non-tradable sector (Hausmann et al., 2014b); and, fourthly, on a more conceptual level, trade is impacted by political economy issues and fluctuations in demand for different types of products or resources.

We find the first two limitations to be particularly critical. Disregarding services and domestic production of goods that are not exported may explain some of the discrepancies found in the data. While the authors argue that a country's inability to export these goods may indicate low productivity or quality (and thus signify lower capabilities or knowledge), the applicability of this statement is likely to differ across countries, depending on size and location. For instance, in large or geographically remote countries, domestic production of goods and services is likely to play an important role in the economy and may differ significantly in composition from exports. This is the case for Australia, a service-oriented economy whose exports are focused on low complexity products, such as minerals and

[17] There has been a recent effort to include some service data in the Atlas for Economic Complexity website. Nevertheless, the data includes only five service categories, presenting a much less granular level than for the case of products. Thus, services remain largely disregarded from ECI measures using export data.

metals. Despite a high income level, Australia ranks poorly in terms of the ECI, appearing behind countries such as Bahrain, South Africa or Cyprus (see Figure 1.3). The omission of services is likely to become an increasingly pressing issue, as services are central to knowledge-based economies, and there is increasingly a blurring of manufacturing and services that cannot be ignored if we are to fully understand the knowledge and capabilities embedded in a country or region.

Moving to the sub-national level, the use of export data can be further questioned. On the one hand, it may be justified from a theoretical perspective. For instance, Kemeny and Storper (2015) defend that a region's income level is strongly influenced by specialisation and trade – firstly, output level is influenced by the tradable sector, as demand is not limited by the region's income and, secondly, the terms of trade (i.e., the relative prices of the region's output compared with the prices of their imports) are set by the tradable sector – thus focusing on exports may capture these dynamics. On the other hand, it poses important drawbacks as the previous limitations still apply and may be even stronger. One region within a large country can be export-oriented, while another one specialises in activities for the domestic market, without this necessarily implying that the first one has more 'capabilities' than the second. Brazil provides a good example of this – the states of São Paulo and Santa Catarina both have high income levels, but while São Paulo has the highest complexity level in the country, Santa Catarina has among the lowest.[18] Finally, incorporating important aspects such as inter-regional linkages in economic complexity analyses becomes impossible since international trade does not capture linkages between regions within the same national borders.

There are alternatives to the use of export data in regional contexts. Employment data (mostly based on industry classifications) and patent data are used in sub-national analyses to calculate complexity indicators using the method of reflections (as will be described in Section 4.3). Nevertheless, while researchers have argued that patents can reflect different technological capabilities across regions (e.g., Balland and Rigby, 2017), there is still a bias in patenting towards certain industries, in particular manufacturing, and regions strongly specialised in advanced services may appear to be lagging behind competitively even when that is not the case. Similarly, for the case of employment data, it was not established or thoroughly discussed whether capabilities can be accurately inferred from industry or occupation data, and empirical applications have abstracted from discussing this, as will be further explored in Chapter 2.

[18] This statement is supported by sub-national economic complexity data derived from Brazilian government sources, available at: http://dataviva.info/en/

Recent research by Stojkoski et al. (2023) relies on patents and research publications to measure economic complexity across world countries. Nevertheless, we argue that this is problematic in a context where countries with different historical backgrounds, languages and cultures are considered together. In particular, what this type of data can capture might be completely different across countries due to varying propensities to patent technologies or publish academic research across different parts of the world.

A second drawback rests on the choice of the two-mode network used to calculate economic complexity and on how the RCA is calculated. It remains unclear whether RCA calculations should be based on larger networks, comparing a region to the rest of the world, as is often the case when using export data, or on smaller networks, comparing a region to the rest of the country to which it belongs, as is done for employment data. As described by De Benedictis and Tamberi (2001), the RCA is calculated for a particular aggregation of products or sectors, and relative to a particular benchmark of countries or regions.[19] French (2017) goes further to argue that the most appropriate point of reference for RCA calculations depends on the purpose. The network and points of reference used will therefore lead to substantially different conclusions (and, ultimately, different interpretations of the ECI). Theoretically, it is important to consider what we want to capture. In the case of a large country with high regional disparities, it may be more sensible to compare regions to the rest of the country (or, for instance, the wider macro-region), than to the rest of world. For a lagging region, adopting technologies and developing capabilities that are already present within the country might be enough to generate economic growth and for the region to continue on its own development trajectory thereafter, and is perhaps more insightful than starting from a point of comparison with the rest of world.

Beyond this, there are questions over the legitimacy of relying on spatial distributions of economic activity in order to measure economic complexity or underlying capabilities. The implications of this are, of course, very different at the national and sub-national levels. As discussed, some recent applications have tried to move away from the reliance on spatial distribution, and this issue will be further explored in Chapter 3.

Finally, it is important to note that these data limitations are likely to explain some of

[19]The lack of comparability of measures of revealed comparative advantage across different countries and other interpretations issues propelled researchers to develop normalised RCA measures (e.g., Proudman and Redding, 1998; Laursen, 2000) and alternative revealed comparative measures (such as the Gravity-Based Index or Bilateral Additive Index, see French, 2017); these transformations bring other important limitations and thus De Benedictis and Tamberi (2001) argue that relying on the traditional RCA measures is preferred, while French (2017) argue that the alternative indices are theoretically appropriate for different tasks. The ECI literature has abstracted from any discussions at this level; the main robustness check in empirical analyses is using alternative cut-offs when calculating the binary matrix.

the inconsistencies identified in our initial analysis. As mentioned, some of the countries that appear to rank too high in terms of economic complexity compared to the rest of their macro region or their income level are tax havens, and this is probably driven from the use of export data and how that data is recorded. Furthermore, the ECI appears to be particularly flawed at capturing relative capabilities in certain parts of the world. For instance, within Africa, it seems to be less accurately reflecting the complexity and income level of countries that rely on natural resources, as Nigeria and Angola rank below Somalia or São Tomé and Príncipe – while natural resources should not be considered 'complex' products, it seems countries that export them are prejudiced and rank below other countries that we may expect to have a lower complexity level (e.g., war-struck, very poor countries or tax havens). In contrast, in Europe and Asia, rankings appear to make more sense intuitively. Still, Norway presents another inconsistent example – it has a relatively low economic complexity level compared to what we would expect from their income and development levels, which appears to be driven from their exports of oil and fish, both of which are associated with low complexity.

Empirical applications

Lastly, there are challenges with empirical applications, in particular the lack of a causal link between initial economic complexity and subsequent economic growth. Most papers avoid discussions regarding the strength of the empirical methods used and thus this literature is often not seen as robust by other groups of scholars. While there have been some recent attempts using System GMM and instrumental variable approaches to argue for causal links, these papers are still not entirely convincing. As economic complexity indicators are increasingly used in country and regional development reports, there is a need to either establish a more robust relationship between economic complexity and income growth or to recognise that this should be simply understood as a descriptive tool.

There is also limited evidence of what drives economic complexity, and what exactly differs between countries or regions at different economic complexity levels. Beyond productive structures, it is important to understand what kind of prerequisites may be needed, such as institutional quality, and what kind of policies may be helpful to achieve higher economic complexity – for instance, is it simply a matter of introducing industrial policy targeting more complex goods or industries that may be within reach of that location given their current production structure (e.g., with the help of relatedness measures), or are educational policies also an effective way through which countries have managed to increase their economic complexity level in the past.

While several country reports focus on industrial strategy, there is no clear empirical evidence that shows that is necessarily the most objective and efficient way to increase economic complexity levels. Instead, there may be other target areas, such as institutional capacity and educational attainment, that could be more promising areas of action. From our exploratory analysis, it appears that these variables are very highly correlated with the ECI, particularly in contexts where a certain threshold has not yet been achieved.

1.6.2 ECI's contribution to regional development literature

In addition to the drawbacks outlined, as evidenced throughout the theoretical grounding discussions, the ideas adopted to introduce the term and methods of economic complexity are not new. Similarly, the empirical findings are not ground-breaking. For instance, the importance of knowledge and capabilities has long been recognised, and it has been clear for several decades that some activities are only found in the largest metropolitan areas, while other more mundane ones are present across all cities or regions. This begs the question of what is new about economic complexity and what this top-down approach has to offer to our understanding of development, particularly in regional contexts.

We argue that the main contribution of the ECI is as a new methodology for quantifying differences across regions, which allows for the identification of capabilities through the lens of the activities that a country or region is able to perform competitively. In doing so, it is based on outcomes, rather than on *a priori* assignment of the 'difficulty' that some activities appear to involve compared to others. As a result, the method does not require pinpointing what capabilities are – they might include many things beyond tacit or codified knowledge, including human capital, organisational skills, institutional contexts, availability of services, among many others that affect a location's capacity to perform an activity competitively. Furthermore, it allows for the identification of productive outcomes, rather than relying on the assumption of a direct link between innovation inputs, such as R&D spending or physical and social infrastructure availability, and innovation or knowledge outputs, that often occurs in the literature.

Still, by capturing 'everything', it might overlook what really matters. In particular, this literature has failed to explain the ways in which economic complexity is connected to economic growth and how it translates into development and prosperity – precisely a shortcoming that has been emphasised by Kogler (2017) for the case of relatedness. Furthermore, researchers have only recently started considering negative path dependencies and the competition for scarce resources – both emphasised by Hassink (2005) as important aspects in regional contexts – which suggests that attempting to continuously increase complexity levels across all regions is not a sustainable goal.

Regarding policy implications, economic complexity offers a ranking of countries or regions, shedding light on how different places fare in relation to others. However, a better understanding of the links between economic complexity and economic growth would be beneficial for the development of policy actions that go beyond what we already knew, or the aspiration towards ever-increasing economic complexity levels (which, moreover, is a moving target). Finally, because capabilities are not explicitly identified, understanding the fundamental differences between places ranked at different levels is challenging, and it is not evident what must change 'qualitatively' in order for a place to become 'more complex'.

All in all, while the ECI provides a promising and improved way of quantifying differences across regions, much work is still needed in terms of improving our understanding of how it works, what it captures in regional settings, how it relates to our vast existing knowledge about regional development, as well as what drives different economic complexity levels across regions. We identify these aspects as the main priorities in this literature area, and as those that have been most overlooked in existing work.

1.7 Conclusion

This chapter presented a comprehensive review of economic complexity, starting from how the original proponents portray the term, how it is grounded in the existing theory, both in terms of national and regional contexts, moving to the methodology, empirical applications and exploratory data analysis, and finalising with a discussion of what we identified as the most pressing challenges, as well as the contribution that the ECI method and concept make to existing research.

This represents the status quo regarding economic complexity and our understanding of this concept, and helps guide the remaining chapters in this thesis. The main priority is to further understand whether this index can be applied across different geographical levels and development contexts. To do this, we will move from the cross-country national level to the sub-national one, starting with a country within the European context in Chapter 2, where we will also explore the applicability of the methods to occupation data.

The next step will be to further explore the link between economic complexity and exports, with a particular focus on countries that are heavily reliant on natural resources. To that end, Chapter 3 will return to cross-country analysis, and explore a different context of oil-dependent countries, with a focus on the Gulf Cooperation Council region. The aim is to explore the usefulness and relevance of the economic complexity methods in this context and the challenges that arise when investigating countries with particular export bases, further challenging the universal applicability of this model.

Lastly, Chapter 4 will turn to linking economic complexity with education and economic growth, with the aim of understanding whether economic complexity can indeed better explain future economic growth than simpler measures capturing human capital across countries.

1.A Appendix

Table 1.A.1: Description and source of variables

Variable	**Definition**	**Source**
ECI	Economic Complexity Index based on HS-92 classification. Own calculations.	Observatory of Economic Complexity
GDP per capita	GDP per capita, PPP (constant 2017 international $)	World Bank Open Data
Natural resource rents	Total natural resources rents (% of GDP)	World Bank Open Data
Exports (% GDP)	Exports of goods and services (% of GDP)	World Bank Open Data
Imports (% GDP)	Imports of goods and services (% of GDP)	World Bank Open Data
Merchandise trade (share GDP)	Merchandise trade (% of GDP)	World Bank Open Data
High-tech exports (% exp)	High-technology exports (% of manufactured exports)	World Bank Open Data
Agriculture VA (% GDP)	Agriculture, forestry, and fishing, value added (% of GDP)	World Bank Open Data
Manufacturing VA (% GDP)	Manufacturing, value added (% of GDP)	World Bank Open Data
Employment share in agriculture	Employment in agriculture (% total employment)	World Bank Open Data
Employment share in industry	Employment in industry (% total employment)	World Bank Open Data
Employment share in services	Employment in services (% total employment)	World Bank Open Data
Secondary school enrolment	School enrolment, secondary (% gross)	World Bank Open Data
Tertiary school enrolment	School enrolment, tertiary (% gross)	World Bank Open Data
FDI inflows (% GDP)	Foreign direct investment, net inflows (% of GDP))	World Bank Open Data
R&D expenditure	Research and development expenditure (% of GDP)	World Bank Open Data
R&D researchers	Researchers in R&D (per million people)	World Bank Open Data
Unemployment	Unemployment, total (% of total labour force)	World Bank Open Data
Population	Total population (regardless of legal status/citizenship)	World Bank Open Data
Urban population	People living in urban areas as defined by national statistical offices	World Bank Open Data
Urban primacy	Population in the largest city (% of urban population)	World Bank Open Data
Income share top 10%	Income share held by highest 10%	World Bank Open Data
Income share bottom 10%	Income share held by lowest 10%	World Bank Open Data
Gini	Gini index	World Bank Open Data
Poverty headcount ratio (1.90 USD line)	Poverty headcount ratio at $1.90 a day (2011 PPP) (% of population)	World Bank Open Data
Poverty headcount ratio (national line)	Poverty headcount ratio at national poverty lines (% of population)	World Bank Open Data
Firms market capitalisation	Market capitalization of listed domestic companies (% of GDP)	World Bank Open Data
Time to start business	Time required to complete procedures to legally operate a business (days)	World Bank Open Data

Table 1.A.1 (continued): Description and source of variables

Variable	**Definition**	**Source**
Human Capital Index (HCI)	Measure of the productivity as a future worker of a child born today relative to the benchmark of full health and complete education (ranges from 0 to 1).	World Bank Open Data
Human Development Index (HDI)	Summary measure of average achievement in key dimensions of human development (long and healthy life, being knowledgeable, decent standard of living)	United Nations Human Development Reports
Voice & Accountability	Perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media	World Bank Worldwide Governance Indicators
Political stability	Perceptions of the likelihood of political instability and/or politically motivated violence, including terrorism	World Bank Worldwide Governance Indicators
Government effectiveness	Perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies	World Bank Worldwide Governance Indicators
Regulatory quality	Perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development	World Bank Worldwide Governance Indicators
Rule of law	Perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence	World Bank Worldwide Governance Indicators
Corruption control	Perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as 'capture' of the state by elites and private interests	World Bank Worldwide Governance Indicators

Table 1.A.2: Descriptive statistics, 1995 to 2019, 179 countries

Variables	N	Mean	SD	Min	Max
ECI	4403	0.00	1.00	-3.87	2.71
GDP per capita (constant intl. $, PPP)	4125	18339	20716	469	161971
Natural resources rents (% GDP)	4211	7.60	11.41	0.00	87.46
Exports (% GDP)	4006	40.31	28.78	0.01	228.99
Imports (% GDP)	4006	44.50	26.09	0.02	221.01
Merchandise trade (% GDP)	4250	64.90	40.91	7.81	419.96
High-tech exports (% exp)	1701	10.39	11.30	0.00	68.14
Agriculture VA (% GDP)	4101	12.82	12.32	0.03	79.04
Manufacturing VA (% GDP)	3888	13.03	6.41	0.23	49.88
Employment share in agriculture	4378	29.24	24.18	0.03	92.37
Employment share in industry	4378	19.70	8.57	0.32	59.58
Employment share in services	4378	51.06	18.68	5.36	89.94
Secondary school enrolment	2957	78.49	30.75	5.28	163.94
Tertiary school enrolment	2812	35.58	26.82	0.20	148.53
FDI inflows (% GDP)	4225	5.19	15.68	-57.60	449.08
R&D expenditure (% GDP)	1922	0.94	0.94	0.01	4.94
R&D researchers (per million people)	1500	1988	1904	5.91	8066
Unemployment	4378	8.06	6.18	0.10	38.80
Population (total)	4403	37829319	136713279	131679	1407745000
Urban population	4378	19095754	59644890	33908	848982855
Urban primacy	3733	33.24	17.89	2.97	100.00
Income share top 10%	1494	29.85	6.78	19.50	61.50
Income share bottom 10%	1493	2.52	0.97	0.10	4.80
Gini index	1497	38.20	8.89	23.00	65.80
Poverty headcount ratio (1.90 USD line)	1498	8.88	16.25	0.00	94.30
Poverty headcount ratio (national line)	893	24.09	14.03	0.60	76.80
Firms market capitalisation (% GDP)	1702	64.60	108.67	0.01	1350
Time to start business (days)	2694	34.00	49.54	0.50	697.00
Human Capital Index	406	0.57	0.15	0.29	0.89
Human Development Index	4105	0.67	0.17	0.24	0.96
Voice and accountability	3676	-0.13	0.99	-2.31	1.80
Political stability	3674	-0.15	0.98	-3.31	1.76
Government effectiveness	3670	-0.06	1.01	-2.48	2.44
Regulatory quality	3671	-0.05	1.00	-2.65	2.26
Rule of law	3676	-0.12	1.01	-2.61	2.13
Corruption control	3673	-0.09	1.01	-1.87	2.47

Chapter 2

Economic complexity and employment in Portuguese municipalities

2.1 Introduction

As described, the ECI was introduced by Hidalgo and Hausmann (2009) a decade ago. Departing from a network connecting countries to the products they export competitively, economic complexity aimed to capture differing levels of 'capabilities' across countries, reflecting the skills, technology, institutions, or anything else that made them able to export certain products competitively. Since then, there has been a significant increase in the literature employing this concept and method. The ECI has been applied across the world, at different geographical levels, relying on different data – including exports, patents and industries – often with limited consideration of social, cultural and development contexts.

In the case of Europe in particular, the concept of economic complexity, often coupled with relatedness, is increasingly regarded as relevant for smart specialisation strategies aimed at determining optimal growth paths in regions (McCann & Ortega-Argilés, 2015; Boschma, 2017). However, despite increased policy interest, the applicability of economic complexity to different contexts, as well as its usefulness in terms of explaining economic growth and development across regions, are still unclear. While relatedness has been the focus of more extensive investigation, existing results are not unequivocal – some papers find that European regions target activities that show intermediate levels of technological relatedness to their existing specialisation, while others find that many regions do not choose paths that are related to their current area of specialisation (Marrocu et al., 2020; Hidalgo, 2021) – and therefore more scrutiny is needed for the case of economic complexity.

Furthermore, many existing applications to regional European contexts either draw on patent data and consider regions across several countries, often disregarding peripheral regions that have very few or no patents, or focus on countries with very different industrial or development contexts (for example, the UK, Italy and Sweden). As will be further discussed, while the use of occupation data is likely to overcome important drawbacks in existing contributions, more work is needed in conceptualising and understanding occupation-based ECI. In particular, the limited number of papers using occupations (e.g., Mealy et al., 2018a; Hane-Weijman et al., 2022) do not discuss the interpretation or implications of occupation-based ECI measures.

This chapter is aimed at advancing the literature towards filling these gaps. By applying this methodology to Portugal, a 'small peripheral' country, at the sub-national level, we explore the meaning and interpretation of the ECI based on occupations and investigate whether it bears any association with employment growth. In particular, we apply the ECI method to occupation data in a European country context, which has only been done for the case of Sweden (Hane-Weijman et al., 2022) for a shorter period (2002-2012). Portugal presents a good empirical context, as it is an under-explored European example, structurally very different from existing applications, it presents significant regional heterogeneity and has comprehensive data.

Furthermore, we bridge the literature on economic complexity with that on occupations and the task-based approach to labour markets, introduced by Autor et al. (2003) and Acemoglu and Autor (2011), to understand whether and how economic complexity relates to this well-established framework and our existing knowledge. Finally, we compare occupation-based and industry-based ECI measures, to understand how each relates to regional outcomes and whether or not they capture similar dynamics. Thus, the aim is to contribute to our understanding of economic complexity applications with occupation data, and explore conceptually the applicability of the ECI at the sub-national level.

To measure economic complexity, we follow the network-based methodology by Hidalgo and Hausmann (2009) and Mealy et al. (2018a), and start from a network linking municipalities to occupations, derive diversity and ubiquity measures, and get a ranking and level of complexity across municipalities and occupations. Focusing on 308 Portuguese municipalities, for the period from 1985 to 2019, we look at the association between our economic complexity measures and regional employment growth. Following this, we explore how occupation complexity relates to the task-based approach to labour markets, as well as to an alternative industry-based indicator of complexity.

The regional disparity patterns observed in terms of economic complexity are similar to those seen with regional Gross Value Added (GVA), with municipalities in coastal areas showing higher economic complexity than in inland ones. The empirical analysis shows a negative association between both economic complexity measures and employment growth, but this association disappears once we control for initial employment level. Our discussion of the results suggests that, in this context, the occupation and industry-based ECI measures may be capturing dynamics related to conventional specialisation patterns too strongly, rather than reflecting the underlying 'capabilities' within a municipality.

The remainder of this chapter is organised as follows. Section 2.2 provides a literature review, focusing on economic complexity applications in European contexts and on bridging the literature on economic complexity with that on occupations and the task-based approach to labour markets, and concludes with our research questions and aims. Section 2.3 provides the research context. Section 2.4 outlines our data and methodology, followed by results and further analysis in Sections 2.5 and 2.6 respectively. Finally, Section 2.7 provides a discussion and Section 2.8 concludes.

2.2 Literature review

Economic complexity and regional economic growth

Despite being introduced as a new method, the ideas underpinning the concept of economic complexity are linked with many theoretical contributions that have been around for several decades. Summarising the broader discussion on economic complexity and regional economic growth, the main idea is that regions that have more complex activities – e.g., products, technologies, industries or occupations – have higher underlying 'capabilities' and knowledge (in particular tacit knowledge), which provide a source of competitive advantage (Lawson & Lorenz, 1999; Asheim & Gertler, 2005) and thus might lead to higher economic growth than that experienced by regions with lower economic complexity levels.

From the *method of reflections* introduced by Hidalgo and Hausmann (2009), two variables are calculated – the ECI, originally developed for the case of countries based on exports, and defined broadly as capturing the capabilities of countries, and the PCI, calculated analogously to the ECI, originally for the case of products, to broadly represent how 'complicated' a product is, proxied by the capabilities it requires.[1] Capabilities are defined in very broad terms, and can essentially encompass anything that makes a country able to export a specific product competitively – the underlying idea being that more complex

[1] The name given to the variables changes across papers in accordance with the network and data used. The ECI is sometimes referred to as Knowledge Complexity Index (KCI) and the PCI as Technological Complexity Index (TCI) for patents, or as Occupation Complexity Index (OCI) for the case of occupations.

products will only be successfully exported by a limited number of countries, which in turn are diverse and able to make other complex products.

While the ECI and PCI are calculated in analogous ways, each focused on one of the sides of the two-mode network used, and thus have similar definitions, the implications behind each of them are different and it is important to distinguish between the two. On the one hand, the ECI relates to countries or regions, providing a ranking that places countries or regions with more similar exports (or patents, industrial concentration, and so on) closer together in the ordering and those with more dissimilar activities further apart (Mealy et al., 2018b, 2019). In regional applications, the papers that focus directly on the ECI tend to depart from the importance of tacit knowledge and localised learning that give some regions advantages over others in their analysis and theoretical grounding. On the other hand, the PCI relates to the products (in the case of export data), technology (for patent data), and industries or occupations (when employment data is used). Some regional applications focus on the PCI component instead, looking at regional entry into new technologies or industries (e.g. Balland et al., 2019; Bishop and Mateos-Garcia, 2019) and grounding their research in topics such as regional diversification, regional branching, related variety and relatedness.

Economic complexity applications to European regions

Focusing on the European context, the concept of economic complexity, often coupled with relatedness[2] measures, is increasingly applied in papers investigating smart specialisation strategies aimed at determining optimal growth paths in regions (McCann & Ortega-Argilés, 2015; Boschma, 2017). As a result, just over the last few years several empirical applications emerged in the literature applying economic complexity methods to the European context. Table 2.A.1 in the Appendix lists all economic complexity empirical applications in European contexts to date, including both papers that focus on EU regions and others that study individual countries at the sub-national level.

A lot of the papers that apply economic complexity methods to all EU regions rely on patent data (Balland et al., 2019; Rigby et al., 2019; Antonelli et al., 2020; Mewes and Broekel, 2022; Pintar and Scherngell, 2022), one draws on scientific publications (Heimeriks et al., 2019), while recent papers by Deegan et al. (2021) and Pinheiro et al. (2022) use employment data, focusing on industries. These papers are concerned mostly with technology or

[2]Relatedness refers to the idea that some activities require similar capabilities and thus are seen as 'related' to each other. There are several different measures of relatedness, including technological relatedness (Breschi et al., 2003), product relatedness (Hidalgo et al., 2007), and skill relatedness (Neffke & Henning, 2013). Different methods are used, including analyses of co-occurrence, co-location, worker flows and skill requirements. For a comprehensive review we refer to the work of Boschma (2017).

priority economic domains, and several of them focus on linking economic complexity with smart specialisation frameworks and on understanding regional paths. In particular, they investigate whether regions moved towards more complex economic activities (and if so, which ones), and also whether those paths led to better regional outcomes.

In contrast, papers that analyse individual countries tend to rely on exports (Basile et al., 2019), employment industries (Bishop & Mateos-Garcia, 2019; Mealy & Coyle, 2022) and the methodology has been expanded to occupation data too (Broekel et al., 2021; Hane-Weijman et al., 2022). More specifically, Hane-Weijman et al. (2022) are the only ones that rely on the ECI method as introduced in Hidalgo and Hausmann (2009) when using occupation data, while Broekel et al. (2021) and Deegan et al. (2021) employ alternative measures of occupation-based economic complexity, introduced in more detail below.

Regarding outcomes of interest, there are two broad groups – first, papers that look at entry and exit of technological or scientific fields; second, papers that look at regional outcomes, such as growth in GDP, employment and labour productivity. Of the second group, some papers focus on the regional complexity metric (ECI), whereas others focus on the other part of the network (PCI or equivalent – reflecting the complexity of products, technologies or industries), and look at whether entry and exit into more complex activities had an impact on regional outcomes.

With the exception of the paper by Rigby et al. (2019), which covers the time period from 1981 to 2019, the papers focused on individual countries cover relatively short periods. Classification changes, as well as data availability, are often blamed and, although the former complicates the analysis, it is a challenge that we aim to overcome in our analysis.

While empirical applications to European regions have mostly relied on patent data, this presents an important drawback as patents are very skewed towards specific economic activities, and also overlook several regions for which the number of patents is too limited to be included in analyses. This is a pressing drawback, as it is often the more peripheral regions, which are of crucial relevance for smart specialisation policies, that are left out. Thus, our research is aimed at addressing this drawback and investigating further whether economic complexity, measured through occupation data, can help us understand regional development paths across Europe.

The case for using occupations in economic complexity measures

While most applications have relied on exports, patents and employment industry data, there are important limitations with each of them. Firstly, export data overlooks the service sector, which played an active role in the most recent process of structural change (Antonelli, 1998), and is increasingly crucial in the world economy and trade. Moreover, services have become increasingly blurred with goods, hindering the distinction and measurement of interactions between them, particularly given the importance of immaterial aspects of goods, and through the standardisation, mass production and trade of services (Evangelista, 2000). In this regard, Pilat and Wölfl (2005) show that the amount of service sector value-added that is embodied in manufactured goods has increased over time and that a growing share of workers employed in the manufacturing sector is performing service-related occupations. Moreover, employment in services has grown and contribute to a considerable share of aggregate labour productivity growth (Castaldi, 2009). Besides, export data does not capture domestic market production.

Secondly, patent data is very biased towards certain sectors and industries, once again largely disregarding services, and their value is variable and highly skewed (Griliches, 1998). Moreover, some regions are completely excluded from analyses because of the very low or non-existing number of patents (e.g., Pintar and Scherngell, 2022 focus strictly on metropolitan regions and have to exclude those that have less than 100 patents in the period studied) – this is particularly pressing in the Portuguese context but also a wider issue in several other European countries and regions.

Thirdly, while employment industry data allows researchers to overcome some of the aforementioned challenges, it covers only the industries in which workers are employed and overlooks the types of job they do. From the urban economics literature, we know there has been a trend towards functional (rather than sectoral) specialisation in cities and increasing separation of company locations by function (Duranton & Puga, 2002). Furthermore, as described by V. Henderson (2010), there is a hierarchy of city sizes, mediated by the relative advantages and costs of cities of varying dimensions. In particular, standardised manufacturing activities benefit mostly from agglomeration in their own industry and thus tend to gravitate towards small specialised cities where economies of scale will be maximised relative to urban size diseconomies (e.g., higher costs, congestion); in contrast, high technology industries, business services and other activities where the overall level of local agglomeration also contributes to productivity gravitate towards larger and more diverse cities (V. Henderson, 2010). Papers investigating functional specialisation and trade have turned to occupations in order to overcome the lack of direct mapping between industries and functions – for example, in order to characterise both direct and indirect value added of exports – by tracing what type of workers (characterised by occupation) are involved in production (Timmer et al., 2019). This provides some evidence that occupations can reflect more accurately the kinds of economic activity taking place in a specific location, and lead to more relevant ECI measures. Moreover, in the context of global value chains, occupations are likely to reflect more closely the types of capabilities present within a region, likely better reflecting the types of jobs workers are doing and the skills they might need than simply the industry in which they work or the resulting exported goods.

Since occupation data is likely to overcome several of these drawbacks, it is pertinent to further investigate whether the ECI measured through occupations is associated with regional outcomes of interest, as well as to shed light on the interpretation and meaning of economic complexity methods applied to this type of data. While occupations may capture more accurately a region's profile in terms of economic activity and functions performed, the question of whether occupation-based complexity measures can be meaningful and useful remains unclear.

While, at the time of writing, there are two existing contributions that rely on occupation data for economic complexity calculations using the method of reflections – Mealy et al. (2018a) for the US and Hane-Weijman et al. (2022) for the case of Sweden – they do not provide an answer to this question and there are important aspects that require further analysis. In particular, neither of the papers discuss the meaning and applicability of the ECI methodology or provide an interpretation beyond the mathematical description of the methods. Furthermore, there is no discussion over the ideal network size or geographic level to use in sub-national level applications.

Linking economic complexity and the task-based approach to labour markets

Focusing on occupation data also provides the opportunity to link the economic complexity literature with the task-based approach to labour markets and technological change. The task-based approach provides a strong case for the importance of tasks, frequently proxied by occupations, for output produced within a region (Autor, 2013), and puts forward a categorisation of task content of different occupations – along two overlapping axes capturing, firstly, whether they involve routine or non-routine tasks and, secondly, predominantly cognitive or manual tasks (Autor et al., 2003; Acemoglu & Autor, 2011).

As defined by Acemoglu and Autor (2011), routine tasks are those that can be performed by following explicit and well-defined rules; they can be cognitive (e.g., bookkeeping), or manual (e.g., repetitive assembly). In contrast, non-routine tasks involve problem-solving and complex communication activities. Non-routine cognitive tasks tend to involve flexibility, creativity, generalised problem-solving, and complex communications (e.g., managing others), whereas non-routine manual tasks require physical flexibility and adaptability, visual recognition, or non-scripted communications (e.g., truck driving).

Making a link with this literature can help us understand to what extent, if any, occupation complexity indicators are aligned with existing knowledge on this topic. From a conceptual perspective, non-routine task intensive occupations, which are known to be less codifiable, require more complex communication, and thus confer higher competitive advantages to regions, are expected to be more complex than routine task intensive occupations. Similarly, cognitive tasks, which require more conscious intellectual activity, are expected to be more complex than manual ones. We therefore expect non-routine cognitive task intensive occupations to have the highest ECI levels and routine manual task intensive ones to show the lowest ECI levels. However, between the other categories, namely routine cognitive and non-routine manual task intensive occupations, the theoretical predictions are not clear cut, and will be further discussed.

There are a few existing papers that look at occupation complexity from different lenses and using alternative methodologies. Caines et al. (2017) depart from the task-based approach view of occupations as bundles of tasks, and measure occupation complexity from the complexity of tasks involved. They define complex tasks as those that "involve higher-order skills such as the ability to abstract, solve problems, make decisions or communicate effectively" (Caines et al., 2017, p. 299) – a definition that is similar to that of non-routine occupations. Using Occupational Information Network (O*NET) occupation descriptors and Principal Component Analysis they create a continuous normalised measure of occupations and compare mean wages, and wage and employment growth between simple and complex occupations in the US from 1980 to 2005. They find a strong positive association between task complexity and both wage levels and growth. Another important finding is that some occupations, for example related to crafted goods, as well as middle-skill occupations in finance and insurance, are found to be complex, despite being classified as routine task intensive occupations (Caines et al., 2017).

In a separate route, Lo Turco and Maggioni (2020) also rely on O*NET job descriptors and requirements, but on a different approach that relates directly with Hidalgo and Hausmann's (2009) product complexity. The authors look at the average complexity of products within three-digit level industries, which they compare with the industries' skill and knowledge work requirements (derived from O*NET information) through rank correlations. They create an indicator based on the average normalised scores of those work requirements for which the rank correlation was highest – namely, Physics, Engineering and Technology, Computer Electronics and Mathematics for knowledge items, and Science, Mathematics and Critical Thinking for skill items – as a measure of the average occupational complexity within an industry (Lo Turco & Maggioni, 2020). They find a strong positive association between the occupation complexity indicator and growth in GDP per capita across US metropolitan statistical areas for the 2001-2017 period.

Two applications to the European context have relied on similar methods. Broekel et al. (2021) look at diversity, relatedness and complexity of occupations and assess their contribution towards explaining growth in Norwegian industry-regions observations over 2009-2014. To measure occupation complexity, they follow approaches similar to the work by Ederer et al. (2015), Caines et al. (2017) and Lo Turco and Maggioni (2020), and find that these measures are not particularly strong in the case of Norway – in their main model, only industrial relatedness has a statistically significant positive effect on employment growth. Importantly, occupation complexity did not show any statistically significant effect on employment growth in any of the models (Broekel et al., 2021). The authors argue this may be due to the specificity of the Norwegian case and its focus on oil and gas, and seafood industries. In parallel, looking at European NUTS-2 regions, Deegan et al. (2021) rely on the method by Caines et al. (2017) to measure occupational skills complexity, which they use to estimate the weighted average complexity for each industry based on their occupational composition (thus their analysis is mostly focused on industries).

By providing different interpretations and measures of occupation complexity, these papers provide a good benchmark of comparison for our economic complexity calculations using occupation data, which will be further discussed in Section 2.7.

Research aims and contribution

In light of the literature and challenges discussed, this chapter answers two broad sets of questions. The first set relates to understanding occupation-based economic complexity measures – in particular, what is meant by occupation complexity, what it does (and does not) capture, as well as comparing and contrasting it with the task-based approach to labour markets. The second set of questions relates to linking economic complexity with regional outcomes – more specifically, investigating whether economic complexity is associated with employment growth across Portuguese municipalities and whether occupation-based ECI measures do indeed matter more than industry-based ones, as suggested above.

As described, following the original intuition behind the ECI as introduced in the literature, the idea is that certain municipalities with higher economic complexity levels have more underlying ‘capabilities’ or knowledge, and thus are expected to experience higher employment growth levels over time. The question of what the indicators capture when applied to occupations remains, and it might not necessarily reflect ‘complexity’ in the conventional sense of the word.

With the ECI methodology, beyond the complexity of occupations, whether this relates to ‘rarity’ or to how ‘complicated’ they are to perform, we are also capturing the extent to which having a certain type of occupational mix – and in particular, a mix that is relatively more diverse and encompassing occupations that are present in a more limited number of regions – tells us something about a region’s characteristics that is relevant for their development paths.

As a result, this chapter contributes not only with an application of economic complexity methods to a new country and dataset, representing a different context and country size, but it also explores further occupation-based complexity measures, and bridges the literature on economic complexity with that on occupations and labour markets. Furthermore, it covers a longer time period than most existing papers, from 1985 to 2019, giving us the opportunity to look at the evolution of economic complexity over time and investigating whether the regional dynamics we are capturing have changed over these three decades.

2.3 Research context: Portuguese regions

There are several characteristics that make Portugal an ideal context to carry out our analysis. Portugal represents an under-explored context of a ‘small peripheral’ country, with significant heterogeneity and inequality across regions (Rego et al., 2021). It is a particularly relevant country from a Smart Specialisation perspective, as a lot of its regions are targeted by such policies, and it is often excluded from papers that use patent data in economic complexity measurements across all EU regions. It also offers a comprehensive dataset, that allows us not only to investigate a significant time period, covering 35 years, but also to draw a direct comparison between occupation- and industry-based measures.

The period of analysis goes back to 1985, just as Portugal was about to join the EU in 1986. Although it experienced GDP growth above the European average during the two initial decades, it still lagged behind most western European countries in terms of GDP, wages, skills and capital even before the 2008 financial crisis hit. Fonseca et al. (2018a) also highlight important characteristics that differentiate Portugal from other countries, even within a European context – in particular, Portugal has lower wages, GDP per capita, capital stock and share of the service sector and it also lagged behind the rest of Europe and even other Southern European countries in terms of its share of college graduates. Despite these characteristics, it experienced job polarisation from the mid-1990s onward (Fonseca et al., 2018a).

Because of good data availability and these characteristics that make it an interesting case study, there are several applications of the dataset we use – in particular papers that investigate regional wage differentials (Pereira & Galego, 2011; Galego & Pereira, 2014; Pereira & Galego, 2015) and others that employ the task-based approach to labour markets and look at job polarisation and technological change (Fonseca et al., 2018a, 2018b).

Despite receiving financial support through the EU's Cohesion Policy for several years, and the increasing recognition of the importance of innovation, Portugal still lags behind its European counterparts, ranking below EU averages in terms of R&D expenditure, patent applications, and employment in high-tech manufacturing industries or human resources in science and technology.[3] There are four key factors that contribute to this. First, a lack of interaction between regional innovation actors, which hinders the capacity to form true regional innovation systems (Santos, 2000; Cooke, 2001; Santos & Simões, 2014; Faria et al., 2020). In particular, as further argued by Santos and Simões (2014), there is a lack of cooperation, with firms acting in individualist ways, a predominance of workers with low qualification levels, as well as low density and quality of innovation infrastructure and erratic innovation policy. Second, despite the presence of a large number of research institutions, the process for technology transfer to industry is lacking and there is a mismatch between knowledge production (predominantly done in universities) and the production or economic sphere. Santos and Simões (2014) argue that this situation is hard to change due to the large majority of SMEs being led by individuals with basic education levels and thus the current innovation actors and infrastructure simply fall too far from their needs and expectations. Third, innovation systems are highly unbalanced, with high concentration in Lisbon. Mapping the location of R&D institutions in Portugal, Santos and Simões (2014) show that they are mostly concentrated in the more developed and densely populated regions, in particular Lisbon and the North and Centro regions (due to the strong university presence in Porto and Minho in the North, and Coimbra and Aveiro in the Centro region). Finally, there is a general predominance of a restrictive and basic notion of innovation among firms in Portugal, focused on very incremental changes, mostly around modernisation and improvement of production processes, rather than innovation per se (Santos, 2000; Santos & Simões, 2014).

[3]For instance, see the 2023 Country Report by the European Commission, available to download here: https://economy-finance.ec.europa.eu/publications/2023-european-semester-country-reports_en.

In a more recent evaluation of smart specialisation strategies, Cooke (2016) found that the Centro and Norte regions were better able to exploit innovation opportunities that ensued from existing R&D infrastructure than the Algarve region, which has difficulty in reaching a common diversification goal and pursuing it, mostly due to disagreement between the national and regional level and its existing specialisation surrounding tourism. This not only reflects the challenges described in the previous paragraph, but also a broader issue in institutions and governance when it comes to innovation, which is further described by Laranja et al. (2020), who highlight the difficulty in making priority choices and a lack of national-regional alignment.

To provide further context of how Portuguese regions evolved over time, Figure 2.1 plots GDP per capita across NUTS-3 regions, showing an increase in income in the vast majority of Portuguese regions over the 1985-2019 period. There was some stagnation and decline in the period after the 2008 financial crisis, but most regions appear to have started experiencing recovery since then.

Figure 2.1: GDP per capita, NUTS-3 regions, 1985 to 2019

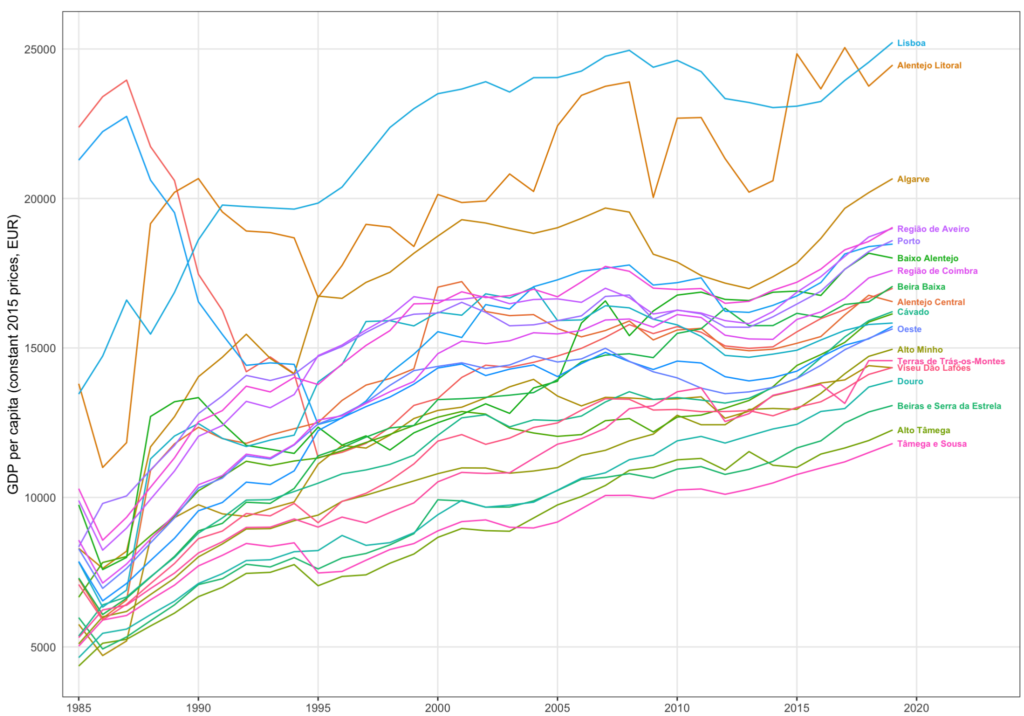


Figure 2.2 shows the employment share in industry, services and agriculture from 1985 to 2018. Regarding the share of employment in industry, Ave, Tâmega e Sousa and Região de Aveiro are the regions where it plays the biggest role. These three NUTS-3 regions

are part of the North and Centre regions that traditionally had an important industrial presence in Portugal; nevertheless, they experienced a decline in the relative importance of employment in industry during this period. As observed across the wider European context, Portugal experienced considerable growth in the share of employment in services throughout this time period. Services play a particularly important role not only across the two biggest metropolitan regions of Lisbon and Porto, but also in Algarve, where tourism plays a crucial role in the economy. Finally, Trás-os-Montes, Douro and Alto Tâmega have the highest shares of employment in agriculture, despite having experienced a sharp decline over this time period.[4]

To shed light on regional disparity patterns in Portugal, Figure 2.3 maps GDP per capita across NUTS-3 regions in the first and final years of the time period. Rather than the north-south divide often seen in other EU countries, Portuguese regions are traditionally split between the more prosperous coastal regions and the poorer and less accessible inland regions, a pattern evident in both 1985 and 2019. Lastly, Figure 2.4 moves from the NUTS-3 to the municipality geographical level and shows GVA per worker in 2019 only (due to limited data availability). As expected, the main patterns are similar, but this provides a clearer picture of the heterogeneity across the country, even within NUTS-3 regions, and allows for some outliers to be identified.

[4]Data was downloaded from ARDECO, the Annual Regional Database of the European Commission's Directorate General for Regional and Urban Policy (available to download: https://knowledge4policy.ec.europa.eu/territorial/ardeco-database). The focus here is on the NUTS-3 regional level due to the lack of data availability at the municipality level, particularly prior to 2010. The plots on employment share only cover the period up to 2018 as this is the latest available year.

Figure 2.2: Employment share in industry, services and agriculture, NUTS-3 regions, 1985 to 2018

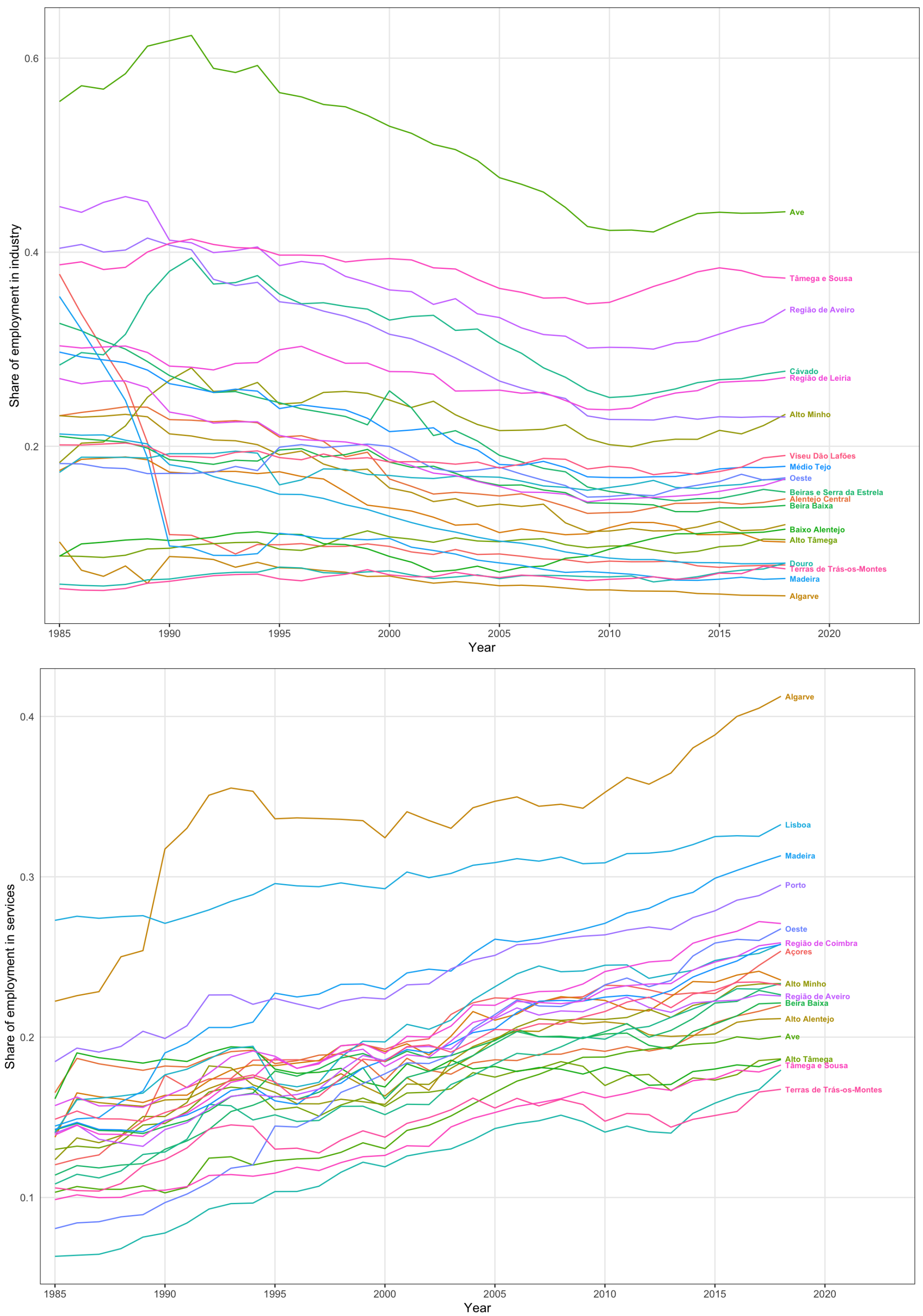


Figure 2.2 (continued): Employment share in industry, services and agriculture, NUTS-3 regions, 1985 to 2018

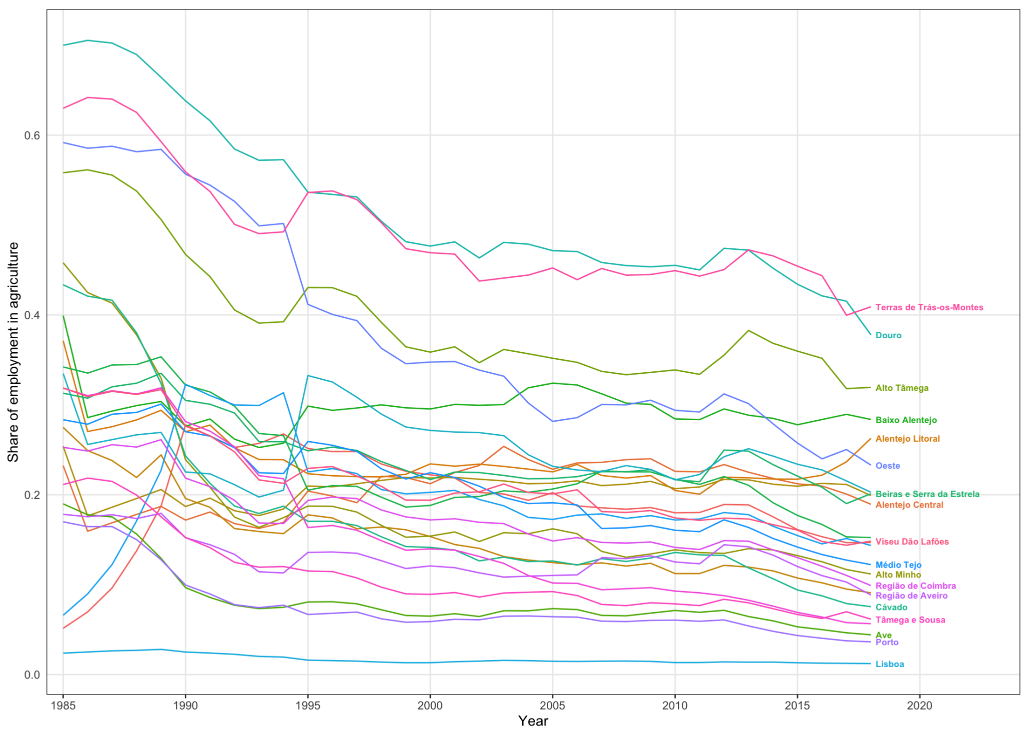


Figure 2.3: Map of NUTS-3 regions, GDP per capita, 1985 and 2019

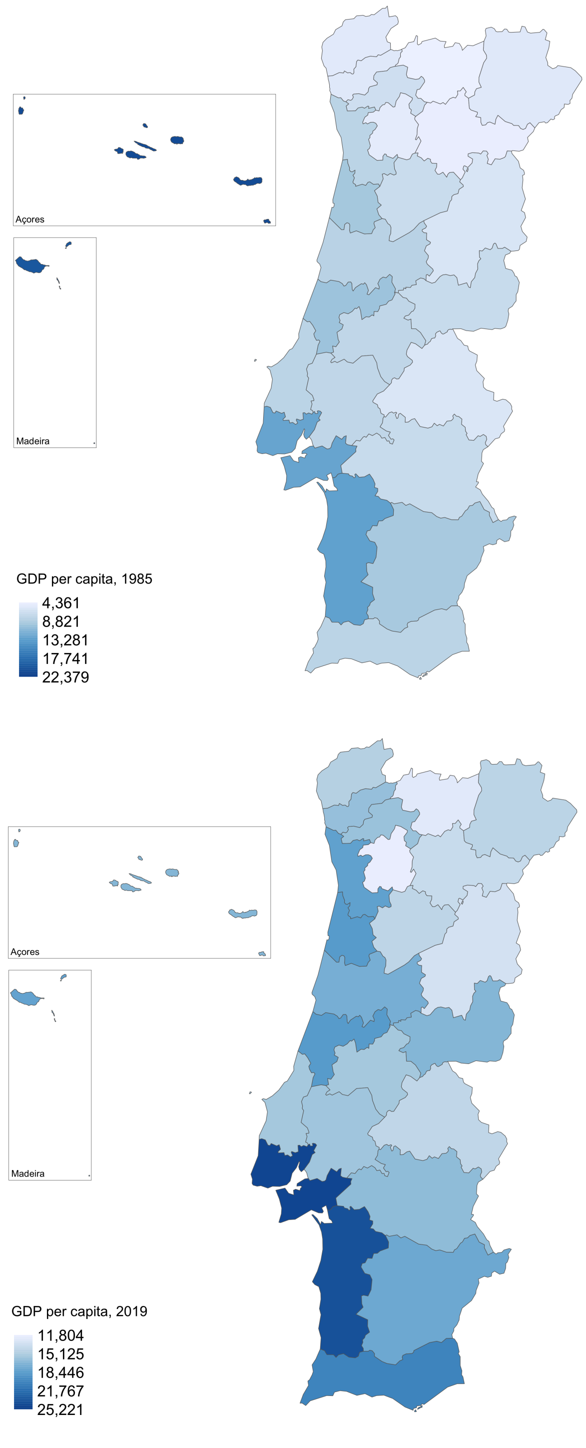
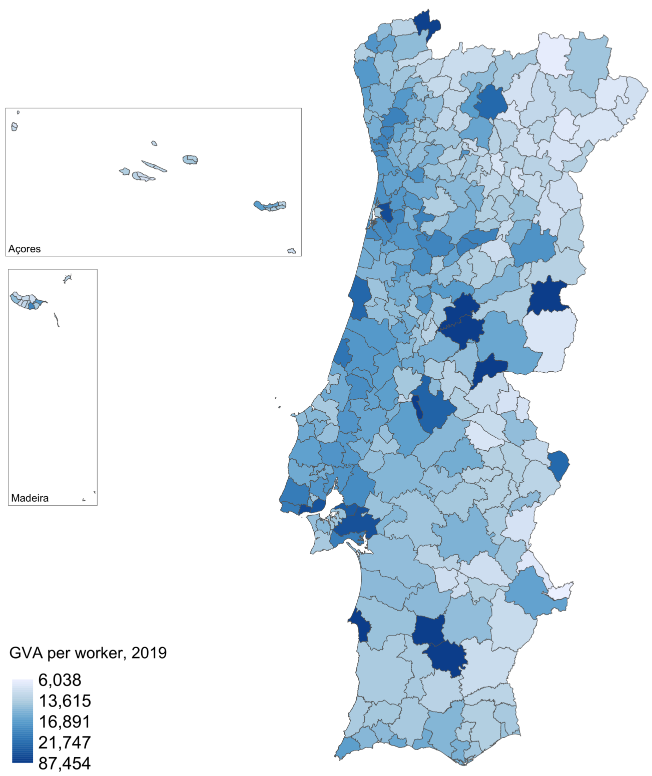


Figure 2.4: Map of municipalities, GVA per worker, 2019



2.4 Data and methods

2.4.1 Data

Our main data source is a Portuguese employer-employee matched dataset, *Quadros de Pessoal*, which consists of a compulsory survey of all firms, conducted annually for the purpose of monitoring compliance with labour regulations by the Portuguese Ministry for Labour, Solidarity and Social Security. The dataset covers the period from 1985 to 2019.[5] For each year, it provides detailed information on all firms, establishments and employees. For establishments and firms, this includes features such as size, location, industry classification and total employees. For employees, it includes information on education, skills, occupation, tenure, monthly wages and hours worked, among others.

The data covers all wage earners, with the exception of civil servants and self-employed workers (as customary with this type of dataset), as these workers are not covered by the same labour regulations and thus are not covered in this survey. In contrast, all firms that employ at least one person need to comply with labour laws and are required to fill out this survey annually, thus it represents a full picture of firms and formal employment in Portugal, and coverage is not biased towards large firms or multinationals.

In order to collect an accurate representation of occupation concentration across municipalities, we restrict the sample of workers, following previous papers (e.g. Fonseca et al., 2018a) and the official publications by the Ministry for Labour, Solidarity and Social Security. In particular, we consider only full-time workers aged 16 to 65.[6] To collect the total number of workers per occupation and industry for each municipality, we link the employees to their establishment, providing us with their work location.[7] To respect anonymity requirements, as well as avoid noise in our calculations, for each of the years, we consider only municipality and occupation combinations that have at least three workers, otherwise it is considered to be zero.

For both industries and occupations, the analysis will be carried out using three-digit level classifications.[8] For each municipality, we built a dataset of the total number of workers for

[5]With the exception of 1990 and 2001, as the survey was not conducted those years.

[6]We define full time workers as those with a full time work contract, who worked at least 120 hours and received their base wage rate in full in the reference month of October.

[7]We used the employee data alongside establishments (rather than firms) to ensure that the geographical place of work we are assigning to the employee is the one where they conduct their job, rather than the headquarters of the firm they work for.

[8]Data was collected at both three- and five-digit levels of occupation and industry classifications. While the three-digit level represents a good level of analysis, five-digit is the most appropriate level for carrying out conversions across different classifications, and is the level at which crosswalks are provided for the

each occupation and industry. Municipality-level data was also collected for the purpose of capturing differences across municipalities, including the share of workers who attained post-secondary education, as well as variables related to the dimension of the municipality (total number of workers and total number of establishments).

Due to the long time period covered, there are two broad challenges in carrying out our analysis. The first relates to changes in activity classifications that have occurred over time. Over the time period, there are three different occupation classifications used – CNP-80 (based on ISCO-68) for years 1985 to 1994, CNP-94 (based on ISCO-88) for 1995 to 2009, and CPP-10 (equivalent to ISCO-08) for data from 2010 onward. Similarly, the dataset cuts across four different industry classifications – CAE-73 (based on NACE 1970) for years 1985 to 1994, CAE-Rev.2 (equivalent to NACE-Rev.1) for 1995 to 2002, CAE-Rev.2.1 (NACE-Rev.1.1) for 2003 to 2006, and CAE-Rev.3 (fully harmonised with NACE-Rev.2) from 2007 onward. Such changes are likely to influence complexity calculations (as it will change significantly the matrix composition and dimensions, and LQ calculations, meaning that the economic complexity values are not directly comparable across the different time periods). Some papers avoid dealing with this issue altogether, by reverting to using shorter time periods (e.g., Hane-Weijman et al., 2022).

For the case of occupations, we look at the periods 1985-1994 and 1995-2019 separately, having converted all the data from 1995 onward to ISCO-08 using the official crosswalk provided by INE (the Portuguese Institute for Statistics). Given the different nature of the older occupation classifications, which followed a logic more similar to the one used in the case of industries, we do not convert the data based on ISCO-68, and instead look at that period separately. For the case of industries, we focus solely on the period from 2007 to 2019, based on NACE-Rev.2 (or the Portuguese CAE-Rev.3). The several Portuguese industry classifications covered prior to this were not fully harmonised with international NACE classifications, and rather 'adapted to the Portuguese context'. Earlier years are therefore complex to analyse, likely not comparable across time, fragmented over several time periods and conversions impossible to carry out methodically.

The second challenge, although far less severe, relates to changes in geographic classifications. There were changes to the numbers and composition of municipalities over the time period. Since municipality is the most detailed geographic unit available for the earlier years there is no way to readjust our data accordingly. More specifically, the number of

earlier Portuguese versions of classifications. We therefore converted the data using the five-digit level, ensuring the three-digit level was fully consistent when there were issues related to merging or separation of occupations.

municipalities is 304 from 1985 to 1993, goes up to 305 until 1998 and up to 308 until present. Furthermore, there were changes in civil parishes and municipalities in 2013, with the aim of reducing the number of civil parishes to reduce administrative costs. For the vast majority of municipalities, this only involved merging of parishes, without changes to the size and composition of the municipality. Nevertheless, in four municipalities, this led to a change in the composition of the municipality, in each case with just one parish changing.[9] Importantly, looking at the data collected, these changes do not represent a change in total number of workers or establishments within these municipalities that is significantly different from the changes occurring in other years.

As will be further described, the economic complexity measures are calculated for the municipality-occupation and municipality-industry two-mode networks, and the main empirical analysis will be carried out at this level. Additional variables were downloaded for the year 2019, the latest period in our analysis (due to the very limited availability of additional data at municipality-level and short time periods). In particular, GVA per worker (total and across different macro-sectors) will be used to investigate any composition effects in our complexity measures. Table 2.A.2 provides the definition and source, and Table 2.A.3 provides summary statistics for the economic complexity measures calculated and the other variables used in our analysis (see Appendix).

2.4.2 Measuring economic complexity

To measure occupation-based economic complexity metrics, we used the methodology introduced by Hidalgo and Hausmann (2009), which has been adapted by Balland and Rigby (2017), and previously applied to occupation data by Mealy et al. (2018a) and Hane-Weijman et al. (2022). The measures were calculated using the EconGeo R package developed by Balland (2017), using the method of reflections.

We construct a binary matrix W of municipal occupational concentration for each period, representing a municipality-occupation network. The network dimensions vary slightly across the two periods, in accordance with the occupation and municipality classifications. More specifically, there are 304 municipalities from 1985 to 1993, 305 municipalities from 1994 to 1998 and 308 municipalities from 1999 to 2019. In terms of occupations, there are 268 for the period from 1985 to 1994 (CNP-80) and 123 three-digit level occupations for the 1995-2019 period (CPP-10) – thus, the economic complexity values should not be directly

[9] Golegã and Santarém had a change in one parish, which passed from Santarém to Golegã; Lisboa and Loures experienced changes with the creation of one parish in Lisboa that involved a small part of territory that was formerly part of Loures.

compared across these two separate periods involving differing occupation classifications, but the municipality rankings can be broadly compared.

Each element W_{mi} relates to a municipality m's location quotient (LQ) in occupation i (where E is the total number of people employed):

$$LQ_{mi} = \frac{E_{mi}/\sum_i E_{mi}}{\sum_m E_{mi}/\sum_m \sum_i E_{mi}} \tag{2.1}$$

$$W_{mi} = \left\{ \begin{array}{ll} 1, & LQmi \geq 1 \\ 0, & LQmi < 1 \end{array} \right\} \tag{2.2}$$

Complexity measures for municipalities and for occupations are calculated from this matrix, using the *method of reflections*. This involves sequentially combining two components:

$$Diversity = K_{m,0} = \sum_i W_{mi} \tag{2.3}$$

$$Ubiquity = K_{i,0} = \sum_m W_{mi} \tag{2.4}$$

Firstly, the diversity of municipalities is given by the number of occupations for which each municipality has $LQ > 1$ and is calculated by summing over the rows of W. Secondly, the ubiquity of occupations (i.e., how common those occupations are across municipalities), is given by the number of municipalities that exhibit $LQ > 1$ in a particular occupation and is calculated by summing over the columns of W.

From there, we calculate the Occupation-based Economic Complexity Index (OECI) for municipalities and the Occupation Complexity Index (OCI) for occupations by correcting the information carried by each of these variables using the other over a series of n iterations. The index is re-scaled from 0 (minimum relative complexity) to 100 (maximum relative complexity).

$$OECI = K_{m,n} = \frac{1}{k_{m,0}} \sum_i W_{mi} * K_{i,n-1} \tag{2.5}$$

$$OCI = K_{i,n} = \frac{1}{K_{i,0}} \sum_m W_{mi} * k_{m,n-1} \tag{2.6}$$

The first iterations of this method are simple to define. In the first iteration, for $n = 1$, the result represents the average ubiquity of the occupations in which a municipality m has $LQ > 1$ for OECI calculations, and the average diversity of municipalities that have $LQ > 1$ in occupation i for the OCI. In the second iteration, for $n = 2$, the result for the OECI represents the average diversity of other municipalities that have an occupation mix similar to municipality m, whereas for the OCI it captures the average ubiquity of occupations developed in municipalities that have $LQ > 1$ in occupation i. As we continue, each additional iteration becomes increasingly harder to define in simple terms; nevertheless, the *method of reflections* provides increasingly accurate measures of the complexity of municipalities and occupations, as noise and size effects are corrected for. Eventually, the ranking of municipalities and occupations remains stable from one step to another and additional iterations do not provide any new information; in our case, we used the default of 20 iterations (we also calculated ECI with fewer iterations and the conclusions did not change).

For the case of industries, the same methodology is followed, using municipal industrial concentration instead. As mentioned, calculations cover the period from 2007 to 2019 only, and the network links 260 industries to 308 municipalities.[10] The metrics derived consist of the Industry-based Economic Complexity Index (IECI), relating to municipalities, and the Industry Complexity Index (ICI) for industries.

2.4.3 Empirical analysis

Beyond analysing our economic complexity calculations, we want to investigate whether the measures show any association with regional paths. Thus, we look at the association between the occupation-based OECI and employment growth in Portuguese municipalities. The baseline regression estimations will be as follows:

$$\Delta y_{it} = \alpha_i + \beta_1 OECI_{it-1} + \beta_j \boldsymbol{X}_{j_{it-1}} + \eta_t + \epsilon_{it} \tag{2.7}$$

Where Δy_{it} is the employment growth rate percentage in municipality i between $t - 1$ and t. $OECI_{it}$ is the occupation-based complexity measure, our independent variable of interest. $\boldsymbol{X}_{j_{it}}$ represents the control variables – the total number of workers (to capture initial employment and proxy for population size) and the share of workers with post-

[10]Four three-digit level industries that involve civil sector activities (related to public administration and defence) were dropped from the analysis given the limited coverage of these industries in the Quadros de Pessoal data.

secondary education (to capture broad education levels), and the natural log is used in both cases. α_i, η_t and ϵ_{it} represent region fixed effects, year fixed effects, and the error term, respectively. We run the regressions for the 308 Portuguese municipalities, focusing mostly on the 1995-2019 period for the occupations-based measure, and on 2007-2019 for the industry-based ECI. In all models, robust standard errors are reported, clustered at the municipality level.

The existing literature follows different specifications when looking at regional outcomes – some papers focus on entry and exit into different economic activities (e.g., Balland et al., 2019), while others focus on employment growth outcomes but consider the ECI of the activities where a region experienced entry or exit (e.g., Rigby et al., 2019; Hane-Weijman et al., 2022). In this chapter, we follow a simple specification to capture whether the overall complexity level of a municipality can explain employment growth, following more closely the contributions by Mewes and Broekel (2022) and Pintar and Scherngell (2022). There is a question of whether to focus on yearly changes or longer periods – Mewes and Broekel (2022) lag regional characteristics by one year following existing literature (while, for patents, they lag by 3 or 5 years); given the focus on occupation and industry data, we lag variables by one year only.

Following this, we turn to the task-based approach to labour markets, and draw on that theory and method to develop the task content of occupations over time (following, for example, Autor et al., 2003), in order to understand how the occupation complexity measures derived using the *method of reflections* compare to existing task-based measures and what, if anything, they can add in terms of our understanding of occupations and regional labour markets.

Lastly, we carry out a similar analysis with the complexity measure constructed with industrial concentration, IECI, in order to investigate whether occupations, which reflect the types of jobs workers are performing within a region, or industries, can help us better understand regional outcomes over time.

2.5 Results

2.5.1 Economic complexity calculations

Occupations – OCI calculation

Figure 2.5 plots the average OCI for each one-digit occupation group, from 1985 to 2019, split in two between the different occupation classifications.[11] Across both panels, the 'Professional' occupation group appears to be the most complex one consistently throughout the years; in the second panel, this group is closely followed or even overtaken by 'legislators, senior officials and managers' and 'technicians and associate professionals' in some years. In contrast, the agricultural and related occupations group appears consistently at the bottom, with the lowest average levels of economic complexity. This tends to be followed by the 'service workers' and the elementary occupation groups in the latest time period.

To further understand which specific occupations are classified as the most and least complex ones according to the OCI, Tables 2.1 and 2.2 show the top and bottom ten three-digit level occupations in 1985 and 2019, respectively. The occupations are ranked by complexity level, and the table also provides ubiquity – i.e., how many municipalities had $LQ > 1$ for that occupation on average – and the average diversity of the municipalities that had $LQ > 1$ in that specific occupation. The numbers are not directly comparable across the two tables and, importantly, there are a lot more occupations in the CNP-80 classification used in the 1985 calculations, thus ubiquity of the most complex occupations tends to be much lower than it is in 2019; nevertheless, we can compare the relative positions of different types of occupations in each of the years.

Occupations with the highest complexity level tend to be in the professional and managerial realms, are quite rare across regions, with low ubiquity, and are present in relatively diverse regions – roughly across both years, these include occupations such as lawyers, engineers and mathematicians, but also artists and librarians. Among the least complex occupations, are those that are more routinised and often associated with lower qualification levels in clerical, elementary or agricultural groups; they are present across many municipalities, and the average level of diversity of the municipalities in which they are present tends to be low – these include service-related occupations, such as shop assistants, waiters and other related occupations, as well as agriculture-related ones.

[11] Complexity measures were aggregated to the one-digit level, so we observe changes over time.

Focusing on Table 2.2, for 2019, we can see some important discrepancies and crucial conceptual implications. For instance, among the ten most complex occupations there are 'street and related service workers', an elementary occupation, and among the bottom ten there are 'professional service managers'. These two examples are counter-intuitive, both from a conceptual perspective, and from looking at the one-digit level groups – elementary and professional, respectively – and how the other occupations within each of these groups rank. Importantly, among the bottom ten occupations in terms of complexity, are 'personal care workers in health services', the type of occupation that may not always require qualifications but is crucial to humanity, far from easy to perform and is recurrently undervalued in society as we move towards a measurement paradigm. While the interpretation of this occupation complexity measure will be further discussed, the evidence thus far points to important questions regarding what is being captured by the OCI. In particular, it appears that the most complex occupations are those that are more geographically concentrated, whereas among the lower complexity levels are occupations that are needed across all municipalities, and thus more geographically disperse. Thus, pointing to the fact that the OCI might be simply reflecting traditional location patterns rather than underlying knowledge or 'capabilities'.

Figure 2.5: Average OCI for one-digit occupation groups, 1985 to 2019

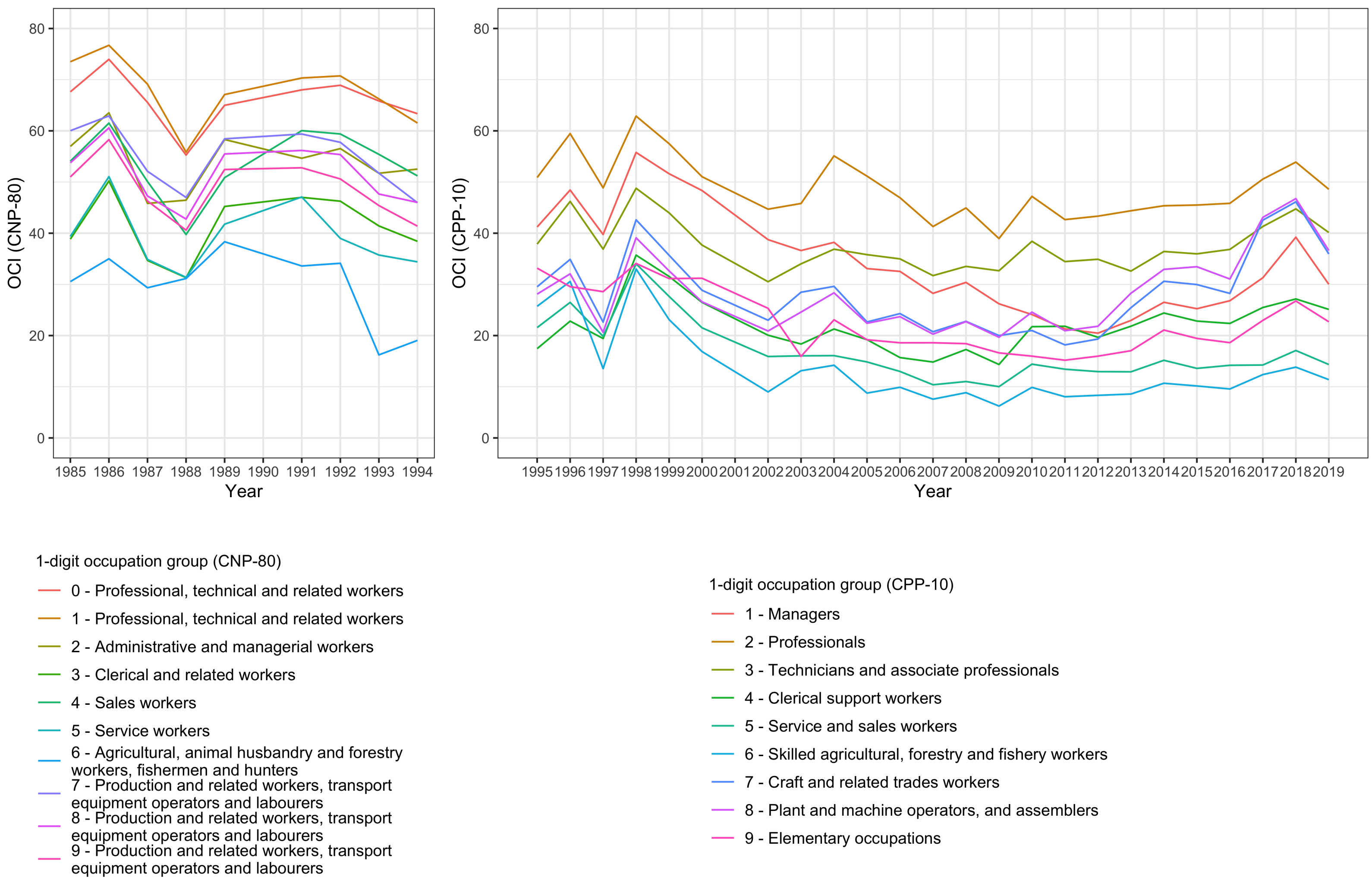


Table 2.1: OCI, ubiquity and average diversity, top and bottom 10 occupations, three-digit level (CNP-80), 1985

Rank	Occupation	OCI	Ubiquity	Avg. diversity	One-digit occupation group
1	Producers, performing arts	100	2	93	Professional, technical and related workers
2	Performing artists not elsewhere classified	100	2	93	Professional, technical and related workers
3	Ships' deck officers and pilots	94.2	1	116	Professional, technical and related workers
4	Optometrists and opticians	94.2	1	116	Professional, technical and related workers
5	Mathematicians and actuaries	94.2	1	116	Professional, technical and related workers
6	Lawyers	94.2	1	116	Professional, technical and related workers
7	Choreographers and dancers	94.2	1	116	Professional, technical and related workers
8	Postmasters	94.2	1	116	Clerical and related workers
9	Auctioneers	94.2	1	116	Sales workers
10	Aircraft engine mechanics	94.2	1	116	Production and related workers, transport equipment operators and labourers
...	...	...	...	...	...
215	Livestock workers	16.19	76	29.11	Agricultural, animal husbandry and forestry workers, fishermen and hunters
216	Motor vehicle drivers	14.77	191	27.17	Production and related workers, transport equipment operators and labourers
217	Farm machinery operators	14.07	80	27.66	Agricultural, animal husbandry and forestry workers, fishermen and hunters
218	Labourers not elsewhere classified	14.07	216	25.2	Production and related workers, transport equipment operators and labourers
219	General farm workers	11.47	135	25.34	Agricultural, animal husbandry and forestry workers, fishermen and hunters
220	Salesmen, shop assistants and demonstrators	10.15	116	29.47	Sales workers
221	Mail distribution clerks	9.64	211	23.48	Clerical and related workers
222	Telephone and telegraph operators	8.27	100	26.74	Clerical and related workers
223	Correspondence and reporting clerks	4.15	34	27.12	Clerical and related workers
224	Bookkeepers, cashiers and related workers	0	162	21.02	Clerical and related workers

Table 2.2: OCI, ubiquity and average diversity, top and bottom 10 occupations, three-digit level (CPP-10), 2019

Rank	Occupation	OCI	Ubiquity	Avg. diversity	One-digit occupation group
1	Traditional and complementary medicine associate professionals	100	6	47.83	Technicians and associate professionals
2	Mathematicians, actuaries and statisticians	99.2	6	51.17	Professionals
3	Sales, marketing and public relations professionals	90.69	10	47.4	Professionals
4	Traditional and complementary medicine professionals	82.4	7	50.71	Professionals
5	Information and communications technology service managers	81.39	19	45.53	Managers
6	Electrotechnology engineers	81.27	24	40.12	Professionals
7	Administration professionals	74.98	11	43.91	Professionals
8	Librarians, archivists and curators	73.9	18	49.28	Professionals
9	University and higher education teachers	73.65	16	46.12	Professionals
10	Street and related service workers	70.71	11	49.27	Elementary occupations
...	...	...	...	...	...
112	Professional services managers	5.25	138	34.57	Managers
113	Personal care workers in health services	5.1	210	32.86	Service and sales workers
114	Protective services workers	3.71	75	30.95	Service and sales workers
115	Domestic, hotel and office cleaners	3.24	144	32.46	Elementary occupations
116	Waiters and bartenders	2.66	120	34.3	Service and sales workers
117	Client information workers	2.4	59	33.75	Clerical support workers
118	Food preparation assistants	1.85	153	34.08	Elementary occupations
119	Cooks	0.76	143	33.11	Service and sales workers
120	Agricultural, forestry and fishery labourers	0.36	165	32.42	Elementary occupations
121	Mixed crop and animal producers	0	137	32.73	Skilled agricultural, forestry and fishery workers

Municipalities – OECI calculation

Turning to the regional component of complexity calculations, Figure 2.6 shows the OECI average across NUTS-3 regions over time.[12] There are significant oscillations in some years, though they happen across all regions, with rankings remaining largely the same. Within each of the time periods, the regions at the top remain relatively more stable, whereas regions at the bottom of the ranking appear to experience more changes in relative positions. Thus, while the OECI might meaningfully reflect regions' relative positions, these oscillations may impact our regression results when looking at changes over time within regions and linking complexity with employment outcomes.

For a visual representation of disparities in complexity levels across municipalities, Figure 2.7 maps OECI levels across municipalities. The regional disparity patterns observed here are similar to those observed with income per capita in Portugal, with coastal municipalities showing higher economic complexity than inland ones. Across both years, the figures show that higher complexity levels tend to be more concentrated among the northern coastal regions, where industry has traditionally played an important role in the economy. This geographical concentration has accentuated over time, and appears to be stronger in 2019 than in 1985. A few outliers can also be identified, with some inland municipalities showing significantly higher levels of complexity than their neighbours within the same NUTS-3 region.

To complement this, Tables 2.3 and 2.4 show the top and bottom ten municipalities ranked by OECI levels for 1985 and 2019 respectively. The tables also show the diversity of municipalities (i.e., the total number of occupations in which the municipalities have an $LQ > 1$); and average ubiquity (which tells us the average ubiquity of the occupations in which the municipalities have $LQ > 1$). While we cannot directly compare values across both tables, we can compare the rankings and the relative positions of different municipalities. In 1985, the top 10 shows several municipalities in the Porto Metropolitan Area, as well as other municipalities in the Norte NUTS-2 region, along with Região de Leiria, in the Centro Region. Among these municipalities, several have very similar levels of OECI, diversity and average ubiquity. In the case of the bottom ten municipalities in 1985, their location appears to be more spread out across the country, including the Algarve, Alentejo and Açores regions, as well as several municipalities in the Norte region, which has municipalities both right at the top and the bottom of the rankings. This is partly driven by the contrast seen within the Norte region, between the Porto Metropolitan

[12]Once again, complexity measures were aggregated and averaged for NUTS-3 regions, so that changes over time can be analysed.

Area and other industry-oriented NUTS-3 regions, such as Ave and Cávado, and the Douro region, which is more reliant on agriculture and specialises in wine production.

Turning to 2019, we can see the Lisbon and Porto Metropolitan Areas represented among the top ten municipalities. The remainder of the municipalities belong once again to the Norte and Centro NUTS-2 regions. Nevertheless, looking at the municipalities themselves, only three feature in the top ten across both years (namely, Guimarães, São João da Madeira and Vila Nova de Famalicão). Among the bottom ten municipalities, only two have remained at the lowest ranks across both time periods. Still, the broader regions represented are similar to 1985, with the Douro, Açores and Alentejo regions featured and the only change being in the absence of any municipalities in Algarve among the bottom ranks in 2019.

Lastly, we can observe the differences in diversity and average ubiquity between the most and least complex regions. As expected, having higher diversity alone does not necessarily translate into higher ECI; rather, the regions at the top of the ranking are those that are able to combine both relatively high diversity and relatively low levels of average ubiquity of occupations.

Figure 2.6: Average OECI across NUTS-3 regions, 1985 to 2019

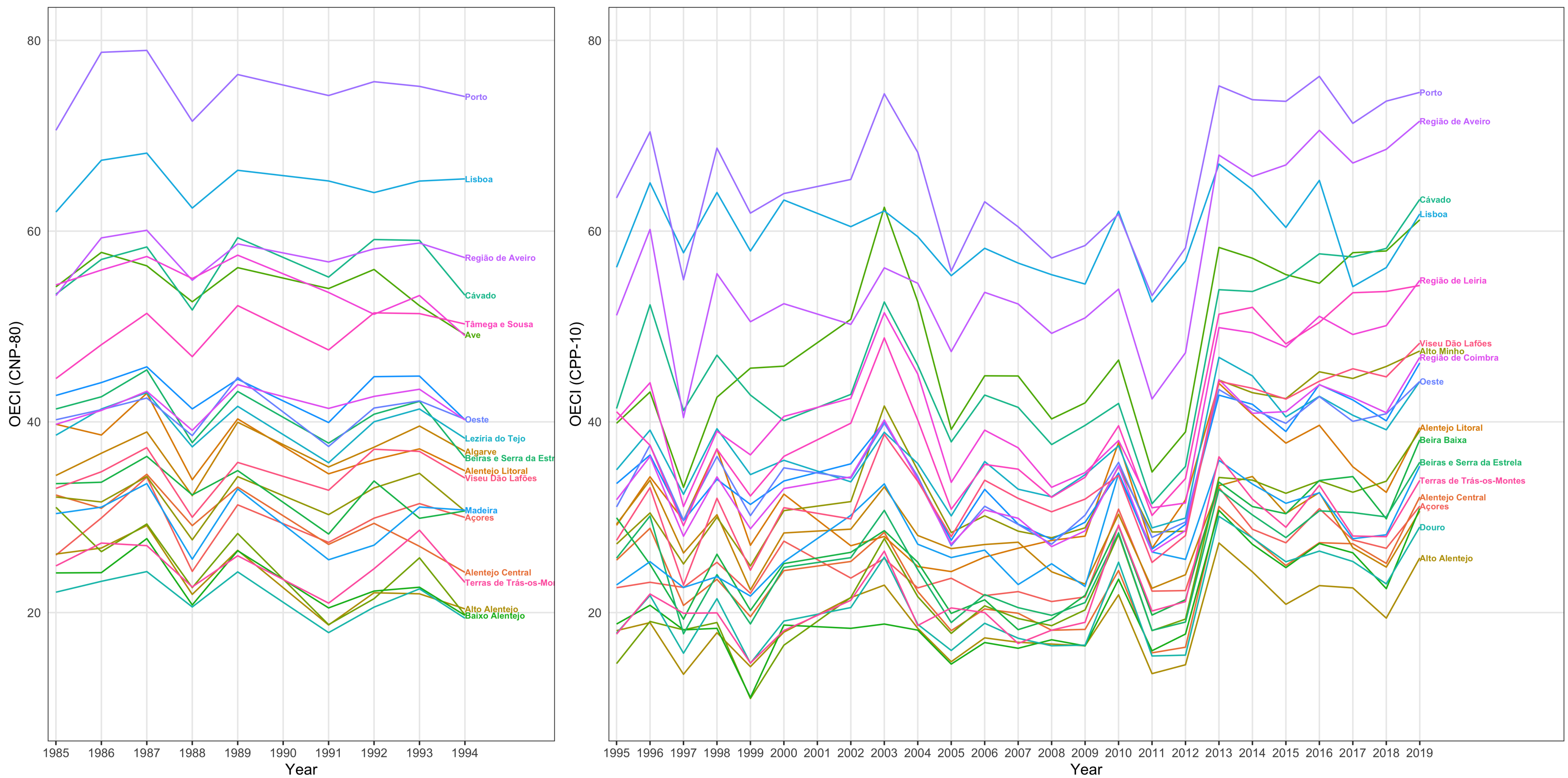


Figure 2.7: Map of OECI across municipalities, 1985 and 2019

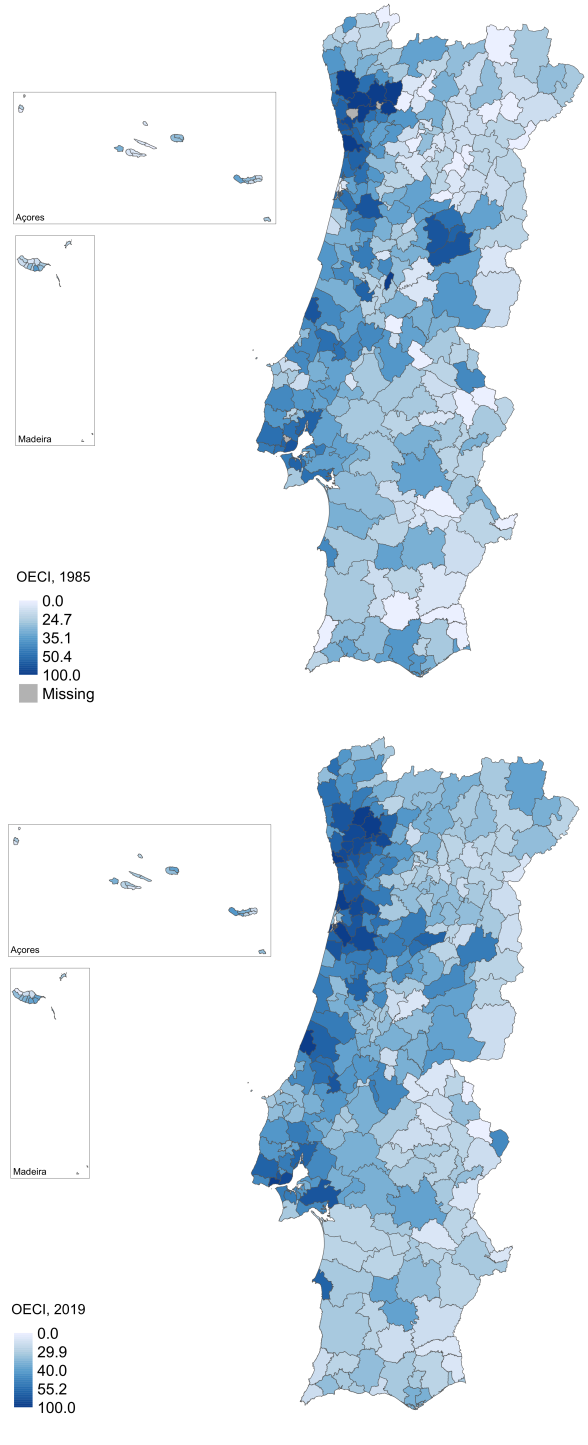


Table 2.3: OECI (CNP-80), top and bottom 10 municipalities, 1985

Rank	Municipality	OECI	Diversity	Avg. ubiquity	NUTS-3 region
1	Castanheira de Pêra	100	10	39.9	Região de Leiria
2	Guimarães	93.41	28	37	Ave
3	Santo Tirso	91.16	37	34.46	Porto (MA)
4	Felgueiras	87.66	27	40.33	Tâmega e Sousa
5	Vila Nova de Famalicão	85.86	34	37.91	Ave
6	Barcelos	85.45	25	38.52	Cávado
7	Fafe	85.16	25	49.08	Ave
8	Vila Nova de Gaia	85.05	70	37.43	Porto (MA)
9	São João da Madeira	83.24	28	40.75	Porto (MA)
10	Maia	80.39	54	45.76	Porto (MA)
...	...	...	...	...	...
295	Ribeira de Pena	8.47	5	174.2	Alto Tâmega
296	Almodôvar	7.79	6	144.5	Baixo Alentejo
297	Resende	7.41	7	133.86	Tâmega e Sousa
298	Freixo de Espada à Cinta	7.22	7	161.57	Douro
299	Castro Marim	6.83	5	139	Algarve
300	Calheta	6.83	4	144.25	Açores
301	Alcoutim	5.97	2	155	Algarve
302	Vila de Rei	4.5	3	196.33	Médio Tejo
303	Lajes das Flores	0.37	2	139	Açores
304	Penedono	0	2	186.5	Douro

Table 2.4: OECI (CPP-10), top and bottom 10 municipalities, 2019

Rank	Municipality	OECI	Diversity	Avg. ubiquity	NUTS-3 region
1	Marinha Grande	100	21	76.33	Região de Leiria
2	Aveiro	93.19	49	76.59	Região de Aveiro
3	Oeiras	92	47	55.57	Lisboa (MA)
4	Guimarães	89.16	32	84.31	Ave
5	Ovar	87.8	29	89.41	Região de Aveiro
6	Matosinhos	87.5	49	67.31	Porto (MA)
7	Braga	85.49	54	88.09	Cávado
8	São João da Madeira	84.96	29	88.76	Porto (MA)
9	Vila Nova de Famalicão	83.96	32	92.78	Ave
10	Lisboa	83.76	53	57.45	Lisboa (MA)
...	...	...	...	...	...
299	Mesão Frio	18.05	25	139.8	Douro
300	Freixo de Espada à Cinta	17.28	16	130.69	Douro
301	Lajes das Flores	16.65	10	132.1	Açores
302	Lajes do Pico	16.2	18	133.5	Açores
303	Gavião	16.04	19	149.89	Alto Alentejo
304	Portel	15.56	28	143.14	Alentejo Central
305	Marvão	14.82	21	137.14	Alto Alentejo
306	Arronches	10.96	15	139.93	Alto Alentejo
307	Porto Moniz	7.23	13	143.54	Madeira
308	Corvo	0	2	165	Açores

Municipalities – OECI and other variables

Here, we go further to explore how the OECI relates to other regional variables. Tables 2.5 to 2.7, show the correlations between the OECI and other variables used in our analysis, in 1985, 2002 and 2019 respectively, capturing the start, mid-point and end of our time period. They show a positive correlation between the OECI and diversity across all years, though it is much stronger in 1985 than in the remaining years, likely due to the different network dimensions and number of occupations. Similarly, there is a strong negative correlation between the OECI and average ubiquity, across all years. Both the post-secondary education share and the total number of workers, a proxy for broader population size, show a positive correlation with the OECI; the only exception to this is the negative correlation between the OECI and education share in 1985.

Due to the limited data available at the municipality level, we focus mostly on our latest year, 2019 and on GVA per worker across different macro-sectors. Table 2.7 includes the GVA per worker variables for the latest period analysed. The OECI shows a positive correlation with total GVA per worker, as well as GVA per worker in manufacturing and ICT services. In contrast, there appears to be no significant correlations with GVA per worker in agriculture and hotel and restaurant related services, the latter of which are present widely across all municipalities and thus only correlates with total number of workers, in line with theoretical expectations.

Figure 2.8 shows scatter plots between OECI and GVA per worker (total and across different macro-sectors) in 2019. In line with the previous table, we can see a positive correlation between the OECI and total GVA per worker, albeit with several outliers. For the case of different macro-sectors, the remaining panels show scatter plots between OECI and GVA per worker in manufacturing, ICT services and agriculture (the accommodation and restaurant service category is not portrayed due to the lack of any meaningful correlation). The positive correlation is strongest for the manufacturing sector, which aligns with the previous findings that the municipalities among the top economic complexity levels were traditionally focused on manufacturing, located in the northern coastal part of Portugal.

Table 2.5: Correlation between OECI (CNP-80) and other variables, municipalities, 1985

	OECI	Diversity	Avg. ubiquity	Education
OECI				
Diversity	0.72***			
Avg. ubiquity	-0.92***	-0.78***		
Education	-0.22***	0.11	0.12*	
Employment	0.37***	0.56***	-0.36***	0.18**

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 2.6: Correlation between OECI (CPP-10) and other variables, municipalities, 2002

	OECI	Diversity	Avg. ubiquity	Education
OECI				
Diversity	0.49***			
Avg. ubiquity	-0.94***	-0.57***		
Education	0.39***	0.61***	-0.45***	
Employment	0.55***	0.50***	-0.53***	0.37***

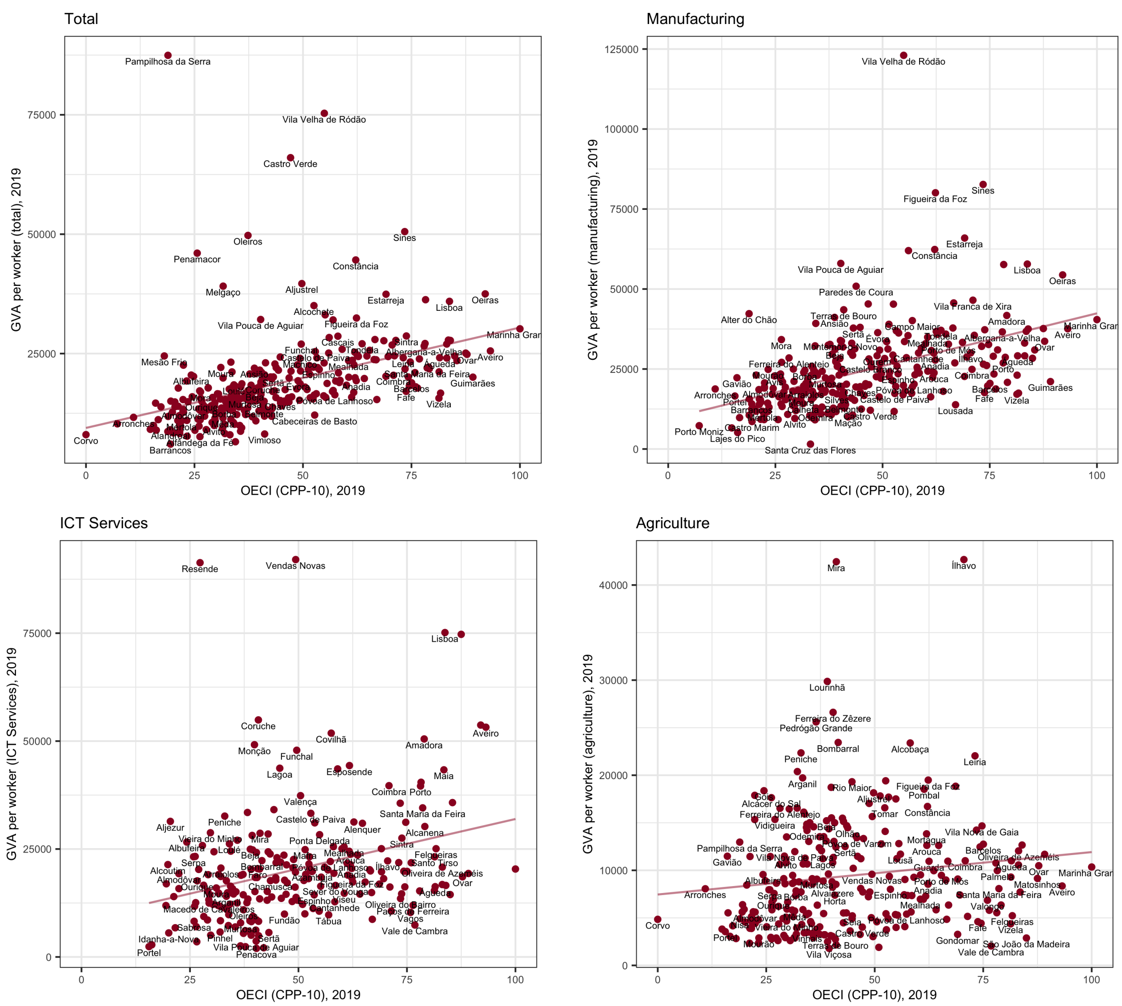
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 2.7: Correlation between OECI (CPP-10) and other variables, municipalities, 2019

	OECI	Diversity	Avg. ubiquity	Education	Employment
OECI					
Diversity	0.38***				
Avg. ubiquity	-0.90***	-0.47***			
Education	0.46***	0.54***	-0.60***		
Employment	0.43***	0.36***	-0.56***	0.55***	
GVA (total)	0.32***	0.28***	-0.44***	0.49***	0.97***
GVA/worker (total)	0.44***	0.14*	-0.40***	0.30***	0.23***
GVA/worker (manuf.)	0.48***	0.23***	-0.43***	0.45***	0.26***
GVA/worker (serv.)	-0.07	0.13*	-0.09	0.10	0.19***
GVA/worker (ICT)	0.32***	0.33***	-0.45***	0.41***	0.44***
GVA/worker (agric.)	0.14*	0.23***	-0.17**	0.08	0.08

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Figure 2.8: Scatter plots, OECI and GVA per worker (total and for different sectors), 2019



2.5.2 Empirical estimation

Focusing on the 1995-2019 period, Table 2.8 presents the results for our main estimation. While there is a negative association between the OECI and employment growth, once we control for initial employment in columns (2) and (3), the coefficient turns positive and is only statistically significant at the 10% level, indicating that in column (1) the OECI might be capturing convergence mechanisms. In both cases, employment shows a negative and statistically significance with employment growth, consistent with usual empirical findings, while there is a positive coefficient on the share of workers with post-secondary education, in line with theoretical expectations. Table 2.A.4 in the Appendix presents the results for 1985-1994, showing broadly similar results, with the exception that once initial employment is included, the OECI is not statistically significant, and neither is the education variable.

Table 2.8: OECI, Fixed Effects estimation results, municipalities, 1995 to 2019

	Employment growth		
Variables	(1)	(2)	(3)
OECI (CPP10)	-0.00107***	0.000681*	0.000616*
	(0.000302)	(0.000346)	(0.000326)
Employment (ln)		-0.236***	-0.233***
		(0.0223)	(0.0217)
Education (ln)			0.0149*
			(0.00808)
Constant	0.0666***	1.730***	1.734***
	(0.0105)	(0.157)	(0.157)
Observations	6,764	6,764	6,764
R-squared	0.148	0.267	0.268
Adjusted R-square	0.145	0.265	0.265
Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

2.6 Further analysis

2.6.1 Occupation complexity and task composition of occupations

This subsection bridges the economic complexity literature with the the task-based approach to labour markets to understand how the occupation-based ECI relates to our existing knowledge surrounding occupations and labour markets. In particular, we measure the task content of occupations in order to classify them into routine/non-routine and manual/cognitive categories and see how this relates to occupation complexity levels.

Methodology

To assess the task content of occupations, we departed from the theoretical framework and methodology introduced by Acemoglu and Autor (2011) and Autor (2013). The task descriptors originate from the US-based O*NET (version 26.1)[13], and we relied on crosswalks constructed by the Institute for Structural Research (Hardy et al., 2018), to convert occupations from the US SOC-10 classification to ISCO-08 classification. Due to the crosswalks available and the applicability of the O*NET work descriptors, we carry out this analysis for the 1995-2019 period, focusing on the ISCO-08 (or CPP-10) occupation classification.

We constructed four different groups of occupations, based on task intensity – non-routine cognitive (abstract), routine cognitive, routine manual and non-routine manual (physical).[14] Table 2.A.5 outlines the occupation task measures and O*NET descriptors of task composition of occupations used here, alongside their source and scale type (see Appendix).

There is a question of whether it is appropriate to use a US-based survey for the case of Portugal. For instance, the nature and task composition of occupations could be different due to different rates of technological adoption. This issue is likely to be amplified the further back we go in time (for instance, given the lower global integration at the start of the period compared to more recent decades). Several papers discuss this at length (e.g., Hardy et al., 2018), and the measures have been applied to the case of Portugal in the past (e.g., Fonseca et al., 2018a).

[13] Available online at: https://www.onetcenter.org/database.html

[14] The original work by Acemoglu and Autor (2011) splits the non-routine cognitive task intensive occupations into two groups (analytical and interpersonal); however, given the overlap between them and the use of a crosswalk, the descriptors under this category are brought together under the non-routine cognitive abstract task measure in other papers (e.g., Fonseca et al., 2018a for Portugal) and we adopt this approach here. There were no significant differences in the evidence and conclusions reached when we split the descriptors into the 'analytical' and 'personal' components.

Results

To compare and contrast how occupations in the different task-composition categories rank in terms of complexity levels, we calculated the average OCI of occupations within each category, and looked at the top and bottom occupations across them. Tables 2.9 and 2.10 show the results for 1995 and 2019 respectively. The task composition groups are ranked by complexity level, and for each group we show the average OCI and the top and bottom three occupations, with their respective OCI and ubiquity.

Table 2.9: Task composition of occupations and OCI (CPP-10), 1995

Task measure	**Average OCI**	**Top and bottom 3 occupations**	**OCI**	**Ubiquity**
Non-routine cognitive (abstract)	45.06	Legal professionals	93.85	2
		Mathematicians, actuaries and statisticians	76.79	5
		Nursing and midwifery associate professionals	75.99	3
		...	...	...
		Primary school and early childhood teachers	18.14	118
		Other personal services workers	17.27	132
		Professional services managers	3.37	161
Routine manual	29.95	Metal processing & finishing plant operators	55.10	37
		Assemblers	50.33	30
		Printing trades workers	50.01	33
		...	...	...
		Food processing and related trades workers	12.31	135
		Domestic, hotel and office cleaners	12.11	96
		Cooks	8.15	152
Non-routine manual (physical)	24.86	Fishery workers, hunters and trappers	74.65	2
		Refuse workers	48.48	19
		Ship and aircraft controllers and technicians	46.33	14
		...	...	...
		Animal producers	9.12	122
		Mining and construction labourers	9.11	210
		Mixed crop and animal producers	7.38	168
Routine cognitive	22.18	Physical and engineering science technicians	49.35	47
		Administrative and specialised secretaries	46.40	27
		Veterinary technicians and assistants	39.55	16
		...	...	...
		Medical and pharmaceutical technicians	10.32	193
		Other clerical support workers	4.51	170
		Numerical clerks	0.00	116

In line with theoretical expectations, the abstract occupations score highest in terms of average OCI in both years. However, within this group, there are also occupations with low

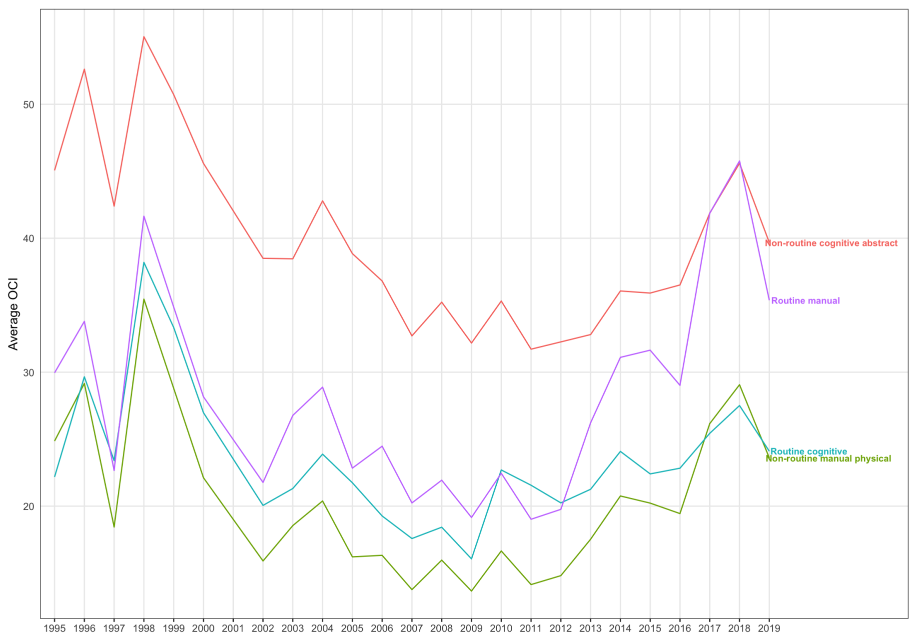
Table 2.10: Task composition of occupations and OCI (CPP-10), 2019

Task measure	Average OCI	Top and bottom 3 occupations	OCI	Ubiquity
Non-routine cognitive	39.68	Traditional & complementary medicine prof.	100.00	6
(abstract)		Mathematicians, actuaries and statisticians	99.20	6
		Sales, marketing & public relations prof.	90.69	10
		...	...	...
		Social and religious professionals	8.16	146
		Professional services managers	5.25	138
		Protective services workers	3.71	75
Routine manual	35.37	Metal processing & finishing plant operators	59.22	52
		Garment and related trades workers	56.53	45
		Blacksmiths, toolmakers & related trades	54.38	74
		...	...	...
		Domestic, hotel and office cleaners	3.24	144
		Food preparation assistants	1.85	153
		Cooks	0.76	143
Routine cognitive	24.10	Physical and engineering science technicians	59.30	71
		Keyboard operators	56.47	32
		Veterinary technicians and assistants	45.05	61
		...	...	...
		Personal care workers in health services	5.10	210
		Waiters and bartenders	2.66	120
		Client information workers	2.40	59
Non-routine manual	23.60	Handicraft workers	53.27	48
(physical)		Locomotive engine drivers & related workers	44.61	32
		Electronics & telecomms installers or repairers	44.52	48
		...	...	...
		Market gardeners and crop growers	7.27	149
		Agricultural, forestry and fishery labourers	0.36	165
		Mixed crop and animal producers	0.00	137

OCI levels, including teachers, service workers and managers. While we expected routine cognitive occupations to have the next highest complexity level, this is not the case and this group shows lower levels in both years. In 1995, the average OCI does not vary a lot across the routine manual, non-routine manual and routine cognitive measures; the highest levels of OCI are similar in all cases and they all include low complexity occupations, with the lowest values seen among the bottom three routine cognitive occupations. In 2019, the abstract and routine manual occupations are closer to each other in terms of average complexity levels, with their top occupations showing higher OCI levels than those at the top in 1995; the routine cognitive and physical task-intensive groups are very close in terms of average OCI, and the physical task-intensive group shows the lowest complexity levels.

Figure 2.9 plots the average OCI across task-content measures from 1995 to 2019, showing how the different occupation groups evolved over time in terms of average complexity levels. The same oscillations as before are mirrored in this plot, even though we would expect such an indicator to be much more stable over time. The abstract task-intensive occupation group shows the highest complexity across the period, though the routine manual task group sees increases over the period, coming much closer to the abstract occupations group. Routine cognitive and physical task measures remain fairly close to each other over the period, with some changes in ranking towards the end. In terms of changes over time, the complexity measure does not capture the structural changes that we might expect to see – for example, that routine manual task intensive occupations would become less complex, while abstract task intensive occupations would become more complex.

Figure 2.9: Average OCI across task-content measures, 1995 to 2019



Conceptually, the expectation would be for 'routine manual' task intensive occupations to be ranked among the lowest complexity levels; however, this is not what we observe. On the contrary, the 'routine manual' group experienced an increase in their average OCI level, vis-à-vis the other groups. The mechanisms for this might relate to how the ECI is measured. For instance, from the tables it appears that this was driven by an increase in the complexity of their top occupations. This might be driven by overall employment across municipalities shifting away from these occupations, thus decreasing their ubiquity level, and ultimately their OCI. Overall, we do not see a clear divide between the task composition measures that we may expect from conceptual standpoint and our occupation-based complexity level may be capturing other dynamics, such as specialisation patterns across the country, rather than 'complexity' in the conventional sense of the word.

2.6.2 Occupation vs. industry-based complexity measures

In this subsection we analyse the industry-based measures for 2007-2019, and compare them with the occupation-based ones.

Industries – ICI calculation

Tables 2.11 and 2.12 show the top and bottom ten industries ranked by ICI, for 2007 and 2019 respectively. Across both tables, the industries with the highest complexity levels are those related to financial services, as well as specific types of manufacturing, administrative services and freight air transport, which are activities that are typically highly geographically agglomerated. Among the lower complexity levels are several industries related to agriculture, forestry and fishing, accommodation and food service activities, mining and manufacturing activities (in this case, related to agriculture), as well as one industry related to residential care activities. These are all activities that tend to be widely spread across countries and regions – on the one hand, there are activities related to agriculture, which are spread across several municipalities within rural areas of the country; on the other hand, there are non-tradable services that tend to locate evenly across the country and in close proximity to consumers. The ICI therefore appears to capture closely the different patterns of specialisation across the country, and it may not ultimately relate to the underlying knowledge and capabilities inherent to these different industries.

Table 2.11: ICI, ubiquity and average diversity, top and bottom 10 industries, three-digit level (CAE-Rev.3), 2007

Rank	Industry (three-digit)	ICI	Ubiquity	Avg. diversity	Industry group (one-digit)
1	Trusts, funds and similar financial entities	100	1	100	Financial and insurance activities
2	Investigation activities	100	1	100	Administrative and support service activities
3	Reinsurance	95.23	2	88	Financial and insurance activities
4	Pension funding	95.23	2	88	Financial and insurance activities
5	Fund management activities	88.48	3	90.33	Financial and insurance activities
6	Freight air transport and space transport	83.23	2	93	Transportation and storage
7	Manufacture of magnetic and optical media	80.26	1	33	Manufacturing
8	Wireless telecommunications activities	70.32	5	85.4	Information and communication
9	Sound recording and music publishing activities	69.78	6	82.33	Information and communication
10	Manufacture of basic pharmaceutical products	67.99	4	80.75	Manufacturing
...	...	...	...	...	...
247	Cutting, shaping and finishing of stone	5.96	115	38.18	Manufacturing
248	Logging	5.46	91	37.08	Agriculture, forestry and fishing
249	Construction of residential and non-residential buildings	5.33	196	35.51	Construction
250	Hunting, trapping and related service activities	5.08	17	42.76	Agriculture, forestry and fishing
251	Support activities for other mining and quarrying	4.5	2	30.5	Mining and quarrying
252	Growing of perennial crops	4.04	114	36.57	Agriculture, forestry and fishing
253	Postal activities under universal service obligation	3.96	155	36.45	Transportation and storage
254	Residential care activities for the elderly and disabled	3.85	202	34.27	Human health and social work activities
255	Mixed farming	3.54	105	35.1	Agriculture, forestry and fishing
256	Monetary intermediation	0	69	31.93	Financial and insurance activities

Table 2.12: ICI, ubiquity and average diversity, top and bottom 10 industries, three-digit level (CAE-Rev.3), 2019

Rank	**Industry (three-digit)**	**ICI**	**Ubiquity**	**Avg. diversity**	**Industry group (one-digit)**
1	Trusts, funds and similar financial entities	100.00	1	64	Financial and insurance activities
2	Pension funding	98.62	1	93	Financial and insurance activities
3	Freight air transport and space transport	93.09	2	83	Transportation and storage
4	Reinsurance	90.00	2	63	Financial and insurance activities
5	Leasing of intellectual property and similar products	82.76	5	79	Administrative and support service activities
6	Fund management activities	80.80	4	69.5	Financial and insurance activities
7	Manufacture of batteries and accumulators	73.33	2	73.5	Manufacturing
8	Other financial service activities	72.80	6	78.17	Financial and insurance activities
9	Manufacture of musical instruments	72.48	3	70	Manufacturing
10	Manufacture of basic iron and steel and of ferro-alloys	70.93	3	60	Manufacturing
...	...	...	...	...	...
245	Hotels and similar accommodation	9.20	91	35	Accommodation and food service activities
246	Hunting, trapping and related service activities	9.08	9	41.44	Agriculture, forestry and fishing
247	Holiday and other short-stay accommodation	8.68	132	36.01	Accommodation and food service activities
248	Support activities to agriculture and post-harvest crop act.	8.41	100	37.68	Agriculture, forestry and fishing
249	Residential care activities for the elderly and disabled	8.09	224	33.88	Human health and social work activities
250	Manufacture of vegetable and animal oils and fats	7.08	45	38.02	Manufacturing
251	Mixed farming	6.28	106	34.18	Agriculture, forestry and fishing
252	Growing of perennial crops	3.79	116	33.77	Agriculture, forestry and fishing
253	Monetary intermediation	1.33	72	30.6	Financial and insurance activities
254	Mining of non-ferrous metal ores	0.00	5	28	Mining and quarrying

Municipalities – IECI calculation

Turning to municipalities, Tables 2.13 and 2.14 show the top and bottom 10 municipalities ranked by their IECI levels. From the tables, we can see a similar location pattern in terms of the most complex municipalities as that observed for the OECI. The most noticeable difference is that municipalities in the metropolitan areas of Lisbon and Porto are much more widely represented in the top ten (along with some manufacturing-focused municipalities, as before). This further suggests that these industry-based complexity measures mirror strongly the patterns of specialisation across the country, and thus capture the fact that some economic activities are only present in large urban areas. Among the bottom ten, there are once again municipalities in the Douro and Alentejo regions, mirroring closely what was found for the occupation-based measure. To complement this, the maps in Figure 2.10 show the IECI across municipalities in 2007 and 2019 respectively. The geographical patterns are similar with those observed for the case of the OECI, though less concentrated in this case, with some municipalities in the Algarve also showing relatively high levels of complexity in 2019.

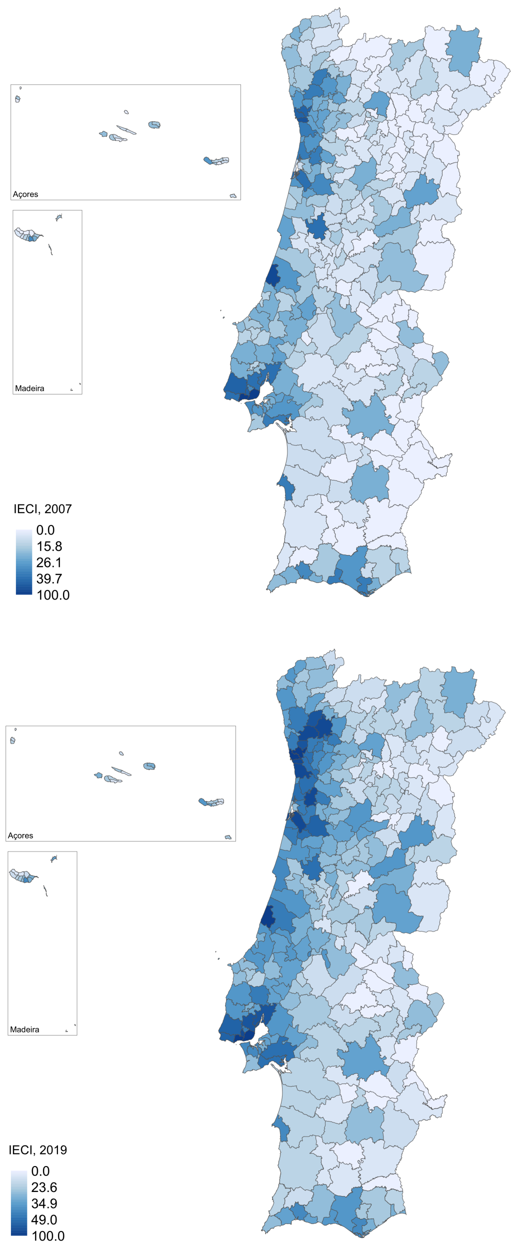
Table 2.13: IECI (CAE-Rev.3), top and bottom 10 municipalities, 2007

Rank	**Municipality**	**IECI**	**Diversity**	**Avg. ubiquity**	**NUTS-3 region**
1	Lisboa	100	100	34.28	Lisboa (MA)
2	Oeiras	92.41	76	39.66	Lisboa (MA)
3	Marinha Grande	84.53	33	39.12	Região de Leiria
4	Porto	79.94	95	44.38	Porto (MA)
5	Matosinhos	73.47	75	53.16	Porto (MA)
6	Maia	73.27	86	49.26	Porto (MA)
7	Sintra	71.18	90	54.1	Lisboa (MA)
8	Amadora	70.62	60	58.12	Lisboa (MA)
9	Vila Franca de Xira	69.26	61	53	Lisboa (MA)
10	Aveiro	68.69	69	52.67	Região de Aveiro
...	...	...	...	...	...
299	Penalva do Castelo	5.34	23	129.7	Viseu Dão Lafões
300	Gavião	5.02	11	125.82	Alto Alentejo
301	Sabrosa	4.86	19	125.58	Douro
302	Vila Flor	4.4	24	121.5	Terras de Trás-os-Montes
303	Alandroal	3.16	18	117.72	Alentejo Central
304	Santa Marta de Penaguião	2.39	16	129.62	Douro
305	Avis	2.07	18	113	Alto Alentejo
306	Mourão	1.5	13	130.69	Alentejo Central
307	Freixo de Espada à Cinta	0.37	12	129.58	Douro
308	Crato	0	17	133.53	Alto Alentejo

Table 2.14: IECI (CAE-Rev.3), top and bottom 10 municipalities, 2019

Rank	**Municipality**	**IECI**	**Diversity**	**Avg. ubiquity**	**NUTS-3 region**
1	Oeiras	100	64	36.2	Lisboa (MA)
2	Lisboa	98.77	93	33.42	Lisboa (MA)
3	Maia	91.22	73	44.45	Porto (MA)
4	Matosinhos	87.3	73	46.37	Porto (MA)
5	Marinha Grande	86.89	33	51.03	Região de Leiria
6	Guimarães	84	44	45.75	Ave
7	Vila Nova de Famalicão	82.17	54	47.02	Ave
8	Vila Nova de Gaia	81.64	95	53.81	Porto (MA)
9	Amadora	81.36	44	53.89	Lisboa (MA)
10	Oliveira de Azeméis	80.29	28	46.82	Porto (MA)
...	...	...	...	...	...
299	Armamar	8.99	23	124.7	Douro
300	Figueira de Castelo Rodrigo	8.14	25	125.24	Beiras e Serra da Estrela
301	Marvão	7.99	14	129	Alto Alentejo
302	Barrancos	7.34	10	135.2	Baixo Alentejo
303	Alfândega da Fé	6.93	21	118.43	Terras de Trás-os-Montes
304	Gavião	3.28	11	106.18	Alto Alentejo
305	Vidigueira	1.8	17	122.41	Baixo Alentejo
306	Alandroal	1.72	18	108.11	Alentejo Central
307	Alter do Chão	1.5	11	115.73	Alto Alentejo
308	Crato	0	16	127.5	Alto Alentejo

Figure 2.10: Map of IECI across municipalities, 2007 and 2019



Municipalities – IECI, OECI and other variables

Tables 2.15 and 2.16 show the correlation between the IECI and other variables for the beginning and end of the period. In line with what was observed for the OECI, industry-based complexity shows strong positive correlations with initial employment and the share of workers with post-secondary education, across both years with higher magnitude in 2019. There is also a strong positive correlation between the IECI and OECI. This is confirmed in the scatter plots between the IECI and OECI for 2007 and 2019 in Figure 2.11. While there are no major outliers, some municipalities show more divergent values and relative positions in terms of industry complexity, when compared to the occupation-based measure.

Turning to the IECI and other variables, Table 2.16 includes the correlation between industry-based complexity and GVA variables. The patterns of correlation are very similar to those obtained for the OECI, with a positive correlation between the IECI and total GVA per worker as well as GVA per worker in manufacturing, ICT services and, to a lesser extent, agriculture, whereas there is no correlation with GVA per worker in accommodation and restaurant related industries. The scatter plots in Figure 2.12, representing these correlations graphically, further confirm that the patterns in correlation across the different panels is similar to what was observed with the OECI.

We also estimated the regression model for the case of the IECI and compared the relative explanatory power of the two variables. Table 2.17 provides estimates that mirror the previous ones, for 2007-2019. Here, we also see a negative association between the IECI and employment growth, but there is no association once employment is included in columns (2) and (3); the coefficient on education is not statistically significant. Table 2.18 shows regression estimation results that compare directly the IECI and OECI, for 2007-2019. The IECI has a stronger explanatory power than the OECI, when we compare the individual regressions in columns (1) and (2), as well as when both variables are included in the horse race regression in the third column, with negative coefficients in all cases. As before, once the employment variable is included, both complexity measures lose statistical significance (see fourth column), suggesting that the negative association may simply be capturing convergence dynamics in both cases.

Table 2.15: Correlation between IECI (CAE-Rev.3) and key variables, municipalities, 2007

	IECI	Diversity	Avg. ubiquity	OECI	Education
IECI					
Diversity	0.77***				
Avg. ubiquity	-0.95***	-0.72***			
OECI	0.87***	0.61***	-0.87***		
Education	0.62***	0.64***	-0.53***	0.58***	
Employment	0.61***	0.55***	-0.53***	0.55***	0.59***

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 2.16: Correlation between IECI (CAE-Rev.3) and key variables, municipalities, 2019

	IECI	Diversity	Avg. ubiquity	OECI	Education	Employment
IECI						
Diversity	0.73***					
Avg. ubiquity	-0.95***	-0.72***				
OECI	0.88***	0.59***	-0.85***			
Education	0.52***	0.53***	-0.51***	0.46***		
Employment	0.54***	0.51***	-0.52***	0.43***	0.55***	
GVA (total)	0.42***	0.40***	-0.39***	0.32***	0.49***	0.97***
GVA/worker (total)	0.37***	0.20***	-0.36***	0.44***	0.30***	0.23***
GVA/worker (manuf.)	0.42***	0.27***	-0.43***	0.48***	0.45***	0.26***
GVA/worker (serv.)	0.08	0.14*	-0.15**	-0.07	0.10	0.19***
GVA/worker (ICT)	0.43***	0.38***	-0.43***	0.32***	0.41***	0.44***
GVA/worker (agric.)	0.20**	0.20***	-0.20**	0.14*	0.08	0.08

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Figure 2.11: Scatter plots of OECI and IECI, 2007 and 2019

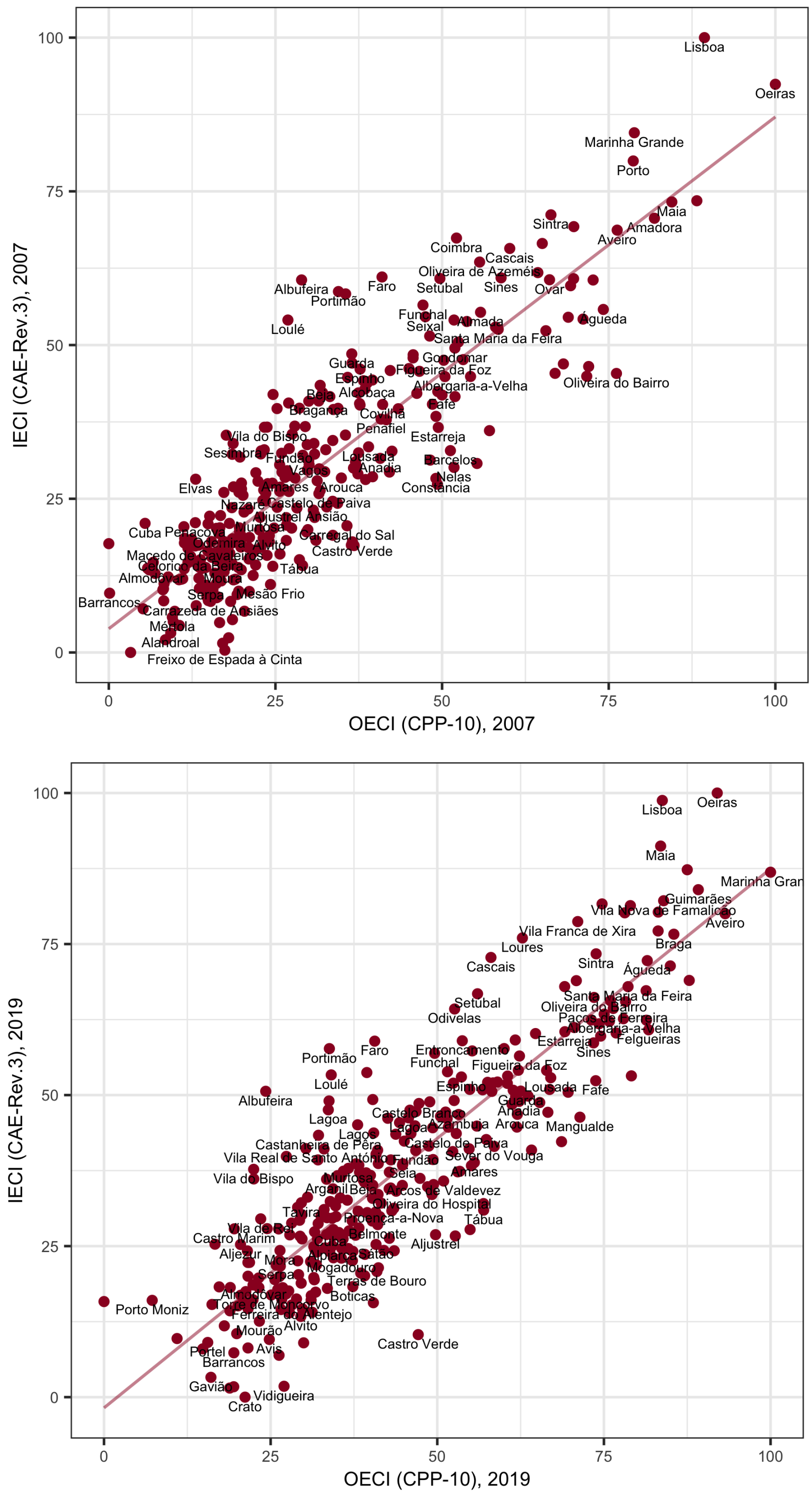


Figure 2.12: Scatter plots, IECI and GVA per worker (total and by sectors), 2019

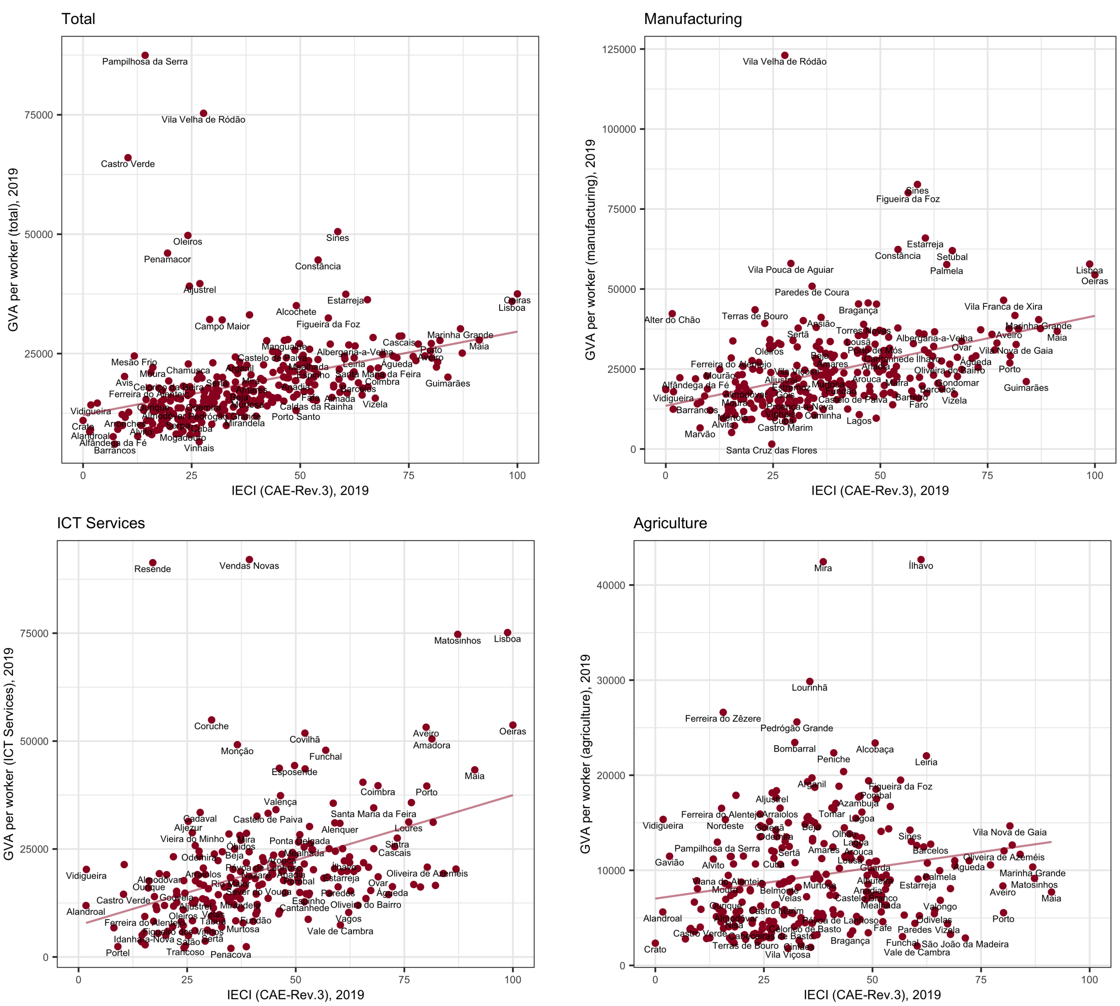


Table 2.17: IECI, Fixed Effects estimation results, municipalities, 2007 to 2019

	Employment growth		
Variables	(1)	(2)	(3)
IECI (CAE-Rev.3)	-0.00124***	0.000255	0.000210
	(0.000415)	(0.000425)	(0.000423)
Employment (ln)		-0.303***	-0.303***
		(0.0219)	(0.0220)
Education (ln)			0.00690
			(0.0109)
Constant	0.0399***	2.369***	2.391***
	(0.0132)	(0.168)	(0.170)
Observations	3,696	3,696	3,696
R-squared	0.190	0.307	0.307
Adjusted R-square	0.187	0.304	0.305
Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.18: OECI and IECI, Fixed Effects estimation results, municipalities, 2007 to 2019

	Employment growth			
Variables	(1)	(2)	(3)	(4)
OECI (CPP10)	-0.000906***		-0.000773**	0.000151
	(0.000331)		(0.000358)	(0.000329)
IECI (CAE-Rev.3)		-0.00124***	-0.000984**	0.000213
		(0.000415)	(0.000456)	(0.000456)
Employment (ln)				-0.305***
				(0.0218)
Constant	0.0312***	0.0399***	0.0559***	2.379***
	(0.0101)	(0.0132)	(0.0133)	(0.168)
Observations	3,696	3,696	3,696	3,696
R-squared	0.191	0.190	0.193	0.307
Adjusted R-square	0.188	0.187	0.190	0.304
Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

2.7 Discussion

What is a complex occupation?

In a straightforward way, the methodology and network used define a complex occupation as one that few regions are able to be specialised in (or have a 'revealed comparative advantage' in), relative to the rest of the country, and in turn those regions also have a concentration of other, relatively 'rare' occupations. The rationale, if applying directly from the definition provided by Hidalgo and Hausmann (2009), is that more complex occupations require 'capabilities' that only some regions are able to develop.

While in a lot of cases the complexity classification makes sense – for example, intuitively, we would think of engineers and certain types of mechanics as having a fairly complex job – in others, the most complex occupations appear not to be necessarily the most 'complicated' ones to perform, but rather those that are relatively rare and only found in some municipalities. Given the methodology at hand, if these relatively rare occupations (e.g., street and related service workers) happen to be present within a municipality that has a relatively high concentration in several other occupations that are relatively rare (and may, or may not, be complex or 'complicated'), these rare occupations will be classified as being 'complex'. Of course, what we want to understand is whether, when a municipality has a larger set of relatively more complex occupations, this has some meaning in terms of economic or structural advantage, or if they are an urban area with lots of people, providing a market for street and related workers. Ultimately, the mechanisms at play are very hard to disentangle, but they are crucial for the relevance of the economic complexity measures.

This is where existing papers that use occupation data fall short, as they do not provide an in-depth discussion regarding the interpretation and implications of economic complexity measures. Whereas for patents or exported goods there is some logic as to why they would be complex – for example, while only some regions may be able to patent in certain technology classes or export a certain product because they have some kind of advantage (e.g., a technological cluster where lots of people with the necessary skills are present), we would expect that technology class or product to be 'complicated' or hard to develop, otherwise it could be present in plenty of other regions that did not have any other complex technologies or products – it is not clear that the same logic applies when we think of occupation concentration.

This argument and broader discussion are closely interlinked with the one on the inclusion of services in ECI calculations. While some, often non-tradable, services are present in every city (e.g., hairdressers or family lawyers), others are only present in primary cities (e.g., specialised finance lawyers, international accounting firms or artists and other culturally-intensive profiles). Nevertheless, this does not directly translate into one being more or less complex than the other, but rather it may simply reflect that demand for some services is only supported by a larger market.

Implications of results

Our analysis comparing the occupation complexity measure derived and the task-composition framework provides further evidence to this point. Rather than following theoretical expectations – i.e., that more complex occupations may involve tasks that are intuitively harder to perform such as non-routine cognitive ones, whereas those involving mostly routine manual tasks would show lower complexity levels – it appears the logic behind occupation complexity follows specialisation patterns more closely, as well as changes over time in patterns of employment. For instance, there is a clear possibility that the employment share may increase for non-routine cognitive task-intensive occupations, and decrease for routine manual occupations (e.g., due to the latter being more likely to be replaced by machines over this period), and thus routine manual occupations become 'rarer' and less ubiquitous across the country, leading to a higher occupation complexity level, through a mechanism that, conceptually, is not what ECI measures aim to capture.

When comparing the occupation-based complexity measure with the industry-based one, this is further amplified, as the evidence points to high levels of the IECI being even more concentrated among the largest metropolitan areas, with the ICI also reflecting the specialisation patterns that are familiar to economic geographers, whereby certain activities tend to agglomerate geographically more than others. Of course, there is still the argument that those industries that are strongly agglomerated in space – due to the classic matching, sharing and learning opportunities they require – may be the most complex ones to perform and thus these network-based measures may still capture, to some extent, the inherent complexity level of industries.

Choosing development and geographical levels

This in itself leads to the question of the network used – if we were analysing municipalities or regions across several countries, we would capture several large or primary cities that have those services that only 'large' cities support, thus when we plot places against each other, we are not only comparing occupation or industry concentration across primary and secondary regions within a single country, but rather we would have several primary cities (e.g., capital cities and major metropolitan areas). While this could alleviate the issue of an apparent bias in terms of higher complexity levels seen among those occupations or industries that agglomerate in the primary cities, it is not clear that this would address broader concerns with the ECI applicability and interpretation. In fact, the existing empirical evidence points to the ECI methods being more accurate in applications to the US or China than when applied to Europe, partly due to the 'duplication' that occurs across countries in terms of economic activities, which does not represent an issue when applying the measure at the sub-national level within a single (large) country.

In this chapter we measure economic complexity based on a specialisation relative to the rest of the country. While it would be beneficial to extend the analysis to compare Portuguese regions with other (both peripheral and advanced) EU regions, the present comparison still makes sense and provides a more realistic and useful ranking of regions than, for example, comparing low-income countries to the rest of the world, where the bulk of complexity is concentrated in the most advanced countries. Conceptually, identifying the level of development becomes essential for the meaning of the ECI, and in this respect a sub-national analysis is likely much stronger.

The chosen geographical level of analysis has a significant effect on the economic complexity metrics derived, as well as on the interpretation of the ECI, due to its methodology. The ECI was originally developed at the country level, calculating a country's productive complexity relative to the rest of the world. As it was adapted to the sub-national level, the ECI has been applied to several different geographical levels, including municipalities, travel to work areas, metropolitan areas, US states and European NUTS-2 regions, often driven by data availability and with limited discussion about the interpretation of the metrics derived or implications for results.

Nevertheless, there are important factors to consider at a conceptual and methodological level: i) policy considerations – at what level is regional and development policy implemented, and thus where it interests us to better understand the current context; ii) methodological considerations – issues such as spatial dependence, overlapping labour markets or metropolitan areas (e.g., several municipalities make up the whole of the Lisbon and Porto metropolitan areas), as well as questions over the required minimum size of the network from which the complexity measures are derived; iii) data availability and constraints. These discussions are overlooked in many of the existing papers and the universality of the method is often taken for granted.

We carried out our analysis at the municipality level, which was the most detailed geographical level available in our data that covers the entire period. While this may involve issues of spatial dependence, there is still interest in measuring local specialisations, even within large metropolitan areas. For instance, it may be good to capture patterns of specialisation between core and peripheral areas of the same region since, as seen in the plots, there are still important divergences to capture within the NUTS-3 and NUTS-2 regional levels. Furthermore, our data relies on firm location, rather than residential one, and thus we are already capturing the place where the worker is productive and correctly matching people to places.

While it would be beneficial to capture local labour markets or travel to work areas, which are often used in other papers, such a geographic definition does not currently exist in Portugal. Moreover, whereas NUTS-3 regions would overcome the issue of separating different areas of the same metropolitan area, it has the drawback of grouping together different areas that are not well integrated – for instance, there are regions in the North in which the different small cities are not integrated with each other, due to mountains or otherwise poor connections. Thus, focusing on municipalities may give a better picture of specialisation and more accurately capture heterogeneity across the country.

Another reason to focus on municipalities is the sparsity of the network. The ECI measures require a large enough network, so that there is enough variation across places. With a high number of occupations and few NUTS-3 regions, several places will have the exact same diversity and ubiquity measures, leading to the same complexity score. In contrast, with the higher number of municipalities, more differences can be captured across places, and thus the complexity indicator becomes more insightful. This can be exemplified by comparing the 1985 and 2019 OCI calculations – in 1985, due to the higher number of occupations in the CNP-80 classification, several occupations have the exact same OCI levels due to the lack of variation (e.g., several occupations have only one region that is relatively specialised in them). If we go up to NUTS-2 region or OECD's functional urban areas, for example, we would have even less variation. The same issue would occur if using more aggregated occupation classification levels. While the original method was applied to a very sparse network, the size of the networks used is rarely discussed across papers, despite the significant impact it can have on the results and implications derived.

2.8 Conclusion

This chapter attempted to move a step closer towards understanding and conceptualising the ECI when applied to occupation data and, in doing so, exploring the specific case of Portugal, which has often been left out from economic complexity applications that rely on patent data. While, as hypothesised, occupation data has several potential benefits in terms of addressing drawbacks of alternative data sources, our understanding of the meaning and applicability of these methods to this type of data have not been explored and discussed in existing contributions.

Our analysis points to a discrepancy between the ways in which the ECI is often portrayed in applications and what it seems to measure. Conceptually, it appears to be much more closely aligned with Mealy's (2019) definition of the ECI as an ordering of regions that places those regions with most similar activities (here, occupational concentration) closer together in the ordering, and those with more dissimilar activities further apart. In particular, occupation-based complexity appears to reflect more closely location patterns within the country than the underlying knowledge or 'capabilities', or more broadly the intuition behind the word complexity, as something complicated or intricate.

Having established this, these measures are still useful in analytical ways. For instance, as discussed, a location quotient tells us how industries or occupations are spread across places but, by itself, does not tell us which locations are currently relatively specialised in industries or occupations that will become central to innovation processes and thus likely to experience economic growth and further prosperity (Kemeny & Storper, 2015). As a result, the ECI might still helps us understand the relative 'value' of a specialisation, based on the idea that those activities in which relatively few municipalities or regions are able to be specialised in somehow provide a competitive edge. The crucial question that follows is whether we want to assign value based on a complex quantitative methodology or, rather, through simpler existing indicators or some other qualitative judgement.

Further to this concern, in our specific context, it appears that the methods are mostly capturing traditional specialisation patters, and thus the most complex occupations and industries are simply those that tend to be present in municipalities within larger metropolitan areas or that are manufacturing-oriented. Possibly as a result of this, they do not appear to be closely related to regional paths, in particular in terms of employment growth.

There are questions that remain unanswered and can benefit from further research. While our broad conceptual discussion applies to empirical applications of this method to other contexts, the results derived and broader implications may be context-specific. Thus, additional investigation into different geographic contexts and levels, for example, comparing all EU regions together, can provide additional insight.

Overall, this chapter points to the need to be clearer about the meaning, conceptualisation and applicability of the ECI methodology when applied to different types of data and geographical contexts – this is often overlooked in existing empirical contributions, which take the universality of the ECI methods for granted.

2.A Appendix

Table 2.A.1: Papers reviewed – empirical applications in European contexts

Authors	**Geographic setting**	**Time period**	**Data**	**Outcome of interest**	**Findings**
Balland et al. (2019)	EU; 282 NUTS-2 regions	1990-2009	Patents	Entry and technological growth	Both relatedness and complexity have a positive, strong and statistically significant impact on technological growth at the regional level.
Antonelli et al. (2020)	EU; 189 NUTS-2 regions	1997-2009	Patents	Regional productivity growth	Positive effect of complexity on the generation of new technological knowledge, but negative one on its exploitation.
Mewes and Broekel (2022)	EU; 159 NUTS-2 regions	2000-2014	Patents	Economic growth	Technological complexity is an important predictor of regional GDP per capita growth (but result not robust when considering spatial dependencies).
Pintar and Scherngell (2022)	EU; 193 city-regions	2005-2014	Patents	Regional economic growth	There is a positive association between knowledge complexity and economic growth across regions, even after controlling for knowledge production.
Rigby et al. (2019)	EU; 145 city-regions	1981-2015	Patents	Employment growth and GDP growth	For employment growth, knowledge relatedness shows the strongest association; for GDP growth, knowledge complexity shows the strongest one.
Heimeriks et al. (2019)	EU; 286 NUTS-2 regions	2000-2014	Scientific publications	Entry (emergence of new scientific subfields)	Scientific complexity is important for developing new knowledge in complex scientific subfields; relatedness increases likelihood of developing more complex knowledge.
Pinheiro et al. (2022)	EU; 274 NUTS-2 regions	2011-2015	Employment (industries) and patents	Entries of new activities and average complexity of entries	Regions with high initial GDP per capita, complexity and population density have consistently entered higher complexity activities, compared with regions with low and medium initial levels.
Deegan et al. (2021)	EU; 128 NUTS-2 regions	2012-2018	Employment (industry and occupations)	Economic domains selected as priorities in Smart Specialisation Strategies	Both economic complexity and relatedness show a positive and statistically significant association with a region's selected priorities; there is no evidence of an interaction effect between complexity and relatedness.
Basile et al. (2019)	Italy; 103 provinces (NUTS-3)	1995-2015	Exports	Labour productivity growth	ECI contributed to increased productivity inequalities in the short and long runs, and was associated with polarisation of regional productivity.
Mealy and Coyle (2022)	UK; 380 local authorities	2011-2016	Employment (industry)	Local average annual earnings (and annualised growth rate)	UK local authorities with higher ECI have higher per capita earnings, growth rates and greater ability to develop further industries with greater earnings potential.
Bishop and Mateos-Garcia (2019)	UK; 218 travel-to-work areas	2015-2017	Employment (industry)	Technological activity emergence; economic performance	Locations with high economic complexity and a strong presence of emergent companies tend to have higher median earnings and GVA per capita.
Hane-Weijman et al. (2022)	Sweden; 72 local labour markets	2002-2012	Employment (occupations)	Regional employment growth	In times of growth, abandoning occupations close to a region's core capabilities slows employment growth, as does entry into more complex occupations. Negative impacts on employment growth are nullified in recessions.
Broekel et al. (2021)	Norway; 6226 industry-regions	2009-2014	Employment (occupations)	Industry-region employment growth	Positive effect of industrial relatedness on employment growth; no statistically significant effect of occupation diversity or occupation complexity.

Table 2.A.2: Description and source of variables (for municipalities)

Variable	Definition	Source	Years available
OECI (CNP-80)	Occupation-based economic complexity index; own calculation based on methodology described; CNP-80 occupation classification	Quadros de Pessoal (MTSSS)	1985 to 1994
OECI (CPP-10)	Occupation-based economic complexity index; own calculation based on methodology described; CPP-10 occupation classification	Quadros de Pessoal (MTSSS)	1995 to 2019
IECI (CAE-Rev.3)	Industry-based economic complexity index; own calculation based on methods of reflection methodology described	Quadros de Pessoal (MTSSS)	2007 to 2019
OCI (CNP-80)	Occupation complexity index; own calculation based on methodology described; CNP-80 occupation classification	Quadros de Pessoal (MTSSS)	1985 to 1994
OCI (CPP-10)	Occupation complexity index; own calculation based on methodology described; CPP-10 occupation classification	Quadros de Pessoal (MTSSS)	1995 to 2019
ICI (CAE-Rev.3)	Industry complexity index; own calculation based on methods of reflection methodology described	Quadros de Pessoal (MTSSS)	2007 to 2019
Employment growth	Calculated as the growth of municipality i between t and t+1 as $growth_{i,t+1} = log(GDP_{i,t+1}/GDP_{i,t})$; own calculation	Quadros de Pessoal (MTSSS)	1985 to 2019
Employment	Total number of workers employed by firms within the municipality	Quadros de Pessoal (MTSSS)	1985 to 2019
Education	Share of workers with post-secondary education. Number of workers with post-secondary education divided by total number of workers within the municipality	Quadros de Pessoal (MTSSS)	1985 to 2019
GVA (total)	Gross Value Added, Thousand EUR, absolute value	INE, downloaded from pordata.pt	2019
GVA per worker (total)	Ratio between Gross Value Added and the total number of workers (based on firm headquarters)	INE, downloaded from pordata.pt	2019
GVA per worker (industry)	Ratio between Gross Value Added in manufacturing (CAE-Rev.3 sector C) and the total number of workers	INE, downloaded from pordata.pt	2019
GVA per worker (services)	Ratio between Value Added in accomodation, restaurants and related services (CAE-Rev.3 sector I) and the total number of workers	INE, downloaded from pordata.pt	2019
GVA per worker (ICT)	Ratio between Value Added in information and communication services (CAE-Rev.3 sector J) and the total number of workers	INE, downloaded from pordata.pt	2019
GVA per worker (agriculture)	Ratio between Gross Value Added in agriculture, animal production, hunting, forestry and fishing (CAE-Rev.3 sector A) and the total number of workers	INE, downloaded from pordata.pt	2019

Notes: MTSSS is the Portuguese Ministry for Labour and Social Security. INE is Portuguese government office for national statistics.

Table 2.A.3: Summary statistics, 1985 to 2019

Variables	(1) N	(2) Mean	(3) SD	(4) Min	(5) Max
OECI (CNP80)	2,737	39.89	21.20	0	100
OECI (CPP10)	7,380	34.62	19.59	0	100
IECI (CAE-Rev.3)	4,004	31.34	17.66	0	100
Diversity (CNP80)	2,737	28.53	17.33	1	118
Diversity (CPP10)	7,380	29.94	9.500	1	74
Diversity (CAE-Rev.3)	4,004	37.45	16.60	2	100
Average ubiquity (CNP80)	2,737	96.12	30.45	27	204
Average ubiquity (CPP10)	7,380	117.6	20.45	41.27	208
Average ubiquity (CAE-Rev.3)	4,004	92.94	21.16	31.11	165
Employment growth (log)	9,809	0.0314	0.136	-1.462	1.778
Employment	10,117	6,090	17,850	12	314,550
Employment (ln)	10,117	7.501	1.506	2.485	12.66
Education	10,117	0.213	0.126	0.0143	0.882
Education (ln)	10,117	-1.722	0.609	-4.248	-0.126

Table 2.A.4: OECI, Fixed Effects estimation results, municipalities, 1985 to 1994

	Employment growth		
Variables	(1)	(2)	(3)
OECI (CNP80)	-0.00869***	-0.000866	-0.000865
	(0.00156)	(0.00108)	(0.00108)
Employment (ln)		-0.400***	-0.400***
		(0.0282)	(0.0285)
Education (ln)			0.00219
			(0.0252)
Constant	0.353***	2.776***	2.777***
	(0.0607)	(0.186)	(0.186)
Observations	2,433	2,433	2,433
R-squared	0.098	0.263	0.263
Adjusted R-square	0.0952	0.260	0.260
Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.A.5: O*NET descriptors for task composition of occupations

Occupation task measures	O*NET descriptors		O*NET source	Scale type
Non-routine cognitive abstract	4.A.2.a.4	Analysing Data or Information	Work Activities	Importance
	4.A.2.b.2	Thinking Creatively	Work Activities	Importance
	4.A.4.a.1	Interpreting the Meaning of Information for Others	Work Activities	Importance
	4.A.4.a.4	Establishing and Maintaining Interpersonal Relationships	Work Activities	Importance
	4.A.4.b.4	Guiding, Directing, and Motivating Subordinates	Work Activities	Importance
	4.A.4.b.5	Coaching and Developing Others	Work Activities	Importance
Routine cognitive	4.C.3.b.7	Importance of Repeating Same Tasks	Work Context	Context
	4.C.3.b.4	Importance of Being Exact or Accurate	Work Context	Context
	4.C.3.b.8	Structured versus Unstructured Work	Work Context	Context
Routine manual	4.C.3.d.3	Pace Determined by Speed of Equipment	Work Context	Context
	4.A.3.a.3	Controlling Machines and Processes	Work Activities	Importance
	4.C.2.d.1.i	Spend Time Making Repetitive Motions	Work Context	Context
Non-routine manual physical	4.A.3.a.4	Operating Vehicles, Mechanized Devices, or Equipment	Work Activities	Importance
	4.C.2.d.1.g	Spend Time Using Your Hands to Handle, Control, or Feel Objects, Tools, or Controls	Work Context	Context
	1.A.2.a.2	Manual Dexterity	Abilities	Importance
	1.A.1.f.1	Spatial Orientation	Abilities	Importance

Chapter 3

Economic complexity, exports and natural resources in the GCC

3.1 Introduction

As described in Chapter 1, the concept of economic complexity, as introduced by Hidalgo and Hausmann (2009), argues that it is possible to infer countries' knowledge and capabilities through their ability to export competitively a wide range of products that are somewhat rare relative to the rest of the world. The associated ECI has since grown in popularity as a way of predicting countries' economic growth, income inequality, human development, among other macroeconomic outcomes (Hidalgo, 2021); the link between economic complexity and future income growth is presented as an empirical regularity and as a key reason to support such indicators (e.g., Hausmann et al., 2014b). However, the applicability of this concept across different contexts has remained largely unquestioned, with research applying it with limited consideration of wider socioeconomic and development characteristics. In this chapter we argue that, once we consider countries that rely disproportionately on natural resources, the ECI method and conceptualisation may become problematic.

In a review of the literature from the past decade, Hidalgo (2021) argues that the strong predictive power of the ECI in explaining long-term economic growth suggests that a country's complexity level pins an equilibrium income level (i.e., that countries will converge towards a certain income level predicted by their complexity, thus the link between the initial ECI and subsequent growth). He further maintains that the direction of this relationship is from economic complexity to income growth, rather than the opposite, and that it would be somewhat improbable that countries with relatively low complexity given their income – such as Qatar, Oman, Bahrain and Kuwait, among others (Hidalgo, 2021, p.16) – will increase their complexity in the future. The implication of this argument is that these economies would grow less than others with similar GDP per capita, reverting to a lower ‘equilibrium’ income level in line with their ECI.

In parallel, Canh et al. (2020) estimated that economic complexity has a significant negative impact on total natural resource rents, and argued that focusing on improving economic complexity could help lessen the dependence on natural resource wealth. They find a strong negative relationship between the ECI and natural resource rents for upper-middle- and high-income countries – the latter group including Kuwait, Qatar, Saudi Arabia and the United Arab Emirates (UAE) together with other economies that do not rely on natural resource exports. Canh et al.’s (2020) results, however, may be biased by pooling together countries that are very different in nature, because of reverse causality between natural resource rents and economic complexity, and by the fact that the analysis did not include oil revenues.

More generally, economic complexity studies, to a different extent, tend to disregard the specific context and unique characteristics of natural-resource dependent countries. Important questions remain over whether the complexity concept and method can be meaningfully interpreted in all contexts, or rather should remain as a ‘big picture’ empirical regularity that should not be relied upon for finer-grained and specific analyses aimed to inform policy-making. Thus, to address these questions, this chapter explores the applicability and usefulness of economic complexity in oil-dependent countries, with a focus on the Gulf Cooperation Council.

Focusing on the GCCs allows us to explore a setting where oil and natural gas play a major role in the economy, and where this has resulted in high income growth and levels. At the same time, although to different extents, the GCCs have not managed to turn this income generation into improvements in education and skills, R&D investment, general openness and overall business environment (Hvidt, 2013; Kumar & van Welsum, 2013; Arman et al., 2021a), which are crucial in knowledge-based and complex economies. As the need to diversify into different economic activities becomes pressing, policymakers in these countries are turning to economic complexity measures to guide their efforts. As a result, it is important to understand whether the ECI can be meaningfully applied in this setting, as well as other oil-dependent countries, and what we can learn from it.

Using data from the Observatory of Economic Complexity for the period from 1995 to 2019, we look at how the ECI evolves over time across the GCCs, as well as where the GCCs stand in the correlations between the ECI and different key variables, such as income and industrial structure. We carry out regression analysis to assess whether and how the relationship between ECI and income growth differs for the GCCs vis-a-vis the rest of the sample, and explore whether the ECI can be meaningfully applied to this context. To further understand the impact of oil and natural gas products on complexity for the GCCs, we calculate the ECI excluding these products and investigate how this impacts economic complexity and its implications in contexts of high dependence on oil and natural gas.

Focusing on growth over 20-, 10- and 5-year periods between 2000 and 2019, we find that the positive cross-country association between the ECI and subsequent economic growth observed in the literature holds for the GCCs and other oil-dependent economies. Nevertheless, the ECI is not able to explain changes in income over time within countries, and is affected in different ways by oil and natural gas exports, suggesting caution and further consideration of context specificity when applying economic complexity measures.

The remainder of this chapter is organised as follows. Section 3.2 provides a literature review, focusing on the natural resource curse and on the economic complexity literature at the country-level. Section 3.3 describes our research aims and motivation, and provides a detailed discussion about the research context, the GCCs, followed by a description of the data and methodology in Section 3.4. Section 3.5 outlines the results, focusing firstly on exploring the economic complexity measures, for countries and products separately, and secondly on the empirical estimation results. Section 3.6 provides further analysis, where we exclude oil and natural gas products from the ECI calculation. Section 3.7 discusses our findings and their implications, followed by the conclusion in Section 3.8.

3.2 Literature review

In this section, we first describe what the natural resource curse is, its mechanisms and existing empirical findings. This is followed by an overview of the economic complexity literature, focusing on contributions at the country level, and existing attempts to link this concept with economic growth and natural resources.

3.2.1 The natural resource curse

Economists such as Rostow (1961) initially argued that natural resources could be a blessing for countries, allowing in particular developing countries to make a transition to industrial take-off, in a similar way to what happened in the UK, US and Australia (Rosser, 2009; Badeeb et al., 2017). The particularity of natural resources, vis-a-vis different resources or other economic activities, is that on the one hand they only need to be extracted, rather than produced, and therefore can occur somewhat independently of other economic changes and with limited employment creation; on the other hand, they are non-renewable and thus they should be seen more like an asset than a source of income (Humphreys et al., 2007; Badeeb et al., 2017). These assets can then finance higher levels of public and private consumption, including towards public goods such as infrastructure (Sachs, 2007), thus leading to the idea of a potential 'blessing' of natural resources.

While positive views were held until the 1980s, by then researchers started increasingly noting the lack of economic growth and worse development outcomes in Africa and Middle East countries which were rich in natural resources, and the 'Dutch disease' emerged (Corden & Neary, 1982; Corden, 1984; Neary & van Wijnbergen, 1986) – a phenomenon named after the discovery of natural gas in Groningen, which led to de-industrialisation and poorer macroeconomic performance – and halted the positive views. In parallel, Gelb (1988) found that oil economies experienced more serious deterioration in the efficiency of their domestic capital formation during the boom period in the 1970s. The first use of the term "resource curse" is attributed to Auty (1993) who, along with Gelb (1988), stressed the volatile nature of revenues from minerals.

The first empirical paper by Sachs and Warner (1995) showed that economic dependence on oil and mineral resources was correlated with slow economic growth in a cross section of countries. Further cross sectional studies by Sachs and Warner (1999, 2001) confirmed the adverse effects of natural resource dependence on economic growth.

Following this initial empirical evidence, researchers turned their attention to the potential channels through which the resource curse operates. Several mechanisms have been proposed in the literature, across both economic and political realms. In terms of economic mechanisms, the literature points to Dutch disease, volatility of commodity prices, failures of economic policy, including the neglect of education, as the key drivers; with regards to political mechanisms, rent seeking, weak institutions and corruption tend to be indicated as the culprits.

Natural resource revenues have a significant impact on the economic structure of a country, particularly when they make up a very large share of exports (Venables, 2016). The Dutch disease occurs when the discovery or boom of in natural resources leads to an increase in income and demand, which in turn generates inflation and a real exchange rate appreciation, making the relative prices of non-resource commodities higher and leading to lower competitiveness in world markets, ultimately also receiving lower investment (Sachs & Warner, 1995; Gylfason, 2001; Frankel, 2010; Badeeb et al., 2017). This can lead to a crowding out of the manufacturing sector, which often needs government intervention in the form of industrial policy, in particular to incentivise important 'learning by doing' mechanisms (Frankel, 2012). Overall, in the presence of natural resources, it is likely that the focus will be on current spending rather than long-term investment (Venables, 2016).

The prices of commodities such as oil and other resources are highly volatile, creating high uncertainty and difficulty in measuring expected revenues, and ultimately hampering planning for economic development (Badeeb et al., 2017). This volatility is in large part due to short-run elasticities of natural resources – for any given increase in price, demand does not fall much in the short-run and supply does not rise significantly in the short term either (Frankel, 2010, 2012). On the economic side, volatile revenues harm innovation especially in contexts of weak financial development, leading to exchange rate volatility; on the political side, this can lead to shortsighted policy making, by inducing a false sense of security (Van der Ploeg & Poelhekke, 2009). Van der Ploeg and Poelhekke (2009) examine the direct effect of natural resource dependence on growth and an indirect effect through volatility of natural resource revenues, and find that high world price volatility leads to volatility in income in countries that depend heavily on natural resources, which in turn leads to a significant negative impact on long-run growth itself. Thus, although a positive direct effect of resources on growth is generally found, the indirect effect through volatility is negative (and often dominant).

There were also theoretical questions about the long-run trend of world commodity prices and an idea – associated with economists Prebisch and Singer in the 1950s – that specialising in natural resources was a bad deal due to their price being expected to decrease over time relative to manufactured and other goods (Frankel, 2012). Nevertheless, the latter point is complicated by the fact that supply is not fixed over time and countries can adapt the extent to which they explore their resources in response to price and other economic changes (Frankel, 2012).

Within the economic realm, there are also challenges related to economic mismanagement and the incentive structure created by the presence of natural resources and the revenues generated. In particular, natural resource dependence reduces pressure on the government to collect taxes and exert fiscal discipline, and it can also lower incentives for human capital accumulation due to high levels of non-wage income or resource-based wages (Badeeb et al., 2017). Gylfason (2001) and Gylfason and Zoega (2006) focused their attention on broader channels through which natural resource dependence could be affecting sustained economic growth including savings, investment and human capital formation.

On the political side, rent seeking caused by a windfall of resource revenue can lead to increased power of elites, triggering very high inequalities (e.g., Gylfason, 2001), and it can lead to money being spent to the benefit of elites' immediate circles, rather than invested in infrastructure or development (Badeeb et al., 2017). Moreover, these revenues may also become a main cause for conflict between different stakeholders, including politicians, citizens and local communities (Davis & Tilton, 2005; Sala-i-Martin & Subramanian, 2013; Bodea et al., 2016). Valuable extractive resources such as oil, which do not require substantial labour and capital inputs (compared to, for instance, production), make factions more likely to fight over them (Frankel, 2012).

Further to this, corruption and institutional quality have been extensively researched in natural resource curse contexts and the evidence is mixed, and likely highly context-dependent. The quality of institutions is among the most hypothesised channels through which natural resources may influence long-run economic growth.

It is important to distinguish between two different lines of thought when it comes to natural resources and institutions. On the one hand, quality of institutions can be a potential mediating factor in turning natural resources either into a curse or a blessing – in the presence of good quality institutions, countries may be able to invest their resource rents in ways that help development; on the other hand, natural resources and their large rents can cause a deterioration in institutional quality (e.g., through the creation of conflict, rent seeking or eliminating the need for taxation and government restraint, among others).

Moreover, as argued by Frankel (2012), in countries like those in the Middle East where governments have access to large rents, they no longer need to tax the population and this might free them from the need for democracy – particularly as the need for tax revenue is believed to require democracy under the theory of 'taxation without representation'. Importantly, Boschini et al. (2007) argue that the extent of the negative effects of poor institutional quality depend on the level of appropriability of the natural resources present in a country, and that some natural resources are more problematic on this front.

Alexeev and Conrad (2009) explore in more detail the relationship between natural resources (focusing mostly on oil), GDP and institutional quality. Focusing on GDP levels, which they argue is more appropriate than growth rates, they find that oil enhanced long-term growth and show that oil and minerals do not have a significant impact on the quality of countries' institutions. They argue that previous results on the negative impact of natural resources on institutions were misleading as they overlooked the positive impact from resource endowments on GDP – if we consider that natural resources increase GDP without affecting other important development-related variables (e.g., institutions), then it is plausible that institutions or other variables will look relatively worse in countries that experienced growth due to natural resources compared to other countries with similar income levels but not reliant on natural resources (Alexeev & Conrad, 2009). Along with Ding and Field (2005) and Brunnschweiler (2008), Alexeev and Conrad (2009) also argue for the use of resource abundance or wealth (e.g., hydrocarbon deposits per capita or oil production per capita) – which show a positive effect of natural resources on economic growth, providing evidence that counters the idea of a resource curse – rather than resource dependence measures. Resource dependence measures are expressed as a share of GDP or total exports and they could be biased due to a country having low GDP for whatever reason, resulting in a high oil to GDP ratio (and a similar bias would be present for measures based on share of exports).

More recently, questions have emerged over the validity of the natural resource curse. The main reasons fall across three broad areas – first, concerns about the empirical strategies used in early papers; second, questioning of the time sensitivity of findings; and third, researchers have started to distinguish between abundance and dependence measures and argued the conclusions for each of these aspects and implications diverge. There are very comprehensive surveys of the literature available, such as Frankel (2010) and Badeeb et al. (2017), as well as Van der Ploeg (2011) and Venables (2016); thus we point to these for a more extensive overview of the existing evidence.

As outlined by Badeeb et al. (2017) empirical contributions are split across three broad groups. First, those who followed from Sachs and Warner, relying on cross-sectional analysis and employing different measures of resource abundance or dependence (e.g., Ding and Field, 2005; Mehlum et al., 2006; Mehrara, 2009). Second, those who focused on different variables related to growth that might be affected by natural resources, such as education and human capital development, savings rate, manufacturing exports, investment, fiscal policy and institutional quality (e.g., Gylfason, 2001; Gylfason and Zoega, 2006; Stijns, 2006; Dietz et al., 2007; Apergis and Payne, 2014). Third, those who question the validity of the resource curse hypothesis, predominantly through the use of more sophisticated identification strategies or alternative ways of capturing natural resources (e.g., Brunnschweiler and Bulte, 2008; Alexeev and Conrad, 2009).

In terms of empirical evidence relating to the GCC countries or broader geographical areas, Apergis and Payne (2014) look at the oil curse and growth in Middle East and North Africa (MENA) countries between 1990 and 2013 – employing panel data, they regress real GDP per capita on oil reserves and control variables using a time-varying cointegration methodology, and they also split their analysis by different groups of countries within the MENA region (based on the extent to which countries are resource rich or poor and labour abundant or importing). Their long-run results for the resource-rich labour-abundant countries support the hypothesis of an oil curse throughout the entire period, whereas for the resource rich labour-importing group, which includes the GCCs, the oil reserves coefficient is positive beyond 2003 to the end of the period; they argue that institutional conditions over time played a significant role in mitigating the adverse effects of an oil curse (Apergis & Payne, 2014). Exploring the case of 30 oil rich countries for the 1992-2005 period, Bjorvatn et al. (2012) find that the association between oil rents and income per capita varies with the balance of political power. In particular, oil rents are less likely to have a positive effect on GDP in countries with a high fractionalisation index, indicating that the government consists of a large number of small parties and thus is considered a 'weak government' (Bjorvatn et al., 2012).

Overall, although there is no general consensus, it seems that the natural resource curse is not inevitable as some countries have managed to avoid such outcomes. Instead, as Badeeb et al. (2017) argue, it is not resource abundance per se that causes the resource curse, but rather how the revenues are managed and the extent of the reliance on such revenues.

Countries such as the GCCs, have often turned to diversification efforts in an attempt to limit resource revenues becoming the sole economic activity and source of income; resource revenues can be used for investments such as human capital, public infrastructure, as well as targeting different sectors specifically (Venables, 2016). This can involve promoting sectors with backward linkages with the resource sector, for instance promoting the use of local inputs – an example of this are internationally competitive national resource companies like Saudi Aramco in the case of GCCs – or forward linkages, which involves processing further the natural resource prior to export or for local use, as well as supporting investment in sectors that are not directly linked with natural resources (Venables, 2016). Despite these efforts, failures to successfully diversify are common.

3.2.2 Economic complexity and natural resources

Chapter 1 provides a comprehensive discussion on the theoretical grounding of economic complexity, with a subsection focused on country-level theory and empirical research. This section therefore provides an overview of existing empirical research, focusing on the contributions that attempt to link economic complexity and the natural resource curse.

The original paper by Hidalgo and Hausmann (2009) introduced the ECI concept and measure as a way of quantifying countries' productive structures and showed that it was a good predictor of future growth. Since then, a vast number of empirical papers have emerged. On the one hand, several papers looked at the links between economic complexity and several different outcomes, including: economic growth (Poncet & Starosta de Waldemar, 2013; Hausmann et al., 2014b; Stojkoski et al., 2016; Tacchella et al., 2018), income inequality (Hartmann et al., 2017; Lee & Vu, 2020), human development (Ferraz et al., 2018), greenhouse gas emissions (Can & Gozgor, 2017; Neagu & Teodoru, 2019), and natural resource rents (Canh et al., 2020). On the other hand, research has started to explore the apparent drivers of economic complexity, looking at variables such as: institutions (Vu, 2019), modes of taxation (Lapatinas et al., 2019), intellectual property rights (Sweet & Maggio, 2015), demographics (Bahar et al., 2020; Vu, 2020), digital connectivity (Lapatinas, 2019), structural reforms (Demir, 2019), and natural resources (Yalta & Yalta, 2021; Ajide, 2022).

The link between economic complexity and natural resources is a complex one. The ECI takes into account all goods, including natural resources, which might make up a very large share of exports in some countries, as is the case of the GCCs. As a result, natural resources are likely to impact, firstly, the ECI calculations themselves and, secondly, the link between the ECI and economic growth or development.

The intuitive description of the methodology used in the ECI in the Atlas of Economic Complexity by Hausmann et al. (2014b) refers to the example of diamonds, as mentioned in Chapter 1. The ECI departs from the ubiquity – how common a product is among countries' exports – and the diversity – how many products a country exports competitively – and combines these two simple variables in an iterative process. The intuition is that a product that is exported competitively by very few countries would require a high level of capabilities (a term which the authors use to reflect pretty much anything a country might have that enables them to produce those exports, e.g., knowledge, skills, institutional settings, among others). Similarly, a country that is very diverse and able to export competitively a high number of products, is expected to have many and varied capabilities. Natural resources are an exception to this logic – as Hausmann et al. (2014b) describe, diamonds are produced in very few places and thus their ubiquity is low for reasons unrelated to knowledge-intensity. The iteration of ubiquity with diversity, they argue, will help correct for this – if diamonds were complex, the countries exporting them would be able to export many other products due to the high capabilities they had from exporting diamonds, but we know that this is not the case and that the countries that export diamonds or other natural resources tend to have very limited diversity and mainly export the natural resources available to them.

Natural resource revenues play an important role in generating higher GDP levels and growth, and thus natural resources are a key control variable included by Hausmann et al. (2017), who regress annualised growth in GDP per capita by decade on initial income per capita and initial ECI, controlling for the increase in net natural resource exports as a share of initial GDP. This was applied in subsequent empirical papers and the link between the ECI and GDP growth remains statistically significant when natural resource exports are included in regressions (e.g., Stojkoski et al., 2016; Hausmann et al., 2017).

Among the few papers that explicitly link economic complexity and natural resources, Canh et al. (2020) investigate whether economic complexity is a solution for the resource curse. Regressing natural resource rents on the ECI and several control variables (namely GDP growth rate, population density, capital formation, government expenditure and net FDI inflows), they found that economic complexity had a statistically significant negative impact on total natural resource rents, and argued that focusing on improving economic complexity could help lessen dependence on natural resource wealth. The authors also split the analysis between different World Bank income classification groups and find a strong negative relationship for upper-middle and high-income countries – the latter group combining Kuwait, Qatar, Saudi Arabia and the UAE together with other economies that have long moved away from relying on natural resource exports.

More recently, researchers explored the impact of natural resources on the ECI. Yalta and Yalta (2021) look at the determinants of economic complexity in MENA countries, focusing on a panel of 12 countries for the period 1970-2015. Drawing on existing literature on economic complexity and on the determinants of high technology exports they regress the ECI on key variables – GDP per capita, capital stock, education, FDI, institutions and natural resources – and find that education plays an important positive role, while a strong negative effect of natural resource rents on economic complexity is interpreted as support for the resource curse hypothesis. Moreover, they explore the interaction between natural resources and education and find that the marginal effect of natural resource rents depends on human capital accumulation. Ajide (2022) carries out a similar analysis, focusing on 32 African countries from 1995 to 2018, and reaches similar results, with a negative association between natural resource rents and the ECI and a mediating effect of education. Lastly, Avom et al. (2022) investigate the effect of natural resources on economic complexity, focusing on the interaction between political regime type and complexity, and find that the presence of democracy can help mitigate the negative effects of natural resource abundance on economic complexity (Avom et al., 2022), stressing once again the importance of governance of natural resources.

Tabash et al. (2022) link economic growth with natural resources and economic complexity, focusing on 24 African countries, and find a negative association between natural resource rents and economic growth and a positive one between economic complexity and growth. While this specification is similar to the one in our analysis, it suffers from some drawbacks as, for example, the authors do not include the ECI and natural resources in the same model and they rely exclusively on a system GMM estimation. While discussing policy implications, the authors argue that "economic analysts from the African region should consider the option of economic complexity as a remedy for low economic growth" (Tabash et al., 2022, p.7).

While these are important initial contributions towards our understanding of the link between economic complexity, natural resources and economic growth, important questions remain. First, results may be affected by pooling together countries with very differentiated economic and productive structures. Second, these papers all consider varying models – with natural resource rents, the ECI and GDP as dependent variables, and some combination of the ECI and natural resources as independent variables – without much discussion about reverse causality between natural resource rents and economic complexity. Third, oil revenues are not always considered in existing papers (focusing only on other natural resources), and the impact that abundance or dependence on natural resources could have on ECI calculations is not sufficiently explored.

3.3 Research aims and context

3.3.1 Research aims

The link between economic complexity and economic growth is seen in the literature as a general law, and the proponents argue that it is the key reason behind the importance of economic complexity indicators (e.g., Hausmann et al., 2014b). Yet, the GCCs have unique economic and political contexts. The aim of our research is to analyse the applicability of the ECI concept and measure to this group of countries. In particular, we investigate whether the link between economic complexity and income growth differs for the GCCs, and for oil-dependent countries more broadly. This will also help us understand to what extent policy implications can be derived from the ECI concept and measure in the case of the GCCs and oil-dependent economies more generally.

Theoretically, there are three broad hypotheses on the link between economic complexity and natural resource dependence. First, economic complexity could be viewed as an explanation or 'driver' of the natural resource curse – i.e., countries that can increase their economic complexity levels manage to break away from dependence on natural resource rents (Canh et al., 2020). Second, the ECI might simply capture the over-dependence on natural resources experienced by some countries – for instance, by reflecting their lack of knowledge and capabilities needed to export more diverse and complex products (as implied in Hidalgo and Hausmann, 2009; Hausmann et al., 2017 and in line with the findings by Yalta and Yalta, 2021). Third, the ECI might not be adequate as a concept in countries that are heavily reliant on natural resources, as it might simply reflect too strongly the fluctuations in commodity prices, which are highly volatile. Furthermore, there is a question of how the ECI is affected by changes in demand and prices of natural resource exports (i.e., volume of exports versus their monetary value), given the ECI reliance on export data.

The first hypothesis is theoretically problematic, not least because of the potential reverse causality between the ECI and natural resources. The second hypothesis is plausible, but incomplete: indeed, the ECI tends to be overly punishing towards high income economies with high shares of natural resource exports – an often-cited example is Australia, which has an advanced knowledge economy focused on services, but whose exports involve a lot of natural resources. The third hypothesis is a plausible one, and has not been explored in the existing literature.

Finally, the ECI in itself cannot predict a natural resource curse. While having a low ECI might indicate that current or initial productive structure is not very sophisticated and moving into more complex products may be hard, a natural resource-based country might still be able to invest in education and innovation and manage to increase both economic complexity and income, transforming dependence into a resource 'blessing' since, as discussed in Section 3.2, the resource curse does not appear to be inevitable.

3.3.2 Research context: GCC countries

Although the broad discussions are applicable to other oil-dependent countries, there are several reasons why this chapter focuses specifically on GCC countries. First, all six countries have very high shares of oil and natural gas exports, ranging from 45% of merchandise

Table 3.1: Country summary tables, GCCs, 1995 and 2019

1995						
Country	**GDP per capita**	**Oil rents (% GDP)**	**Oil & gas exp. share**	**Population**	**ECI**	**Rank**
Saudi Arabia	42 855.8	29.6	0.80	18 638 790	0.658	38
Oman	33 168.5	32.4	0.74	2 204 267	0.011	81
Kuwait	63 724.7	39.0	0.95	1 605 907	-0.006	83
UAE	101 571.0	17.5	0.75	2 415 099	-0.063	87
Bahrain	47 157.2	3.6	0.26	563 698	-0.109	89
Qatar	86 566.3*	28.0	0.85	513 447	-0.500	115

* GDP per capita value reported is from 2000, the earliest year for which GDP data is available in Qatar

2019						
Country	**GDP per capita**	**Oil rents (% GDP)**	**Oil & gas exp. share**	**Population**	**ECI**	**Rank**
Saudi Arabia	46 962.1	24.2	0.75	34 268 529	0.803	40
Kuwait	49 853.7	42.1	0.85	4 207 077	0.556	53
Bahrain	45 311.9	2.2	0.41	1 641 164	0.495	56
UAE	68 263.7	16.2	0.39	9 770 526	0.382	63
Qatar	89 966.4	16.9	0.88	2 832 071	0.022	88
Oman	31 284.0	24.9	0.69	4 974 992	-0.232	105

exports in Bahrain and 88% in Qatar in 2019. Second, they are all classified as high income countries by the World Bank and have roughly comparable GDP per capita levels (World Bank, 2022). Table 3.1 shows selected indicators for the GCC countries, illustrating these points. Population size varies significantly across the six countries and all have experienced vast population growth over the period of analysis. Third, the GCCs share a similar geographical context and have broad similarities in terms of history and culture (Valenta & Jakobsen, 2016).

While there are other countries with similar or even higher levels of exports share oil and natural gas, including South Sudan, Libya, Nigeria and Algeria among others, they represent significantly different socioeconomic and geographic contexts and thus including them would hinder detailed and careful analysis. Papers that analyse the MENA region typically split the countries into three distinct groups, as described by Samans and Zahidi (2017): i) the natural resource-rich (in particular in terms of oil and natural gas), high-income countries of the GCC; ii) labour-abundant, middle income countries (Egypt, Tunisia, Algeria, Morrocco); iii) conflict-affected areas (Syria, Iraq, Yemen). In this chapter we are therefore focusing our analysis on the first group of countries within the MENA region. This group is also sometimes referred to as the 'resource-rich labour-importing' countries (e.g., Apergis and Payne, 2014), with immigrant workers accounting for a very big share of the population and workforce, as will be further discussed.

As evidenced in Table 3.1, oil and natural gas play a crucial role in the economy of the GCCs. Figure 3.1 shows oil and natural gas exports as a share of total merchandise exports across the GCC countries from 1995 to 2019.[1] These products make up a very significant share of total exports across all countries, ranging from lower levels in Bahrain (although above 30 percent most years), to over 80 percent in Kuwait and Qatar over time. The largest change occurred in the UAE, with a significant overall decrease over the period, while Bahrain saw the sharpest oscillation, with a sharp increase in 2014 due to a significant increase in oil production, followed by decreases, as in the rest of the group, due to low oil prices in those years. There was relative stability in Kuwait and Qatar, and some decline in the last decade in Oman and Saudi Arabia.

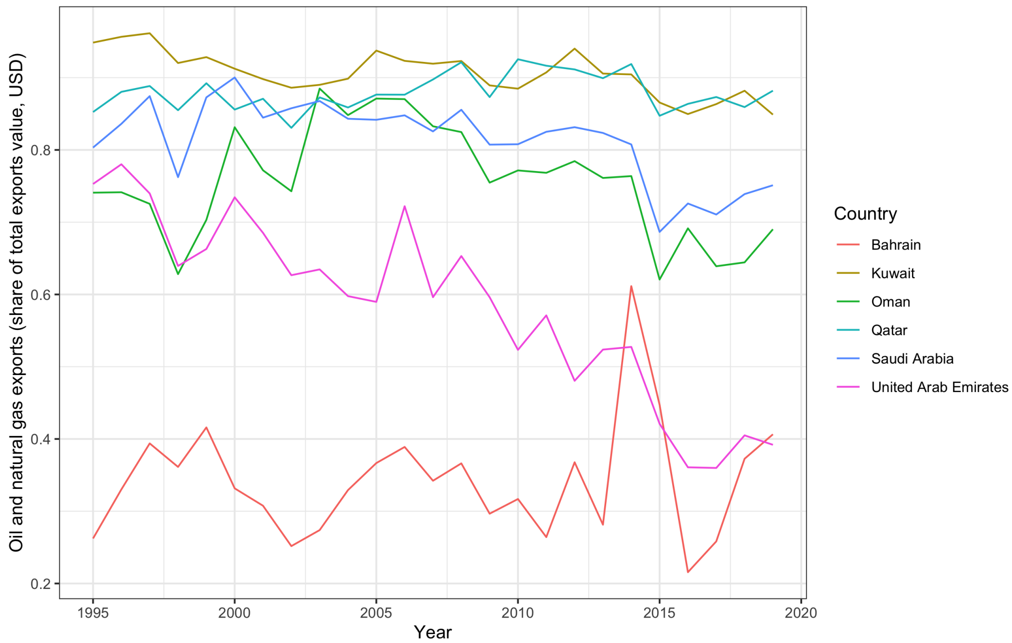
To investigate the role of oil rents in the GCCs, Figure 3.2 shows oil rents as a percentage of GDP from 1995 to 2019 – significant oscillations are observed, with sharp increases and drops every five to ten years that occur with similar patterns across all countries, mimicking the business cycle. Kuwait is the country where oil rents play the biggest role in the economy. Figure 3.A.1 in the Appendix shows total natural resource rents (percentage of GDP), from 1995 to 2019, showing that oil rents are the most important natural resource in these countries, with trends and levels that mirror closely those seen in Figure 3.2.[2]

A crucial implication of such a heavy reliance on oil and natural gas in the GCCs is that, as described by Beblawi (2011), oil extraction is not simply another economic activity that exists in addition to an advanced productive structure (as in the case of Canada, Australia or Scandinavian countries, for example), but it dominates the economy and is the almost exclusive source of wealth. Importantly, the GCC area has achieved high GDP growth and ultimately high income levels through natural resource availability, though without recording much improvement in other socio-economic conditions vis-a-vis other high-income economies. GCCs, although to different extents, show limited advancement in aspects considered crucial for knowledge-based and complex economies, such as education and skills, R&D investment, general openness and overall business environment (Hvidt, 2013; Kumar & van Welsum, 2013; Arman et al., 2021a). The main explanations relate to the existence of an 'allocation state' model, which is driven purely by the state, focused on wealth distribution, the extensive reliance on migrant labour, and a significant underdevelopment of productive assets, resulting in a failure to deliver further development, as it does not

[1] Data for this figure originates from the Observatory of Economic Complexity. We consider the following oil and gas products (and respective HS-92 four-digit level classification code): Petroleum oils, crude (2709); Petroleum oils, refined (2710); Petroleum gases (2711); Petroleum jelly (2712); Petroleum coke (2713).

[2] Both oil rents and natural resource rents indicators are estimated by the World Bank. Oil rents are the difference between the value of crude oil production and total costs of production. Total natural resource rents is the sum of oil rents, natural gas rents, coal rents (hard and soft), mineral rents, and forest rents.

Figure 3.1: Oil and natural gas exports (share of total exports), GCCs, 1995 to 2019



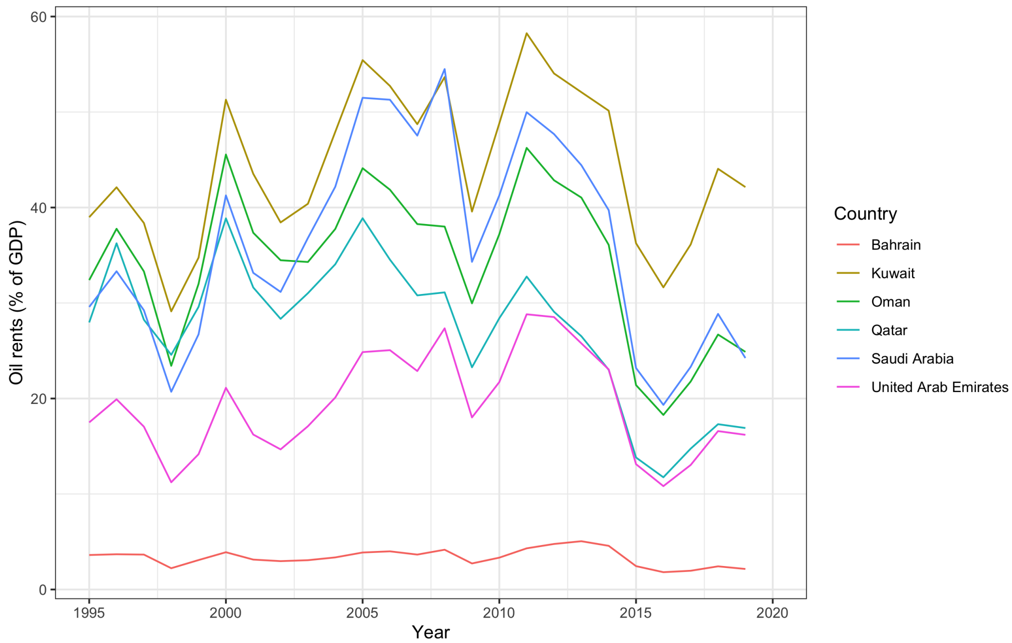
generate a stable and sufficient income for the population or job opportunities for the increasingly young and well-educated population (Hvidt, 2011, 2013).

As Figure 3.3 shows, GDP per capita has remained stable over the 1995 to 2019 period in Bahrain, Qatar and Saudi Arabia, whereas Kuwait, Oman and the UAE saw oscillations and an overall decrease – these trends result from an overall GDP increase across all countries, but an even larger population growth, illustrated in Figures 3.A.2 and 3.A.3 in the Appendix, which show total GDP and population changes over the period.[3]

In addition to the reliance on oil extraction and on foreign workers, the GCCs depend extensively on the public sector to generate employment, particularly for native workers. In recent decades, all countries have tried to address this issue in their development and diversification plans, identifying the need to move away from this reality in order to ensure longer term sustainability, with the GCCs with more limited oil resources expressing more urgency and emphasis than the others (Hvidt, 2013). Efforts to increase private sector R&D in Kuwait, Oman and Qatar have tried to address precisely the two key challenges of lack of diversification and over-reliance on the public sector (Ennis, 2015; Arman et al., 2021b). However, despite these efforts over the past few years, private sector R&D remains low in Kuwait and other GCCs, and often at levels that would be expected in much poorer economies (Arman et al., 2021b).

[3]Figures 3.3 and 3.4 rely on data from the World Bank (see Section 3.4 below for details).

Figure 3.2: Oil rents (percentage of GDP), GCCs, 1995 to 2019



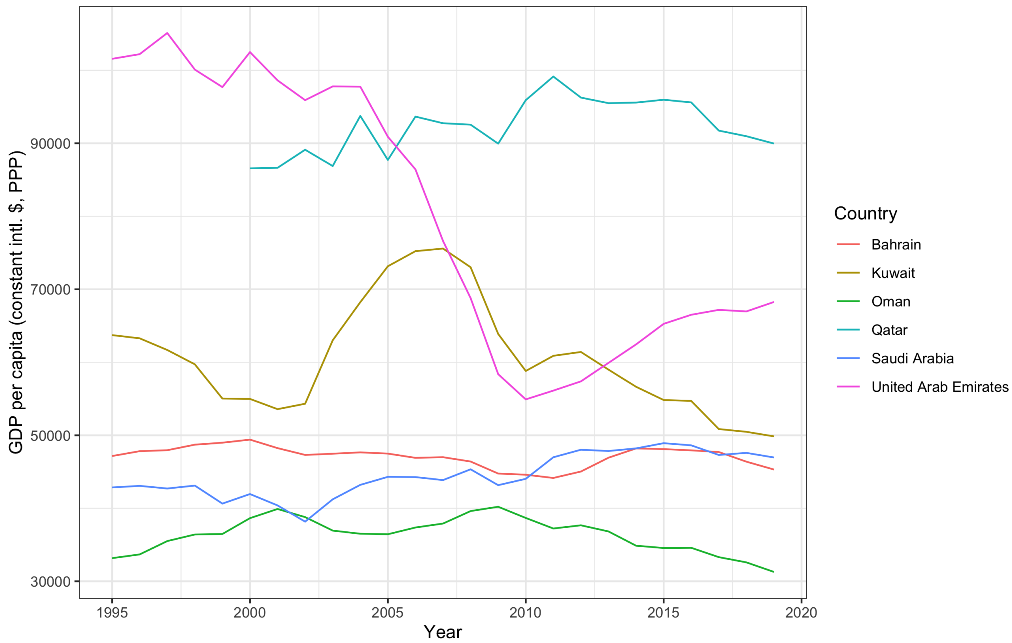
As mentioned, the GCC countries have all experienced large inflows of foreign workers, including many from the labour-abundant countries within the MENA region (as well as from India and South-East Asia), dating back to the time of fossil fuel discoveries and an influx of physical and service workers (Samans & Zahidi, 2017). As a result, all GCCs show a strong duality in the labour force, with migrant workers performing most of the technical, service and manual jobs predominantly in the private sector, and native workers mostly occupied in in the public sector (Samans & Zahidi, 2017).

Figure 3.4 illustrates the changes in employment shares in industry and services over time. Employment in industry increased in Qatar, Oman and to lesser extent in Bahrain between 1995 and 2019, while it remained relatively stable across the other countries, accounting for roughly a quarter of jobs in the GCCs. Services are by far the most important source of jobs – they saw a decrease in Qatar, Oman and Bahrain that mirror closely the increase seen in employment in industry, while the UAE experienced a small increase, which could be related to the country's diversification efforts towards services, including banking and tourism industries over the past decades.[4]

Given the dependence on oil, diversification has been at the forefront of economic policy in the GCCs for several decades. As described in detail by Hvidt (2013), the reasoning

[4]Agriculture (not pictured), makes up a very small share of employment across the GCC area and decreased over the time period.

Figure 3.3: GDP per capita, GCCs, 1995 to 2019



behind this has been twofold, with slight changes experienced over time. On the one hand, diversification away from oil is highly desirable given the limited lifespan of oil and natural gas, and this was the main reasoning behind diversification in the 1970s (Koren & Tenreyro, 2010); moreover, the imminent need to move towards cleaner sources of energy has emerged more recently with the climate crisis. On the other hand, diversification of the economy can be beneficial even in the presence of extensive oil reserves, as it can help alleviate the sharp oscillations generated by oil market volatility. The sharp decline in oil prices experienced in the 1980s and the volatility that persisted also throughout the 1990s shifted the focus towards this second rationale for diversification (Hvidt, 2013).

Despite this focus, the drive towards diversification has not been particularly successful, with several shortcomings and challenges outlined in existing literature. For example, assessing previous development reports on Kuwait, which focused mostly on the transition to a knowledge economy, Brinkley et al. (2012) point to the lack of a systemic approach, ignoring the institutional basis of a knowledge economy, as one of the key reasons why previous development strategies had not led to a better economic outcome in Kuwait. Similarly, Arman et al. (2021b) identify several challenges related to the National System of Innovation in Kuwait, in particular related to limited workforce skills level compared to world standards, lack of collaboration and economic linkages. Moreover, assessing the progress towards a knowledge-based economy in the GCC countries, Kumar and van Welsum

(2013) find that the important balance needed between physical and human capital has not been achieved because the focus on ICT infrastructure did not occur in parallel with sufficient development of knowledge, skills and other factors that are essential to a knowledge economy.

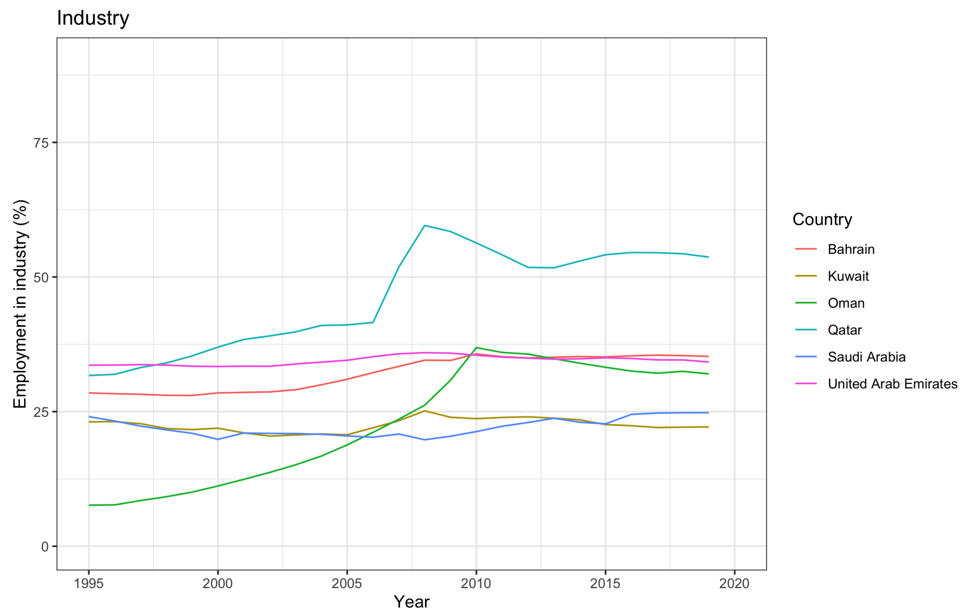
A key aspect to consider is the measurement of diversification. While diversification can usually be measured in very simple and crude ways across countries, it is not an easy exercise in the case of the GCCs, due to both the lack of high quality data for some variables and the impact of changing oil prices on available variables (Hvidt, 2013). The variables that are typically applied in the case of the GCCs include: i) the percentage contribution of oil and non-oil sectors to GDP; ii) the proportion of oil revenues as a share of total government revenues; iii) the share of non-oil exports in total export earnings; iv) the relative contribution of public and private sector to GDP; v) GDP volatility in relation to oil price fluctuations (Hvidt, 2013).

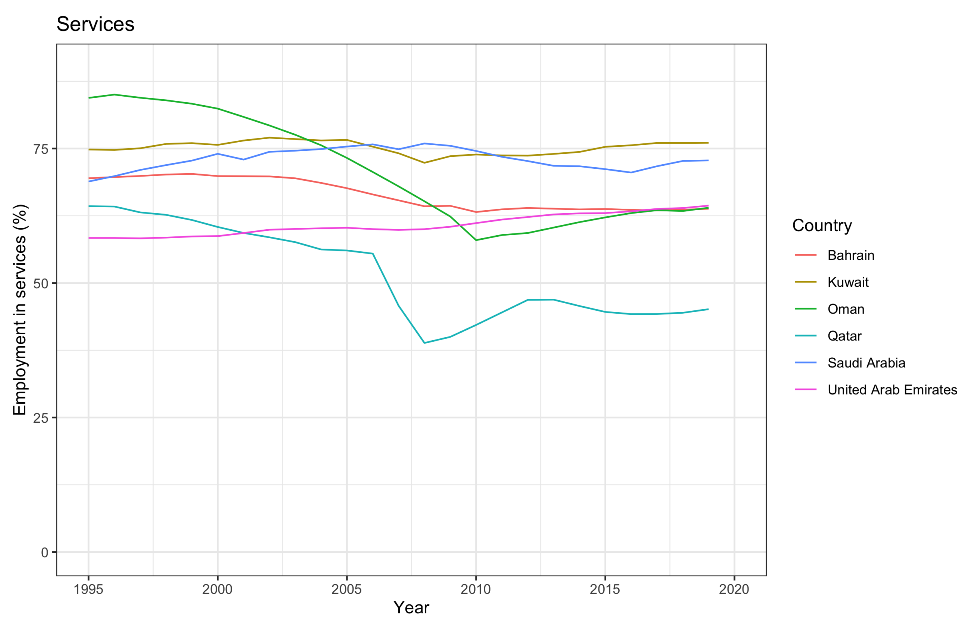
In this regard, economic complexity measures based on exports suffer the same issues and challenges, though this is disregarded in the literature, with existing contributions simply addressing this challenge by controlling for natural resources when assessing the impact of economic complexity on income growth. In particular, we expect the reliance on oil and natural gas exports, which experience different surges in demand and prices due to forces external to these countries, to affect their ECI. As they are based on export values, the RCA and ECI calculations might be impacted by the volatility of these exports, affecting the accuracy of the complexity measure and the extent to which it captures underlying capabilities in the GCCs.

In the area of exports, several papers investigate the validity of the export-led growth hypothesis in the GCC context (Kalaitzi & Chamberlain, 2021). Broadly speaking, these studies analyse the relationship between merchandise exports and economic growth, focusing on several GCCs (e.g., Kalaitzi and Chamberlain, 2021), or specific ones (e.g., Kalaitzi and Cleeve, 2018; Kalaitzi and Chamberlain, 2019; Chamberlain and Kalaitzi, 2020; Kalaitzi and Chamberlain, 2020, for the UAE). Some of these papers focus on specific types of exports, such as the case of fuel-mining exports (Chamberlain & Kalaitzi, 2020), or non-oil exports and re-exports (Kalaitzi & Chamberlain, 2019) to identify their link with growth, while the remainder look at merchandise exports more broadly. The results tend to show rather disparate patterns, with Kalaitzi and Chamberlain (2021) finding differing results across the GCC countries, as well as different short- and long-run patterns in causality between exports and economic growth between 1975 and 2016.

Overall, the volatility experienced by the GCC group is a crucial aspect: in part due to their strong dependence on oil, these economies are intrinsically more volatile than others at the same level of development (Koren & Tenreyro, 2010). Beyond exports or economic performance, there is further literature exploring the links between volatility transmission among GCC countries, looking for example at stock markets and oil shocks (e.g., Jouini and Harrathi, 2014), which emphasises the wider recognition that the GCC countries are undoubtedly highly subject to consequences from oil price shocks in a wide range of socioeconomic areas.

Figure 3.4: Share of employment in industry and services, GCCs, 1995 to 2019





3.4 Data and methods

3.4.1 Data

This chapter relies on the export data and complexity calculations introduced in Chapter 1. Thus, we focus on complexity measures calculated from a network with 1241 products and 179 countries, for the period from 1995 to 2019, relying on the HS-1992 four-digit level product classification.

In addition to this, we calculated the ECI and PCI based on a network that excludes the following oil and natural gas products (and respective HS-92 four-digit level classification code): Petroleum oils, crude (2709); Petroleum oils, refined (2710); Petroleum gases (2711); Petroleum jelly (2712); Petroleum coke (2713). The ECI measure excluding oil is used in our further analysis section, to explore the impact that these products have in complexity measures for the GCCs and oil-dependent countries.

While we could have used the OEC or the Atlas of Economic Complexity calculations, we computed the ECI ourselves, which provides two major advantages as it allows us to: first, to work with a more stable sample of countries over the time period and to leave out very small countries or territories; second, not only to calculate the ECI including all products, but also to exclude oil products and explore what happens to the GCC and other oil-dependent countries.

The control variables were downloaded from the World Bank Open Data, with the exception of the Human Development Index which originates from the United Nation's Human Development Reports. Table 3.A.1 in the Appendix provides the definitions and sources of our variables, Tables 3.A.2 to 3.A.4 provide summary statistics, while Table 3.A.5 lists the 179 countries included in our analysis.

3.4.2 Methodology

The objective of this chapter is to explore the applicability of the ECI concept and measure to the GCC countries. On the one hand, we want to investigate the ECI's internal validity for the GCC group, focusing on how economic complexity levels changed over time and what may be driving oscillations. To this end, we start by exploring how the ECI evolves from 1995 to 2019 and how it correlates with other key variables, with a focus on where the GCC countries stand vis-a-vis the other countries in the sample.

On the other hand, we want to explore external validity and look at the relationship between the ECI and economic growth, and whether it differs for this particular group of countries. Here, we turn to regression analysis. We replicate the common specification in the economic complexity literature, originally done by Hidalgo and Hausmann (2009) and Hausmann et al. (2014b) and later by Stojkoski et al. (2016) among others, of regressing economic growth over long time periods on the initial income level, initial ECI and control variables capturing natural resource dependence and trade openness, and explore whether the results differ for GCC countries and oil-dependent countries more broadly.

Our main specification is as follows:

$$growth_{i,t+n} = \alpha + \beta_1 ECI_{i,t} + \beta_j \boldsymbol{X}_{j_{i,t}} + \beta_7 GCC + \eta_t + \epsilon_{it}$$

where $growth_{i,t+n}$ is the GDP per capita growth between t and $t+n$ for country i, calculated as $growth_{i,t+n} = log(GDPpc_{i,t+n}/GDPpc_{i,t})$. $ECI_{i,t}$ is the initial ECI, our independent variable of interest. $\boldsymbol{X}_{j_{it}}$ is a vector representing the control variables – in line with existing contributions, they include: i) initial GDP per capita (natural logarithm) to control for convergence across countries; ii) increase in natural resource exports over the period (as a share of initial GDP) to capture the importance of natural resources; iii) increase in total exports (as a share of initial GDP) to capture the growth in exports and show that, despite being based on exports, the predictive power of the ECI is not lost due to controlling for increases in exports over the period; iv) initial ratio of exports to GDP to control for different levels of openness across countries; and v) initial population (natural logarithm) to control for any country size effects. η_t and ϵ_{it} represent time fixed effects and the error term, respectively. The transformations of the variables used are in line with the aforementioned papers.

To fulfil the second research aim, we add a dummy variable representing whether a country is part of the GCC, as well as an interaction term between the GCC variable and the initial ECI. Moreover, we perform the same analysis on the sample of oil-dependent countries only – they were identified by looking at the share of exports in oil and natural gas products in total exports over the period, and selecting the countries where these products make up over 30 percent of exports on average (this was identified by carefully exploring the data; there were no countries immediately under this threshold); Table 3.A.5 in the Appendix identifies the oil-dependent countries considered in our analysis. Overall, the aim is to understand whether and how the relationship between the ECI and income growth differs for GCC or oil-dependent countries and the rest of the sample.

We focus on the cross-country association between initial economic complexity and subsequent income growth for 20-, 10- and 5-year periods between 2000 and 2019.[5] To check for within-country association, we also estimate our model using Fixed Effects. This is in line with existing literature, including the original contribution by Hidalgo and Hausmann (2009) and more recent ones at the regional level e.g., by Mewes and Broekel (2022).[6] While we cannot include the GCC dummy variable and interaction term in the Fixed Effects estimations, we also do the analysis with the full sample and with oil-dependent countries separately. Across all models, we use and report robust standard errors, clustered at the country level, to avoid violations of the Ordinary Least Squares (OLS) assumptions.[7]

3.5 Results

3.5.1 Exploring the data

This subsection takes a first look at the economic complexity indicators. We focus firstly on the country-based measures and the GCCs. Following this, we explore the product complexity measures, in an attempt to understand further what might explain the changes observed in the GCCs over time.

Economic complexity across the GCCs

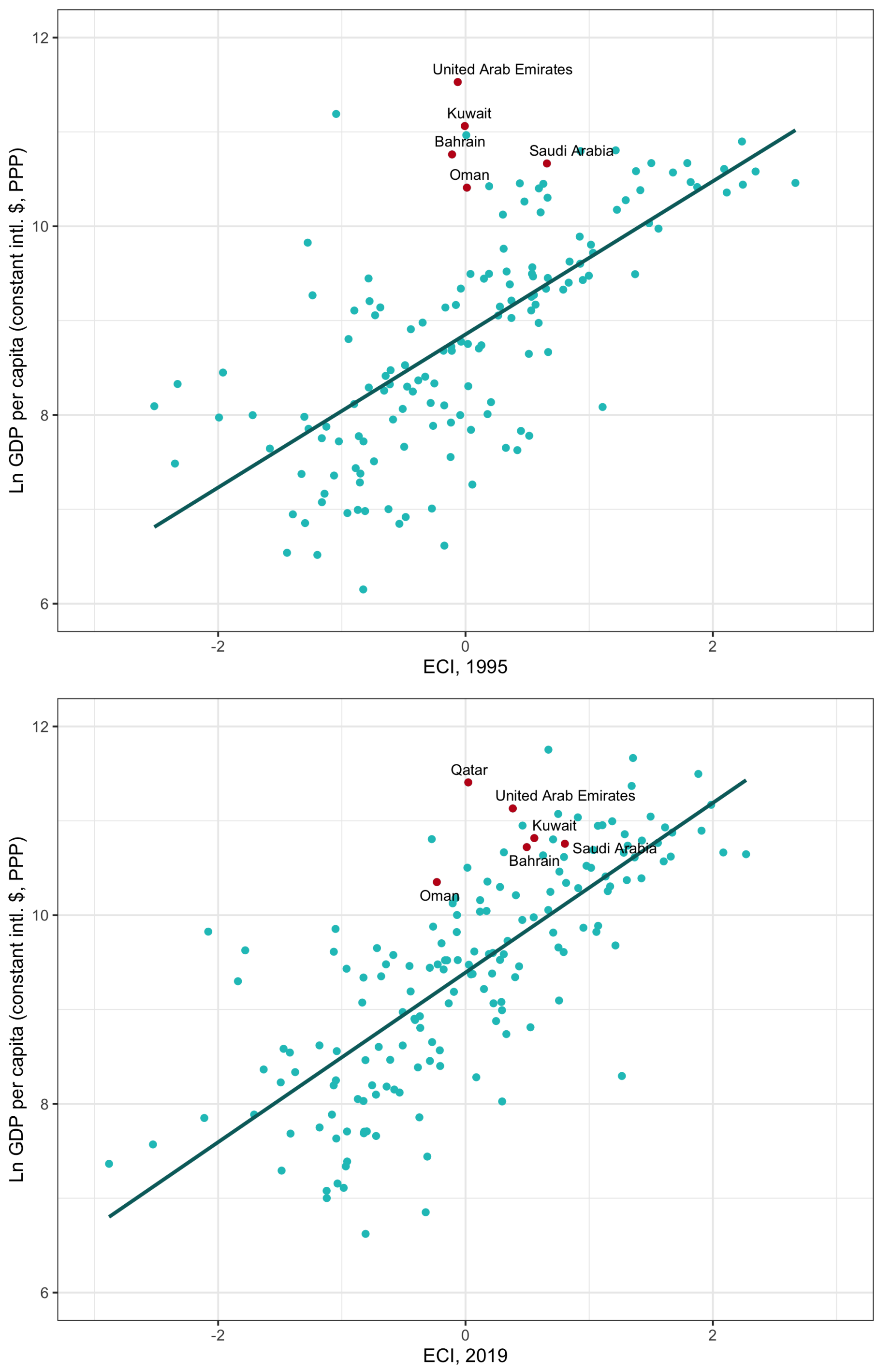
Figure 3.5 plots the correlations between the ECI and GDP per capita for 1995 and 2019. The GCCs have lower economic complexity than expected for their income level, particularly in 1995, as Hidalgo alluded to. Turning to oil rents, Figure 3.6 shows the scatter plots between the ECI and oil rents (as a percentage of GDP) in 1995 and 2019. In addition to the GCCs, countries with oil rents above 10% of GDP are labelled for comparison. The GCCs are marginally more complex than other, mostly lower income, countries with similar levels of oil dependence, showing the expected negative correlation between natural resource rents and levels of complexity. Kuwait stands out, with the highest level of oil rents within the GCC and, at the same time, a higher ECI than every other country with natural resource rents over 35 percent of GDP.

[5] The analysis starts in 2000 due to missing GDP data for Qatar until then. The use of averages to attenuate the impact of business cycles is common, and the time lengths used are in line with existing literature.

[6] We also ran the Hausman test, to check whether the Random Effects or Fixed Effects estimation is the most appropriate for our data; based on the test, we rejected the null hypothesis that the difference in coefficients is not systematic, and thus the Fixed Effects specification is preferred.

[7] To test for autocorrelation, we ran the Wooldridge test for autocorrelation in panel data – the null hypothesis of ‘no first-order autocorrelation’ was rejected at the 5% significance level, indicating the presence of serial correlation. Furthermore, due to our limited number of time periods (and larger number of countries), it is advisable to cluster the standard errors. This also allows for the violation of the homoskedasticity assumption.

Figure 3.5: ECI and GDP per capita, 179 countries, 1995 and 2019



Note: Qatar is missing from the first panel due to missing GDP data for 1995.

Figure 3.6: ECI and oil rents, 179 countries, 1995 and 2019

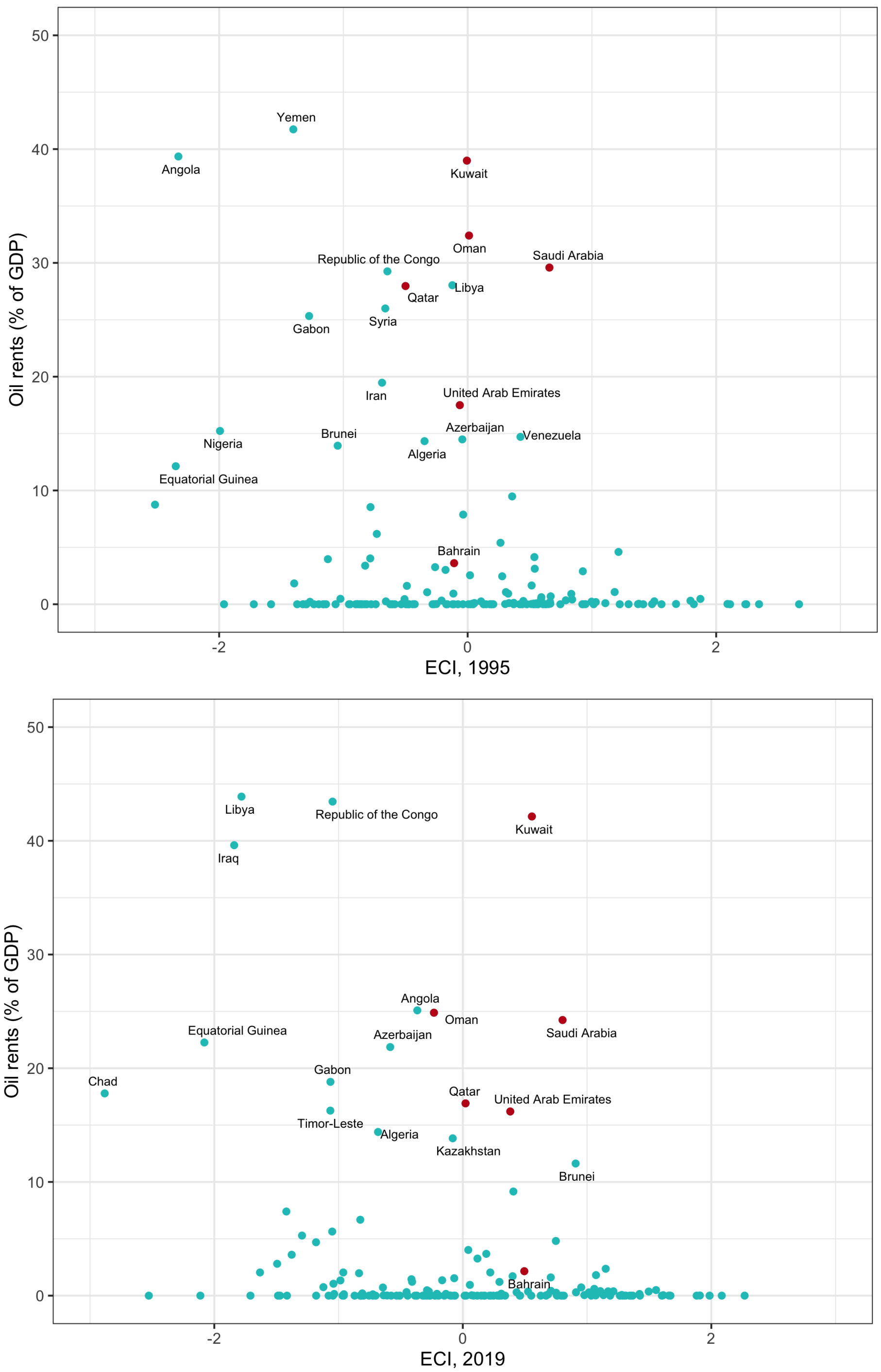


Figure 3.7 shows scatter plots of the ECI and employment in industry and services in 1995 and 2019. The patterns follow theoretical expectations, with a positive correlation between the ECI and employment in industry in both years (though weaker in 2019), and a positive correlation between the ECI and employment in services. Interestingly, not only is the positive correlation between ECI and employment in services stronger in 2019, but the GCC countries appear to be very closely aligned with the linear world trend, and very close to their expected level of employment in services given their ECI level. As the ECI measure is based on goods' export data, this positive correlation with employment in services may be reflecting development levels more broadly, as higher income countries tend to have higher shares of employment in services and to move structurally into service-led economies. Interestingly, Mishra et al. (2020), who incorporate services into ECI calculations, found that resource-rich countries improve in complexity rankings when services are added to the ECI, vis-a-vis measures relying on products only. This is perhaps unsurprising for the case of the GCCs, given the importance of employment in services, as previously discussed.

Figure 3.8 shows the correlations between the ECI and secondary and tertiary schooling separately, in 1995 and 2019. As expected, the correlation between the ECI and education variables is positive in all cases. For tertiary schooling, the enrolment level appears to be aligned with ECI level. Thus, overall, while we may expect higher ECI levels across the GCCs given their GDP per capita level, when it comes to other variables that may also reflect economic development (rather than simply growth or income), the GCCs appear closer to what would be expected given their ECI level, following Hidalgo's (2021) rationale.

Turning to the evolution of economic complexity, Figure 3.9 shows relatively high variation in ECI values over time in GCC countries. For example, between 2010 and 2012, Kuwait saw a drop in economic complexity from a high of 0.77 to a low of -0.47, followed by a recovery and more recent rise. Other GCCs experienced similarly volatile patterns, though not as sharp. From Figure 3.10, showing the dynamics of economic complexity country rankings, we can see the oscillation is present too and does not simply reflect changes in values that are also experienced by other countries. As reflected in the grey lines, the rest of the sample also experiences oscillations in values and rankings, with the latter remaining much more stable for countries with the highest complexity levels. The changes over time seem erratic, and therefore the next subsection explores further the potential drivers for these oscillations.

Figure 3.7: ECI and employment in industry and services, 1995 and 2019

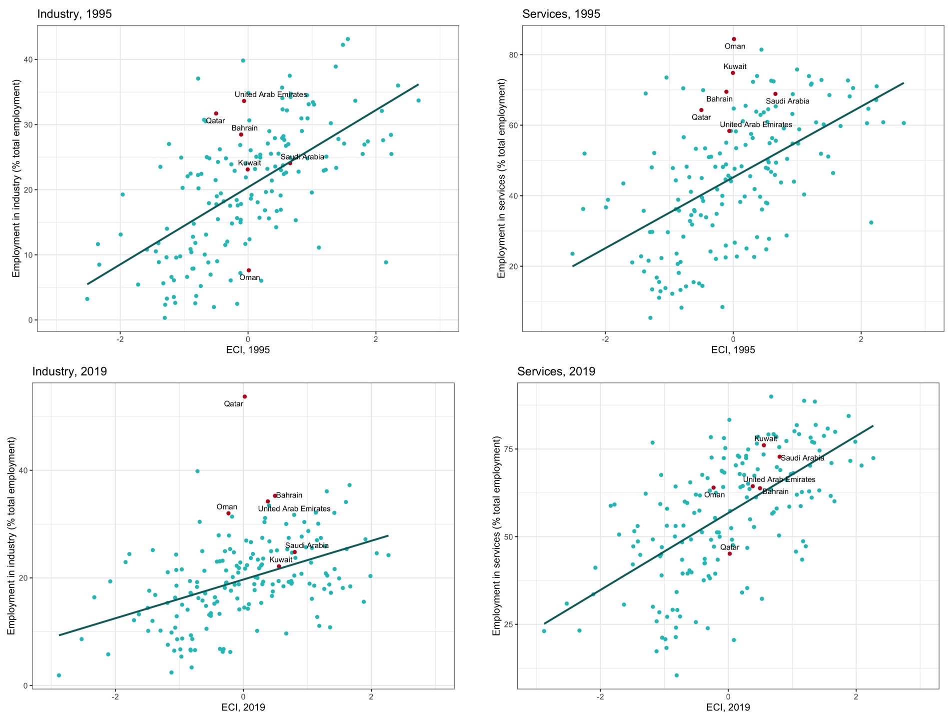


Figure 3.8: ECI and school enrolment (secondary and tertiary), 1995 and 2019

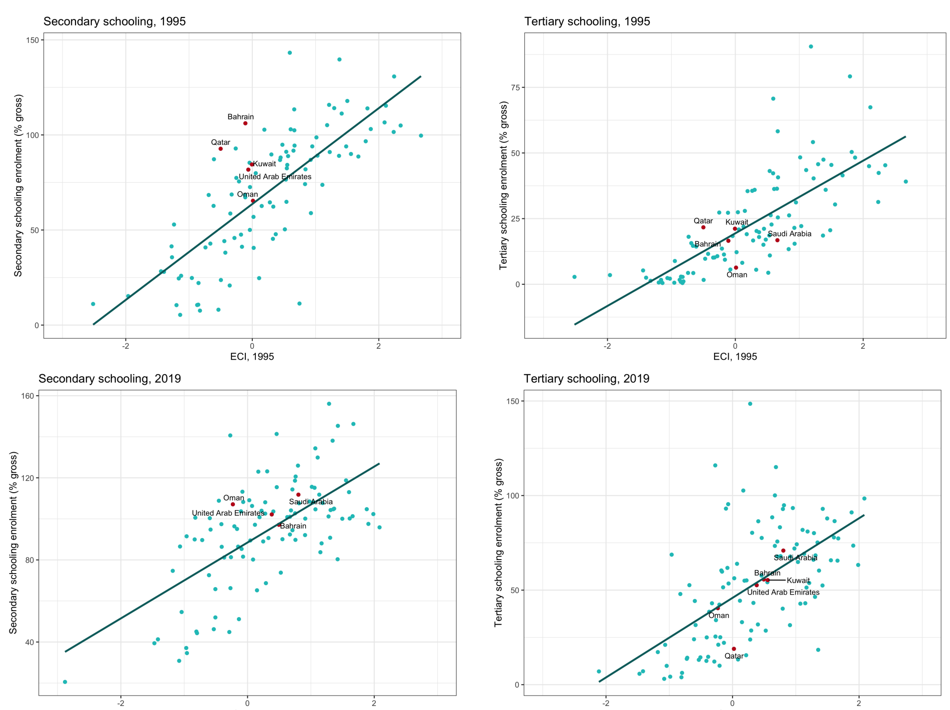


Figure 3.9: ECI value, 1995 to 2019, 179 countries (GCCs highlighted)

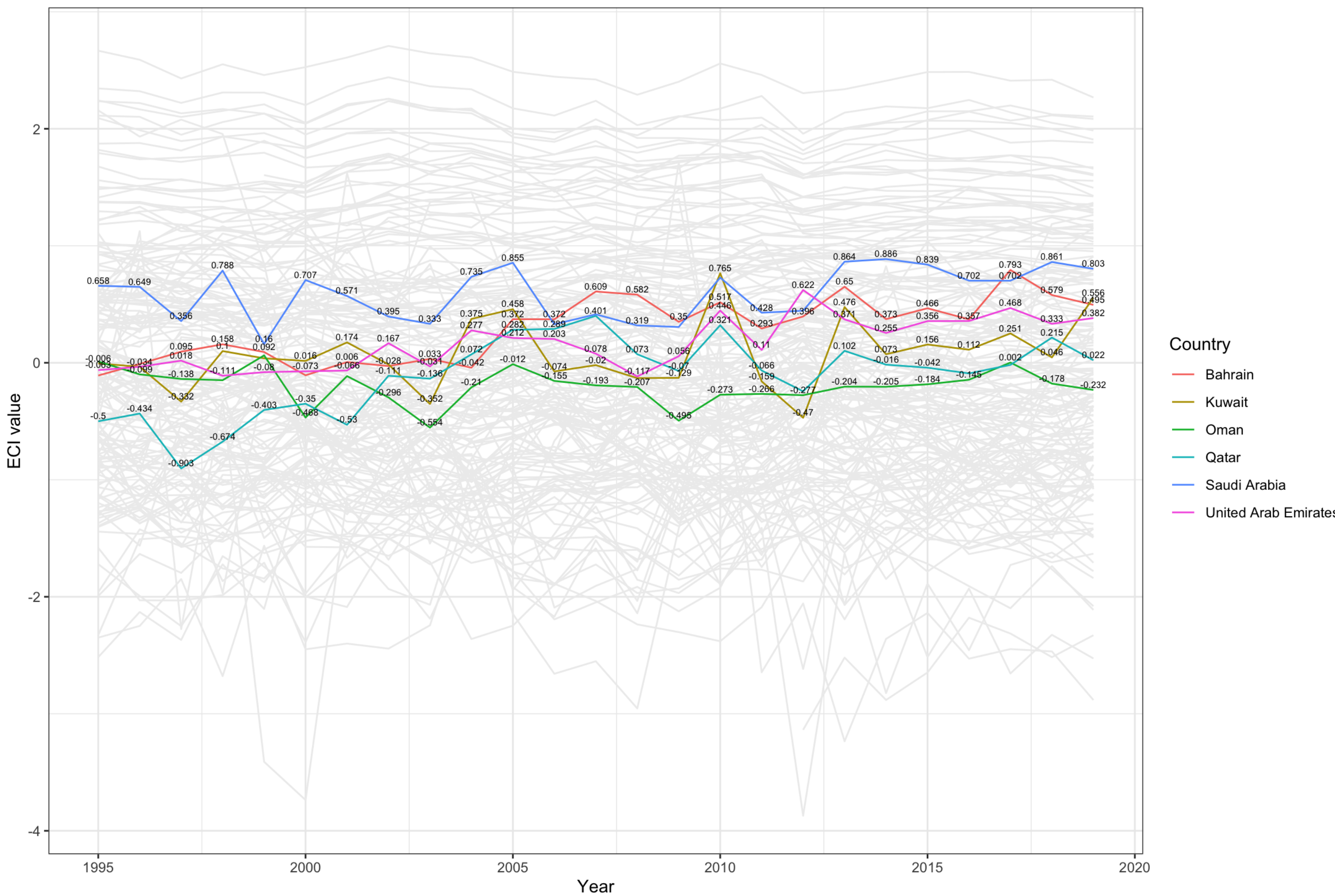
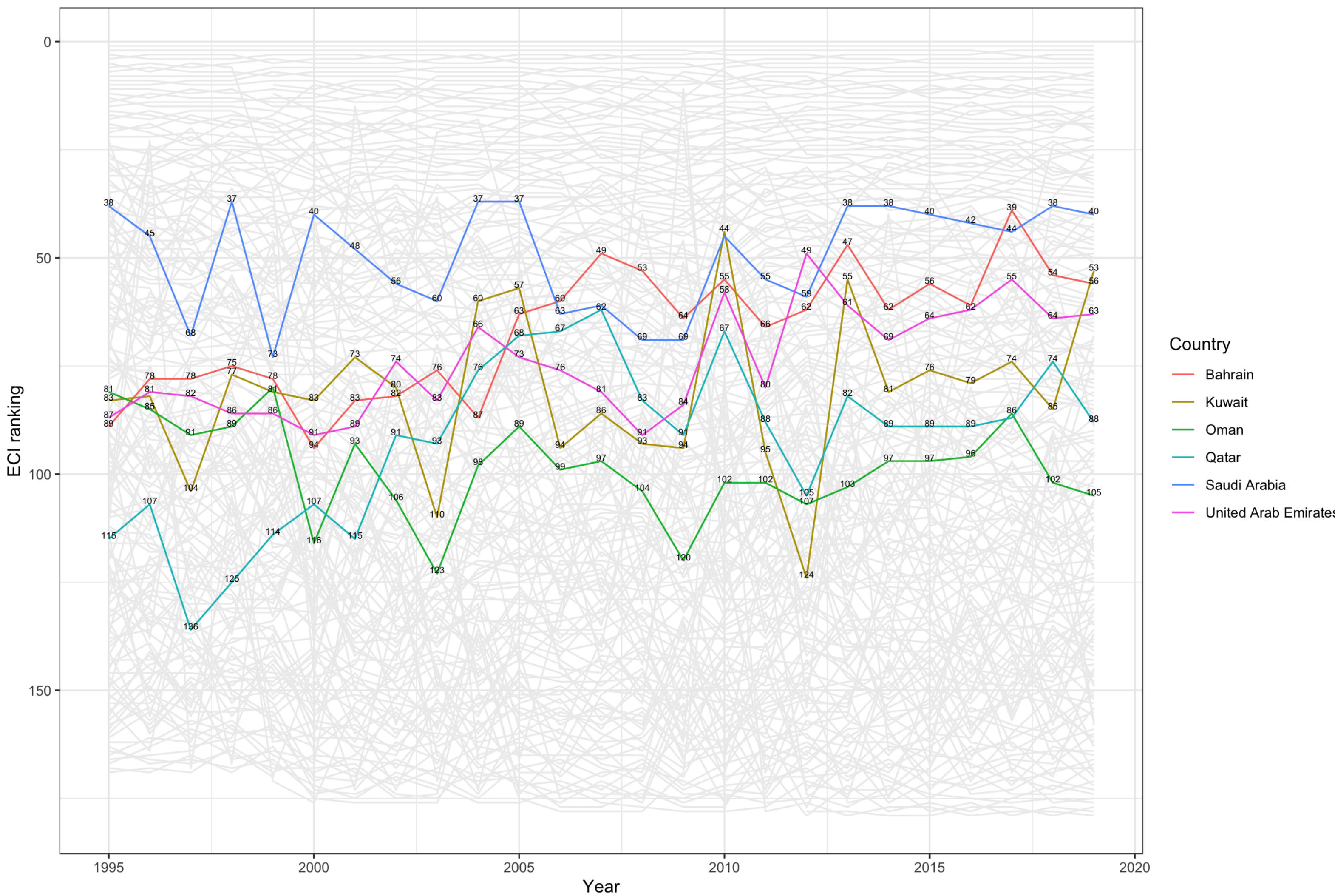


Figure 3.10: ECI rankings, 1995 to 2019, 179 countries (GCCs highlighted)



Understanding changes in economic complexity – the product side

To investigate further what may be driving the ECI fluctuations over time, this subsection takes a closer look at the product-based complexity measures, which are calculated analogously to the ECI for countries.

We start by looking at diversity, one of the two key variables behind the ECI measure, which is simply the total number of products that a country exports competitively relative to the rest of the world in a particular year (i.e., Balassa's (1965) Revealed Comparative Advantage, RCA index, using a threshold of 1). Figure 3.11 shows changes in diversity across the GCCs from 1995 to 2019. There are sharp oscillations in diversity across the time period, which are somewhat reflected in the ECI values and rankings, though not always perfectly aligned – e.g., Kuwait's ECI value dropped significantly from 2010 to 2011, whereas diversity only decreased in 2012. A similar pattern can be observed for the case of the UAE with sharp oscillations in both its ECI level and diversity throughout the time period. An interesting aspect to note is that, despite Saudi Arabia showing the highest complexity levels among the GCC countries for the majority of the years in our period of study, they are not the most diverse country, with the UAE showing the highest levels of diversity throughout the time period.

Looking at diversity in isolation does not tell us much about the types of products the GCCs are exporting competitively and how complex they are. Thus, we look at the Product Complexity Index, with a focus on the key oil and natural gas products that play a crucial role in these countries' economies, making up a large share of their exports. Figure 3.12 plots the PCI of all products, highlighting oil and natural gas products over the period. The biggest oscillations in the PCI are seen in petroleum gases (which includes natural gas) and, to a lesser extent, in crude petroleum oil. Looking at the evolution of the PCI for crude oil, we can see the changes over time are closely aligned with those seen in Kuwait's diversity levels.[8]

To investigate what may be behind the changes in PCI values, Figure 3.13 plots ubiquity – the number of countries that have a relative comparative advantage in a particular product each year – for the observed period, highlighting the oil and gas products. Among these, the product with the highest ubiquity is refined petroleum oils, with a significantly higher level than the remainder of the products, although with a decline towards the end of the period. This is followed by crude petroleum oils, with around 40 countries from 1995 to 2019,

[8]Petroleum jelly and petroleum coke are not shown in plots due to the very limited size of exports, but are considered in our data and plots when referring to oil and natural gas products.

and by petroleum gases with an average of 30 across the time period. The oscillations here are not as sharp as in diversity, partly because of the nature of the network used, which has a much higher number of products than countries. Despite having the highest ubiquity level of these three products, and being one of the most ubiquitous products in the sample overall (which we can see by comparing to the grey lines in the plot), refined petroleum oil has the highest PCI, suggesting that it may be exported by countries with higher complexity levels (i.e., those that are relatively diverse and export other products that relatively few other countries export).

To explore in more detail the products that each of the GCCs exports competitively and their complexity levels, we look at the top and bottom products in terms of PCI across the GCCs. Table 3.A.6 in the Appendix shows the top and bottom five products in terms of PCI, for each country, in 1995 and 2019. All six countries have at least one oil-related product among their bottom five products each year (except Bahrain in 1995).

In his assessment of diversification in the Gulf region, Hvidt (2013) finds that even after decades of diversification policies, hydrocarbon exports still play a significant role, with a lot of diversification pursued still being in the oil-related sector – both upstream (searching for and recovering hydrocarbons) and downstream (refining, selling and distributing hydrocarbons). This is evidenced in these top and bottom tables – Qatar and Kuwait both have hydrocarbon products among their most complex exports, possibly due to these exports' low ubiquity and, while related to oil dependence, these products may well be increasing the countries' overall complexity level given their current specialisation patterns.

Overall, this analysis raises questions on what the ECI and PCI measures capture for the GCCs and how influenced they are by these countries' high dependence on oil and natural gas. Changes in economic complexity levels experienced in the GCC group may be driven by oil-related goods' shifts, particularly in terms of a drop in the relative complexity of such products. This points to some limitations of complexity indicators in the case of oil-dependent economies, since in 'absolute' terms there is no reason to believe that the capabilities needed to produce oil-related products (or their 'complexity' level) suddenly decreased only for a limited number of years, relative to all other products. Thus, we next explore further the association between economic complexity and economic growth for the case of GCC and oil-dependent countries.

Figure 3.11: Diversity (number of products with $RCA > 1$), GCCs, 1995 to 2019

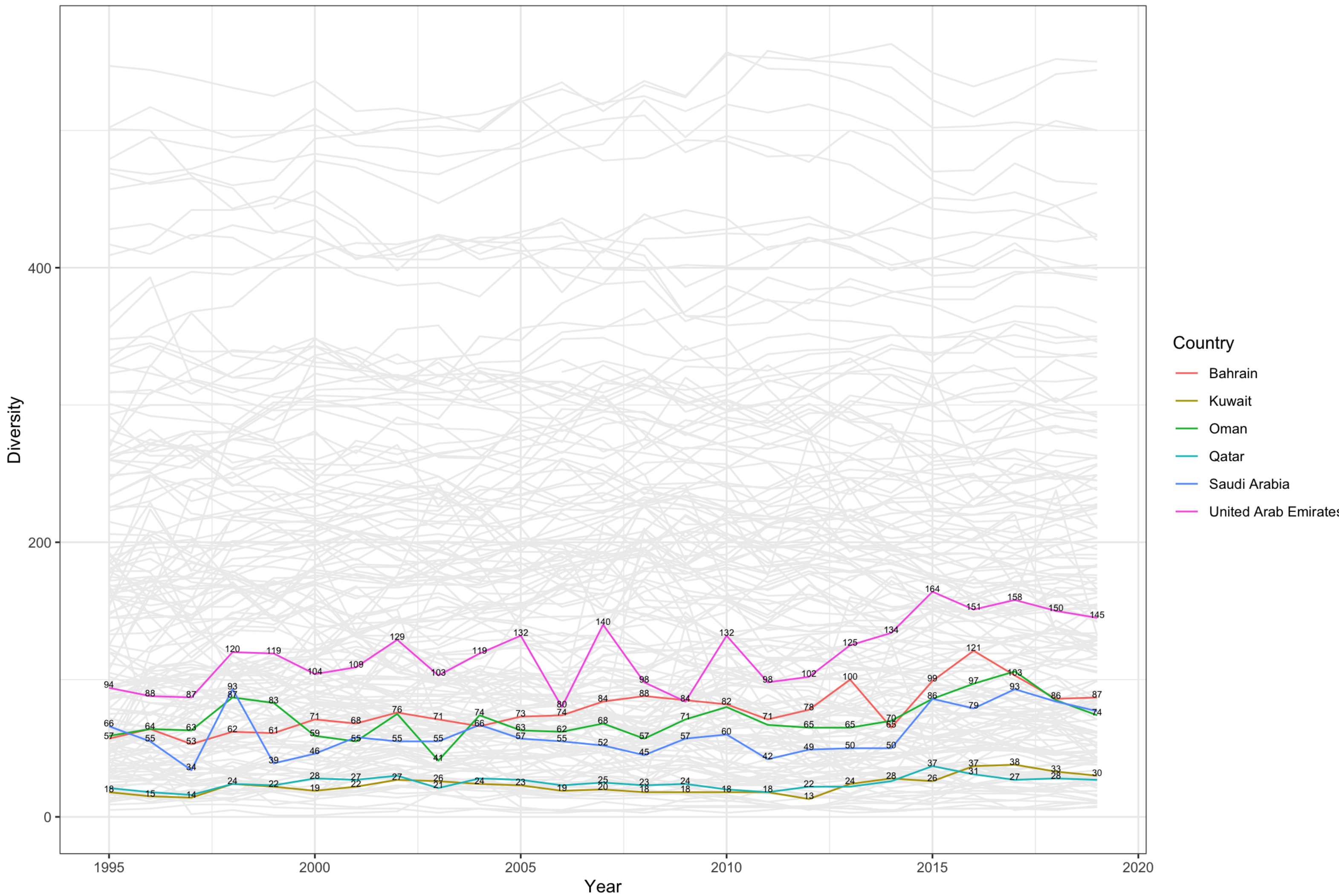


Figure 3.12: Product Complexity Index, oil and gas products highlighted, 1995 to 2019

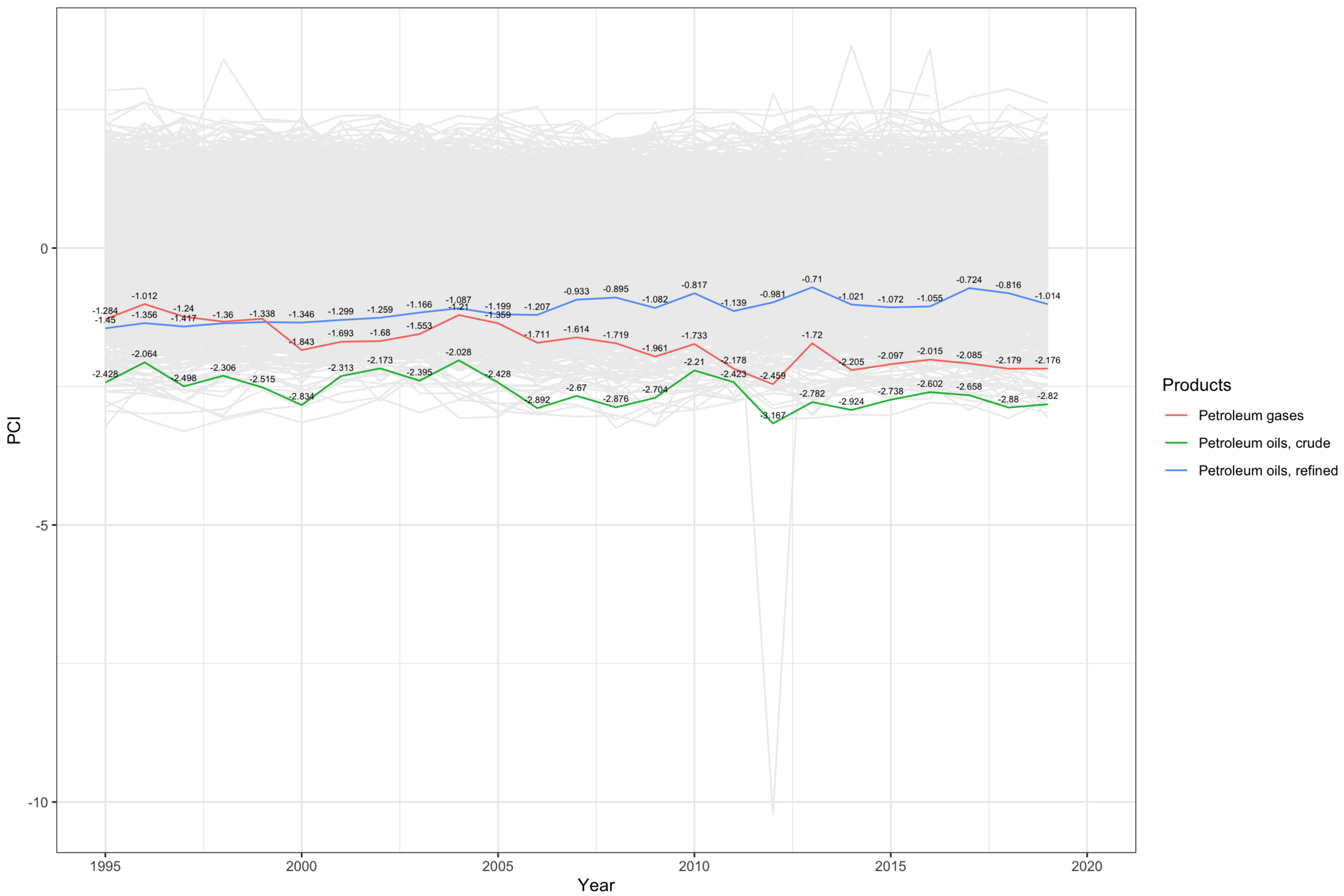
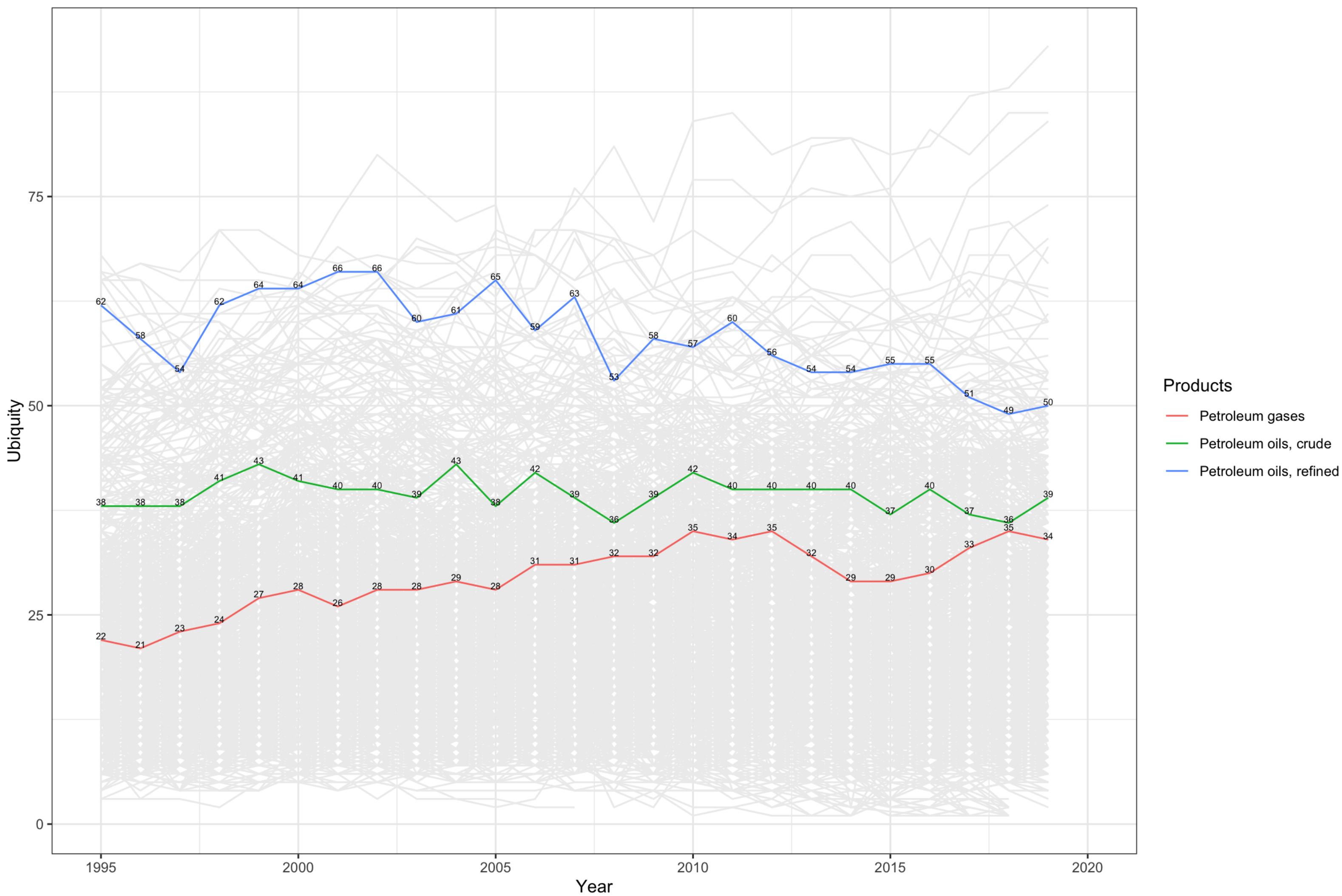


Figure 3.13: Ubiquity (number of countries with $RCA > 1$), oil and gas products highlighted, 1995 to 2019



3.5.2 Empirical estimation

ECI and economic growth

Our starting point is to replicate the most common analysis in the economic complexity literature, of regressing economic growth on the initial ECI and income level. We do this for different time lengths – 20-, 10- and 5-year growth periods – from 2000 to 2019. In each case, we start from the full model and, following this, explore how much of the variation in growth across countries is explained by the ECI.[9]

Starting from the 20-year growth period, Table 3.2 presents the cross-section results, while Figure 3.14 shows a graphical illustration of this association, with a partial regression scatter plot between GDP per capita growth and the initial ECI conditional on initial GDP per capita.[10] Both show a positive association between the initial ECI and income growth, which is statistically significant across all the model specifications. When the ECI is removed from the model in column (5) in Table 3.2, there is a drop in the adjusted R-squared from 0.344 to 0.248, indicating that 9.6% of the variance in economic growth that is not accounted for by initial income and increase in exports is explained by the ECI.

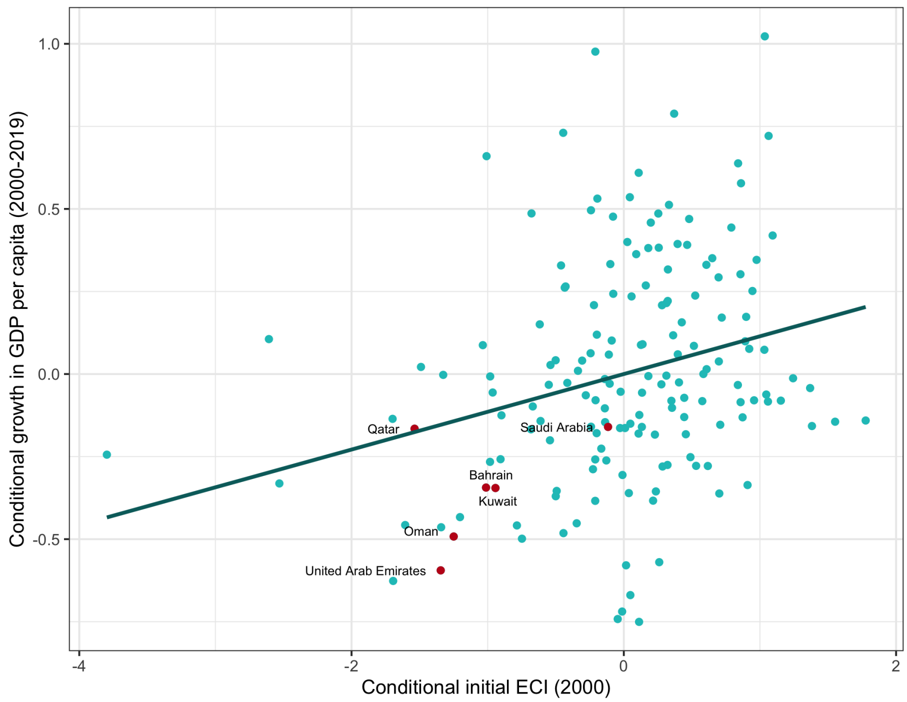
With regards to the control variables, as expected there is a negative and statistically significant coefficient on initial GDP per capita across all specifications, as well as a positive coefficient on the increase in exports as a share of initial income. The other control variables – the increase in natural resource exports (as a share of initial income) and population – do not appear statistically significant and, as columns (2) and (3) show, they have a negligible impact on the overall variance explained by the model. The coefficient on the dummy variable for GCC countries is negative and statistically significant − being a GCC country is associated with a lower GDP per capita compared to the rest of the sample – and its interaction term with the ECI is positive and statistically significant.

Turning to the 10-year growth periods, Table 3.3 presents the results. The first five columns show the cross-country Pooled OLS estimation, following the same specifications as before – the results are aligned with those for the 20-year growth period, with the exception of the interaction term between the GCC dummy variable and the ECI, which is not statistically significant. The Fixed Effects estimation, in columns (6) and (7), analyses the association

[9]To avoid concerns over specific variables, we also introduced each additional variable individually to investigate its impact on the association between ECI and growth over the time period (available upon request).

[10]Both GDP per capita growth and the initial ECI are first regressed on the initial level of GDP per capita, and the residual of these regressions are used in the plot, thus capturing only the variation in the two variables that cannot be accounted for by the initial income level.

Figure 3.14: GDPpc growth (2000-2019) and initial ECI, conditional on initial GDPpc



between economic complexity and income growth within countries. Unlike some existing findings, our analysis shows no association between the ECI and growth in GDP per capita, regardless of the control variables included in the models. The coefficients for the control variables follow the same patterns as before, though in this case the increase in natural resource exports is the most dominant export-based variable.[11]

Table 3.4 shows regression results for the 5-year growth period, mirroring the previous specifications, though in this case we do not remove the increase in natural resource exports, as it is statistically significant and has larger coefficients than the increase in exports. The results are otherwise fully aligned with those for the 10-year periods.

Overall, our results confirm a positive association between ECI and income growth across countries, but in contrast with existing research, we do not find such an association for changes within countries. The interaction term between the GCC dummy variable and the initial ECI was statistically significant only in the 20-year growth regression, suggesting that the positive association between the ECI and growth in GDP per capita does not differ for GCC countries vis-a-vis the rest of the sample, and that in the long term such an association is stronger for them.

[11] To attenuate concerns over the results being driven by the 10-year periods selected, for instance due to the 2008 financial crisis, we tried alternative cut-offs; in each case the results were in line with the ones presented here (available upon request).

Table 3.2: Economic complexity and 20-year growth in GDP per capita, full sample

	GDPpc growth (2000-2019)				
Variables	(1)	(2)	(3)	(4)	(5)
Initial ECI	0.115***	0.104***	0.117***	0.143***	
	(0.0296)	(0.0274)	(0.0279)	(0.0287)	
Initial GDPpc (log)	-0.113***	-0.105***	-0.116***	-0.150***	-0.0796***
	(0.0309)	(0.0275)	(0.0281)	(0.0282)	(0.0222)
Increase in exports	0.119***	0.135***	0.136***	0.127***	0.113***
	(0.0289)	(0.0230)	(0.0225)	(0.0209)	(0.0203)
Increase in NR exports	0.0438				
	(0.0464)				
Exports to GDP (initial)	0.0348				
	(0.124)				
Initial population (log)	0.0221	0.0211			
	(0.0155)	(0.0150)			
GCC	-0.419***	-0.404***	-0.408***		
	(0.0797)	(0.0775)	(0.0748)		
GCC * Initial ECI	0.245***	0.228***	0.267***		
	(0.0769)	(0.0685)	(0.0686)		
Constant	0.960**	0.915**	1.342***	1.643***	1.027***
	(0.379)	(0.367)	(0.259)	(0.258)	(0.213)
Observations	164	164	164	164	164
R-squared	0.412	0.408	0.399	0.357	0.257
Adjusted R-square	0.382	0.385	0.380	0.344	0.248

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.3: Economic complexity and 10-year growth in GDP per capita, full sample

	GDPpc growth (2000-2009 and 2010-2019)						
	Pooled OLS					Fixed Effects	
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Initial ECI	0.0509***	0.0515***	0.0583***	0.0694***		-0.0104	-0.0829
	(0.0166)	(0.0173)	(0.0175)	(0.0169)		(0.0273)	(0.0519)
Initial GDPpc (log)	-0.0542***	-0.0561***	-0.0623***	-0.0765***	-0.0385***	-0.548***	-0.767***
	(0.0165)	(0.0163)	(0.0167)	(0.0162)	(0.0108)	(0.0740)	(0.114)
Increase in exports	0.214***	0.214***	0.213***	0.211***	0.192***	0.0398	
	(0.0452)	(0.0231)	(0.0227)	(0.0220)	(0.0242)	(0.0361)	
Increase in NR exports	0.00652					0.171***	
	(0.0522)					(0.0469)	
Exports to GDP (initial)	-0.0353					-0.0874	
	(0.0667)					(0.0647)	
Initial population (log)	0.0128*	0.0134*				0.0788	
	(0.00730)	(0.00718)				(0.105)	
GCC	-0.202***	-0.202***	-0.207***				
	(0.0532)	(0.0519)	(0.0507)				
GCC * Initial ECI	0.145	0.140	0.152				
	(0.106)	(0.105)	(0.107)				
Constant	0.431**	0.429**	0.698***	0.819***	0.486***	3.897*	7.135***
	(0.193)	(0.192)	(0.152)	(0.147)	(0.103)	(2.033)	(1.025)
Observations	332	332	332	332	332	332	332
R-squared	0.365	0.364	0.354	0.332	0.279	0.722	0.543
Adjusted R-square	0.347	0.350	0.342	0.323	0.272	0.716	0.539
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	No	No	No	No	No	Yes	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3.4: Economic complexity and 5-year growth in GDP per capita, full sample

	GDPpc growth (2000-2004, 2005-2009, 2010-2014, 2015-2019)						
	Pooled OLS					Fixed Effects	
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Initial ECI	0.0247***	0.0252***	0.0282***	0.0329***		0.00712	-0.00810
	(0.00850)	(0.00861)	(0.00882)	(0.00839)		(0.0112)	(0.0213)
Initial GDPpc (log)	-0.0253***	-0.0264***	-0.0291***	-0.0363***	-0.0176***	-0.222***	-0.321***
	(0.00799)	(0.00786)	(0.00813)	(0.00791)	(0.00490)	(0.0315)	(0.0443)
Increase in exports	0.203***	0.195***	0.191***	0.196***	0.213***	0.115**	
	(0.0654)	(0.0564)	(0.0567)	(0.0568)	(0.0592)	(0.0567)	
Increase in NR exports	0.241***	0.244***	0.246***	0.226***	0.162**	0.314***	
	(0.0693)	(0.0658)	(0.0666)	(0.0659)	(0.0680)	(0.0716)	
Exports to GDP (initial)	-0.0204					0.0296	
	(0.0303)					(0.0438)	
Initial population (log)	0.00496	0.00526				-0.0521	
	(0.00354)	(0.00342)				(0.0393)	
GCC	-0.102***	-0.103***	-0.105***				
	(0.0221)	(0.0223)	(0.0212)				
GCC * Initial ECI	-0.0146	-0.0149	-0.00942				
	(0.0455)	(0.0455)	(0.0470)				
Constant	0.212**	0.212**	0.320***	0.382***	0.213***	2.877***	2.990***
	(0.0898)	(0.0902)	(0.0737)	(0.0723)	(0.0467)	(0.752)	(0.401)
Observations	668	668	668	668	668	668	668
R-squared	0.318	0.317	0.312	0.287	0.251	0.470	0.230
Adjusted R-square	0.306	0.306	0.302	0.280	0.244	0.463	0.224
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	No	No	No	No	No	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

ECI and economic growth in oil-dependent countries

We saw above that the ECI has a positive association with economic growth across countries that does not appear to differ for the case of the GCCs – here we investigate this association for oil-dependent countries. Table 3.5 replicates our analysis with ten-year growth periods for this sub-sample. The positive association between the ECI and economic growth is confirmed, in some of the specifications showing a much stronger coefficient than in the case of the full sample (along with higher adjusted R-squares). Turning to the within-country association, once again no significant association emerges between initial ECI and growth in GDP per capita. Results follow the same patterns for both the 20-year and 5-year growth regressions.

Overall, and generally in line with existing research, economic complexity shows a positive association with growth in GDP per capita across the full sample, including the GCCs and oil-dependent countries alike. Nevertheless, once we look at within-country associations, we do not find a statistically significant association between the ECI and economic growth.

While the latter finding contradicts the existing literature, it is perhaps not entirely surprising, given the oscillations observed in the descriptive analysis. Furthermore, even though the association between economic complexity and economic growth is in line with the existing literature, some of our concerns are still present – from the research context, we know that the GCCs had relatively low (and in some cases negative) economic complexity levels and that they observed a decrease in GDP per capita (experiencing negative economic growth, though due to rising population) over the period. Thus, overall, questions remain on the limitations of the ECI concept and measure in the GCCs and other oil-dependent countries. The next section explores this question further, by looking at the links between economic complexity and oil and natural gas dependence.

Table 3.5: Economic complexity and 10-year growth in GDP per capita, oil-dependent countries

	GDPpc growth (2000-2009 and 2010-2019)						
	Pooled OLS					Fixed Effects	
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Initial ECI	0.119***	0.113***	0.116***	0.118***		0.0110	-0.111
	(0.0302)	(0.0289)	(0.0321)	(0.0340)		(0.0529)	(0.0859)
Initial GDPpc (log)	-0.104**	-0.104**	-0.119***	-0.145***	-0.0988***	-0.655***	-1.161***
	(0.0500)	(0.0412)	(0.0415)	(0.0368)	(0.0260)	(0.132)	(0.164)
Increase in exports	-0.135	0.215***	0.207***	0.207***	0.182***	0.0475	
	(0.301)	(0.0311)	(0.0265)	(0.0269)	(0.0291)	(0.477)	
Increase in NR exports	0.368					0.146	
	(0.312)					(0.514)	
Exports to GDP (initial)	0.0178					-0.221*	
	(0.185)					(0.111)	
Initial population (log)	0.0263	0.0248				0.223*	
	(0.0158)	(0.0177)				(0.111)	
GCC	-0.0891	-0.113*	-0.126*				
	(0.0695)	(0.0637)	(0.0632)				
GCC * Initial ECI	0.0817	0.101	0.133				
	(0.108)	(0.108)	(0.107)				
Constant	0.725	0.730	1.266***	1.483***	1.001***	2.808	11.04***
	(0.589)	(0.589)	(0.395)	(0.361)	(0.250)	(2.421)	(1.511)
Observations	60	60	60	60	60	60	60
R-squared	0.674	0.667	0.655	0.641	0.547	0.907	0.804
Adjusted R-square	0.615	0.622	0.616	0.615	0.523	0.894	0.794
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	No	No	No	No	No	Yes	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

3.6 Further analysis – ECI and oil dependence

In this section, we investigate the links between economic complexity measures and oil dependence in the GCC countries. First, we explore the ECI excluding oil and natural gas products; second, we analyse the correlation between the different ECI measures, income and oil-related variables.

Excluding oil and natural gas products from ECI calculations

As described in Section 4.1, to explore the impact that oil and natural gas products have in complexity measures for the GCC countries, we constructed an alternative measure, based on a network that excludes oil and natural gas products.

The ECI is based on a binary matrix of countries and products, based on RCA calculations, taking the threshold of 1.[12] In the RCA calculation itself, total exports enter twice in the denominator. Since oil and related products make up a very large share of these countries' exports, when there is an increase in the value of those exports (which could be caused by both an increase in oil export volume or price), the denominator will increase a lot and it will become harder for countries to achieve a threshold of 1 in the RCA, leading to a decrease in diversity, and ultimately to a lower ECI level would be expected. Thus, removing the five key oil and natural gas products from the ECI calculations allows us to investigate what happens to ECI levels and changes over time in the GCC countries, as well as to explore the correlations with other key variables.

Figures 3.15 and 3.16 show the ECI values and rankings over time respectively, mirroring the previous plots. Here, we can see that the GCC countries tend to show higher levels of ECI on average and that, while there are still some oscillations, they are not as sharp as they were before, particularly if we look at the examples of Kuwait and Saudi Arabia (whereas Qatar, for example, still experiences significant fluctuations); the rankings reflect the same patterns.

In addition, Figure 3.17 plots diversity, calculated from the network excluding oil products over time for the GCCs. As we are simply looking at the number of products in which countries are competitive, based on the $RCA \geq 1$ cut-off, and given the mechanism outlined above, there are sharp differences between this plot and our original one, as expected. Diversity is higher overall across all GCC countries. While the UAE displays again the

[12] Formally, where X_{cp} represents the exports of product p by country c, the RCA that country c has in product p is expressed as: $RCA_{cp} = \frac{X_{cp}}{\sum_c X_{cp}} / \frac{\sum_p X_{cp}}{\sum_{c,p} X_{cp}}$

highest diversity, Kuwait is no longer at the bottom of this group (and is much further away from the bottom vis-a-vis the rest of the world, plotted in grey), and Saudi Arabia has also seen a relative increase compared to the GCC group; in contrast, Bahrain saw a relative decrease, whilst Qatar shows the lowest levels of diversity (as it did previously, along with Kuwait).

Overall, the descriptive analysis of these plots points to some important differences, particularly for the GCCs with the highest oil dependence, such as Kuwait and Saudi Arabia. Nevertheless, this does not appear to always be the case – for instance, Qatar also has a very high share of oil and natural gas exports (0.88 in 2019, as shown in the country summary tables) and while it had an increase in diversity when oil was excluded from the indicators, it still remains relatively less diverse vis-a-vis the other GCCs.

Oil, economic complexity and income

Finally, we explore the correlation between both ECI measures and key variables in the GCC group. Figure 3.18 shows the correlation values and scatter plots between the ECI, ECI excluding oil, GDP per capita, oil and gas exports share and oil price. In each case, it shows the overall correlation and the correlation in each country.

Unsurprisingly, there is a strong positive correlation between the standard ECI and the ECI measure excluding oil and natural gas products. This is the case for the overall correlation as well as for each individual country, with the exception of Kuwait where the correlation is nearly nonexistent.

Regarding the ECI and GDP per capita, the picture is more mixed, with a positive correlation in some countries and a negative one in others; in the most oil-dependent countries – Qatar and Kuwait – there is no correlation. A similarly mixed pattern emerges for the correlation between the ECI excluding oil and GDP per capita – once again, there are countries with positive and negative correlations and overall there appears to be no correlation for the group; interestingly, both Qatar and Kuwait have positive correlations, in contrast to what we saw for the standard ECI.

Turning to oil dependence, a positive relationship emerges between both ECI measures and the oil and gas exports share and oil price; for the standard ECI measure, this is particularly strong in Bahrain and Qatar with both oil variables. To explore more clearly the links between the the ECI and crude oil price over time, Figure 3.A.4 in the Appendix shows the ECI (values in the primary Y-axis) and the price of crude oil (values in the secondary Y-axis), over the time period and Figure 3.A.5 in the Appendix shows scatter plots between crude oil price and ECI in different years, with each of the GCC countries in a separate panel. From these, we can see that there is some association between oil prices and oscillations in the ECI across these countries, though it is not always clear-cut – in Bahrain, Qatar and the UAE, there appears to be a positive correlation in the trends over time, whereas for Kuwait, Saudi Arabia and Oman there is no correlation.

Overall, while we found a cross-country association between the ECI and subsequent economic growth in the previous section, there is still a substantial impact on the ECI from oil price volatility and from having a large share of exports in oil and related products in the GCCs. This provides some support towards our hypothesis that the ECI may not be accurately reflecting capabilities in economies heavily reliant on natural resources, particularly those with highly volatile prices.

Figure 3.15: ECI (excl. oil) value, 1995 to 2019, 179 countries (GCCs highlighted)

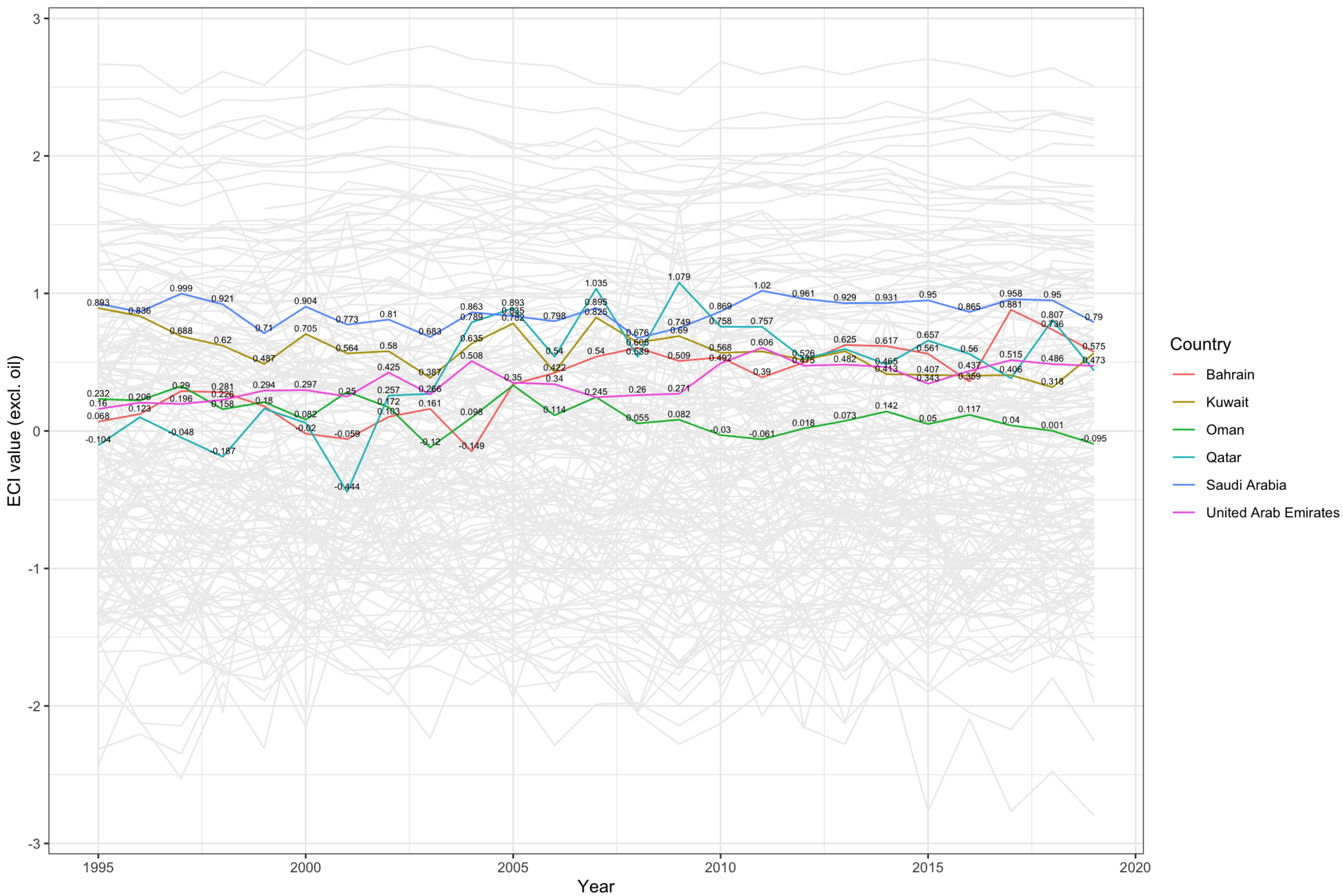


Figure 3.16: ECI (excl. oil) rankings, 1995 to 2019, 179 countries (GCCs highlighted)

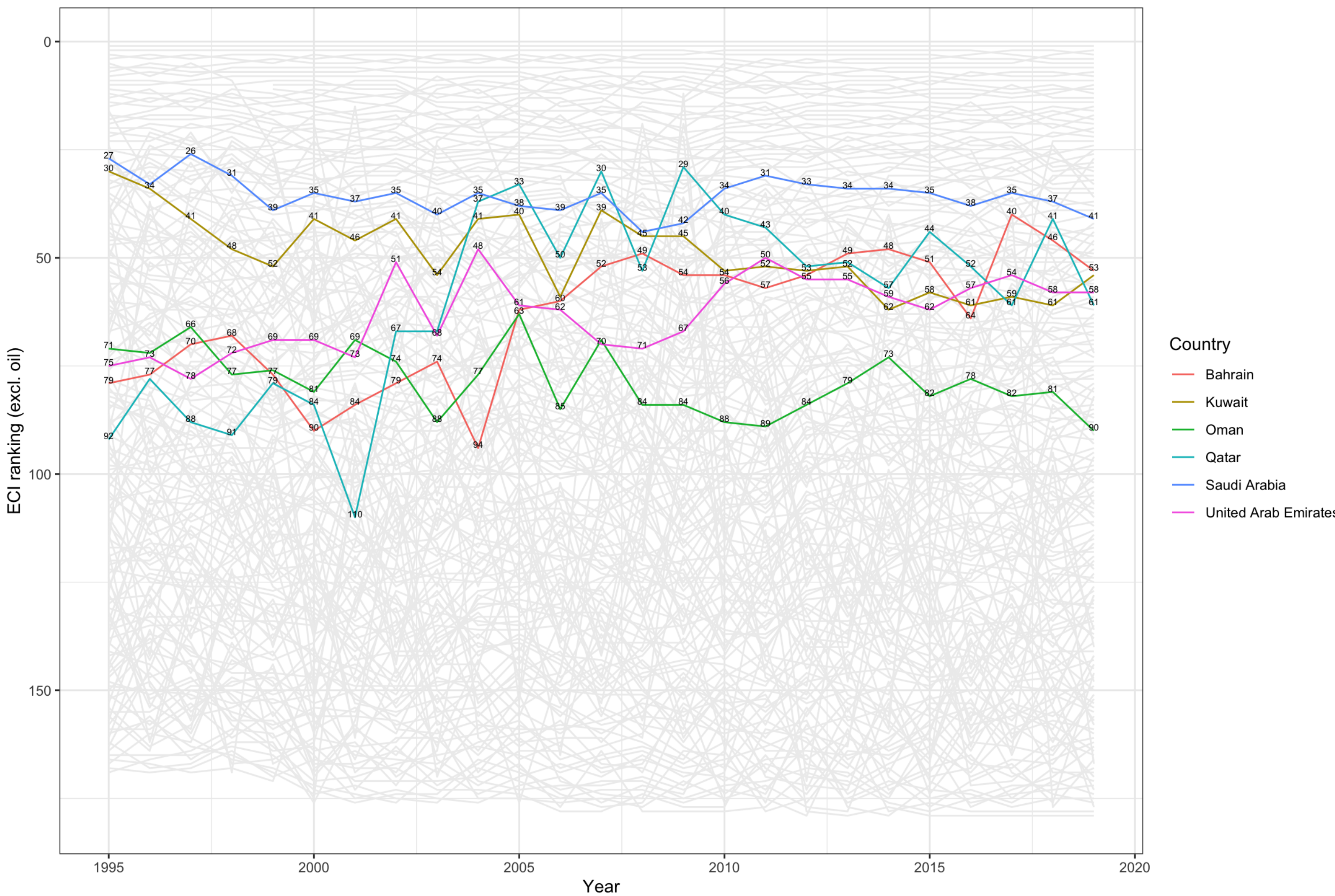


Figure 3.17: Diversity (excl. oil), GCCs, 1995 to 2019

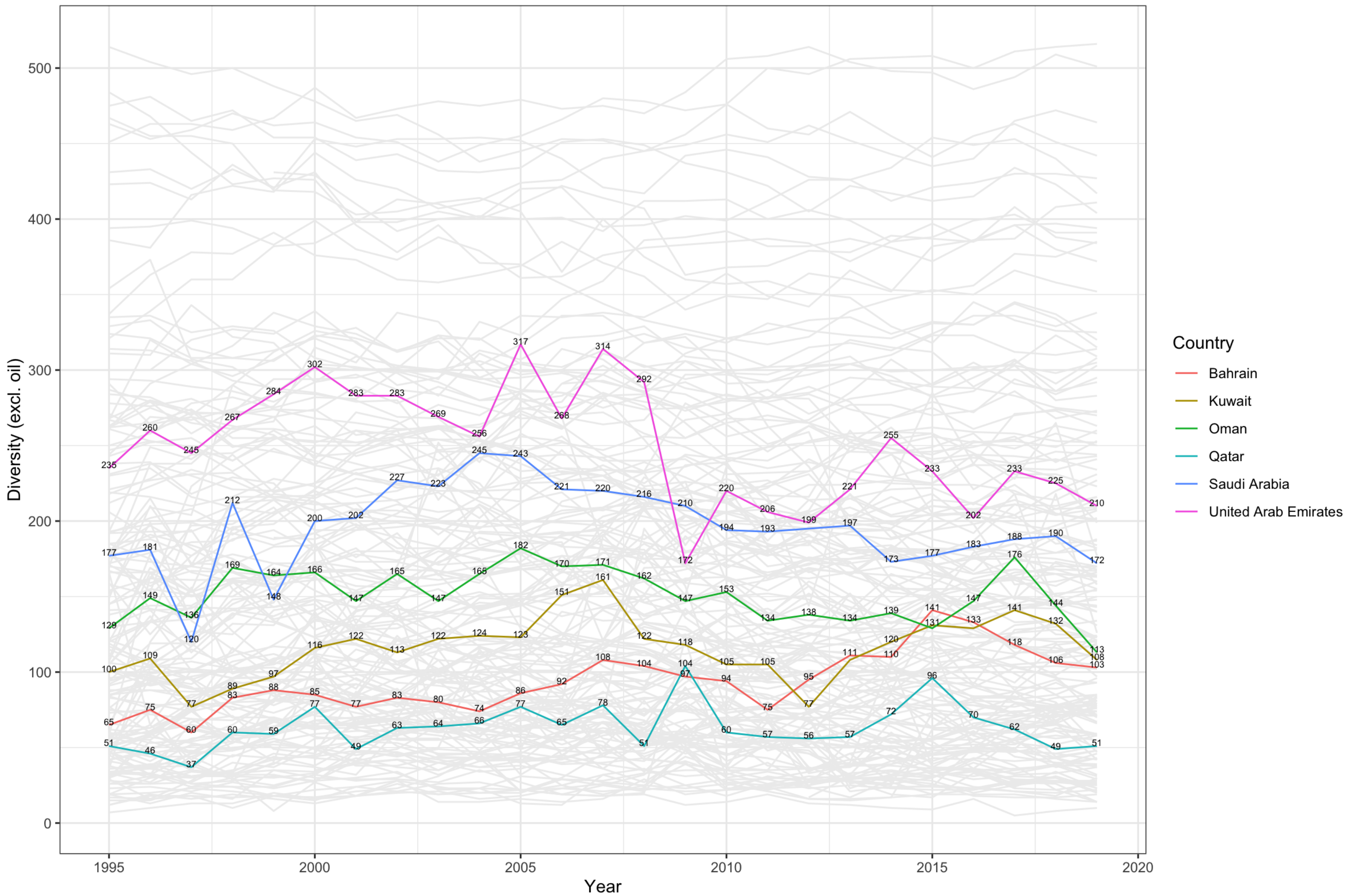
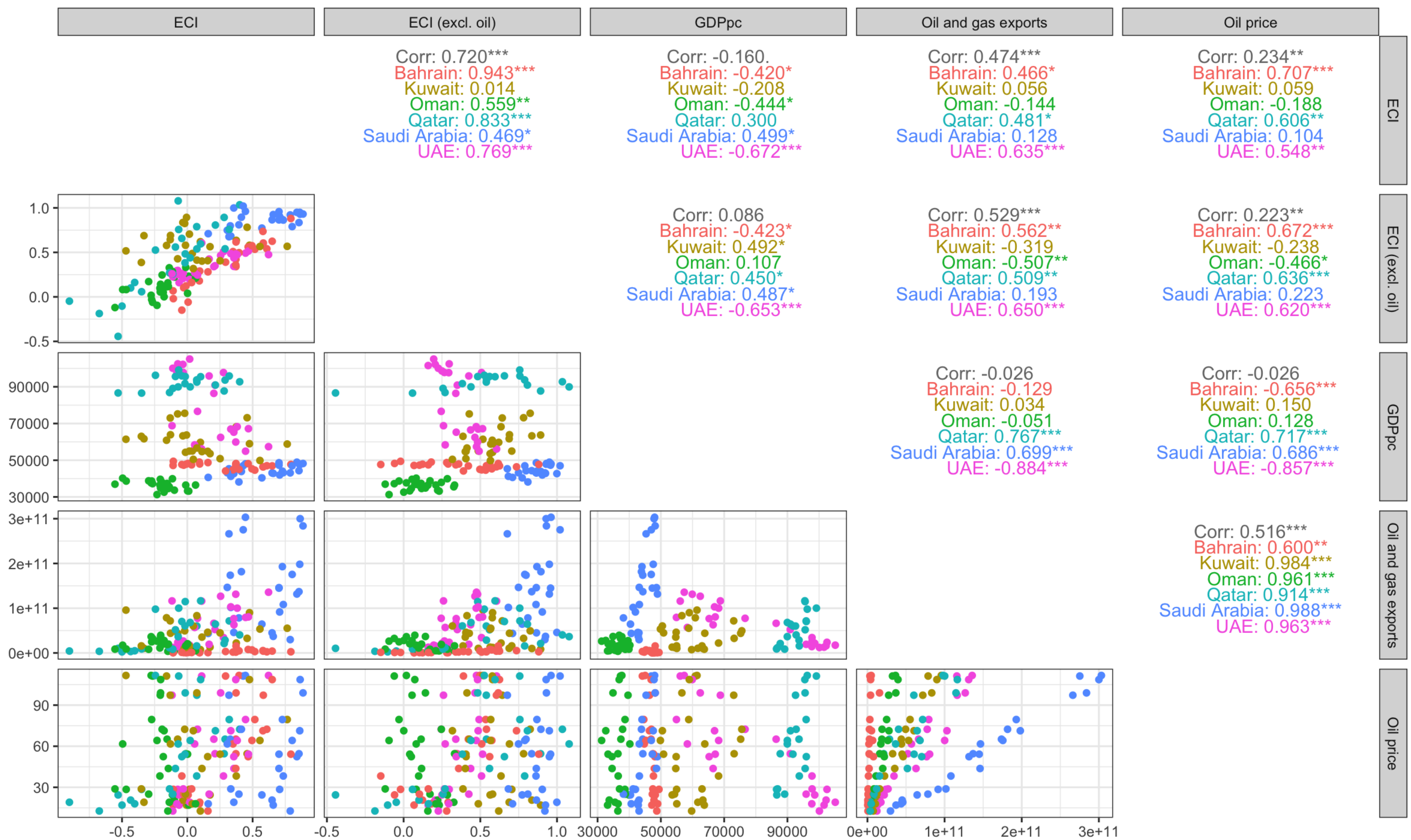


Figure 3.18: Correlations and scatter plots between selected variables, 1995 to 2019



3.7 Discussion

Policy implications

Even though there is an association between the ECI and economic growth across countries, including for the GCCs and other oil-dependent countries, there is still evidence of different ways in which the ECI is affected by the large dependence on oil and natural gas in these countries. This, along with the lack of explanatory power of the ECI for within-country changes in economic growth, presents important challenges for policy.

For policymakers, tracking progress over time is vital – beyond simply assessing the relative position vis-a-vis all world economies, policymakers want and need to understand changes over time within their own country. The oscillations observed in the ECI, along with the impact that oil and related products have in economic complexity levels, means that changes over time may be meaningless in oil-dependent countries due to the volatility in oil prices, along with changes in demand and political aspects that have a big impact on oil exports.

Overall, even if some specific issues with the ECI can be addressed – there have been previous attempts, for example, of measuring the ECI with value-added exports in order to better capture underlying capabilities by Koch (2021) – there are broader important questions unanswered in the context of ECI measures, such as the long-standing question in economic development of 'how'. In particular, policy and practical implications derived from ECI analyses may be too high level and not place-based – even if we know that diversification is beneficial, and that countries should move into more complex production (and we couple it with measures such as relatedness, which aim to show countries the products that they are more likely to be able to develop successfully given their current product space), this disregards existing, and often longstanding, context-dependent issues such as natural resources dependence, institutional constraints, and lack of or incomplete national and local systems of innovation, among others.

Implications for the GCCs & issues with a top-down approach

The GCCs have long recognised the need for diversification away from the oil sector. In his analysis of their efforts, Hvidt (2013) notes a high level of uniformity in the assessment of the challenges and recommendations for the GCC countries in their development paths – which he argues may be related to the role of international consultancy in advising countries – and questions whether indeed there are no alternative pathways for them, beyond the dominant discourse in development strategies.

Recently, some consultancy efforts have relied on ECI metrics to devise plans for this group of countries. As argued in the literature, the GCCs need a wider set of actors, relationships and institutions to evolve and manage to diversify into different activities. Economic complexity, as an indicator and concept, has been argued as capturing these underlying interactions and linkages and reflect them neglecting investigation into local capabilities and conditions (often more costly and longer-term). This has led to a certain attractiveness for policymakers to rely on such indicators, prioritising approaches unlikely to provide specificity for effective policy action. As with other top-down approaches to policy, the course of action and initiatives proposed are unlikely to survive or to generate much change and development, particularly if they are introduced as stand-alone policies.

While some of the issues identified with the ECI are unique to its construction and the data employed, the broader challenges apply in other cases. For example, the World Bank's Knowledge Economy Index is also used in other reports on the GCCs and the MENA region, and has been criticised in similar ways. Brinkley et al. (2012) summarises some of the issues – in particular, the fact that they are based on several indicators with different degrees of volatility and sensitivity to economic cycles, and as a result they can move in odd and unpredictable ways, making their use limited for analysis and policy guidance in developing or emerging contexts. Furthermore, if used in isolation such indicators do not provide enough insights into economic and innovation systems, with limited guidance for policymakers on the changes required to achieve a sustainable development trajectory (Brinkley et al., 2012). Clearly, this critique applies closely to the case of the ECI, which provides a top-down overview of 'capabilities' in GCC countries vis-a-vis the rest of the world, but which appears to move erratically over time, in ways that at least in the short-run do not necessarily reflect capabilities within countries. Moreover, even if or when they do provide an accurate representation, it is not clear how countries can strive to achieve higher complexity (and sometimes even how some countries managed to do it in hindsight).

In addition to these challenges, the fact that this indicator is based on world exports raises further concerns, as it might simply reflect political economy aspects rather than development or production capacity – as seen in the recent Russian invasion of Ukraine, key oil-exporting countries can use oil production as a political weapon, and their price can fluctuate as a result, in ways that are completely independent of local capabilities or even of availability of resources.

3.8 Conclusion

This chapter looked at economic complexity indicators based on exports and explored the case of the GCC countries, which are highly dependent on oil and natural gas products. While we found that the link between the ECI and subsequent economic growth observed across countries holds for the GCC and other oil-dependent countries, our analysis exposed important ways in which these measures are affected by the high dependence on oil and its price volatility. We found no association between economic complexity and economic growth within countries over time, which we argue presents challenges in terms of policy-making. Our analysis points to the need for more caution when relying on economic complexity measures for policy-making, and highlight the need for additional and more granular analysis of different contexts, particularly in heavily oil-dependent countries.

Here, we addressed the specific case of oil and related products, which make up an important share of world exports, experience high price volatility, and are susceptible to changes in economic and political conditions across the world. While we looked at the GCC countries in detail, the broad lessons apply to other oil-dependent countries. In terms of other natural resources, the implications may be different – natural resources are highly diversified and, while our call for caution might apply more broadly, other natural resources are not as impacted by highly volatile prices or political economy conditions (for instance, diamonds have seen steady increases in price over time, and do not experience oscillations to the same degree), thus their impact on ECI measures is likely to be more moderate.

Finally, our study has limitations. A key aspect missing is the a more explicit consideration of the natural resource curse in the GCCs context which, unlike other highly natural resource dependent countries, have managed to achieve high income. This can have important implications in assessing the link between economic complexity and economic growth. While we tried to address this by looking at alternative indicators beyond GDP per capita, such as the HDI, the implications and results were the exact same, likely because GDP plays a large role in the HDI. We also explored the possible use of alternative skills measures, but availability is very limited, particularly over time.

Further research could therefore look into the links between economic complexity and different natural resources, to try and assess the applicability of the ECI to other contexts. Moreover, it could further assess the links between economic complexity and the natural resource curse, exploring not just growth but broader aspects of economic development.

3.A Appendix

Table 3.A.1: Description and source of variables. All available for the period 1995-2019.

Variable	Definition	Source
ECI	Economic Complexity Index based on HS-92 classification. Own calculations.	The Observatory of Economic Complexity
PCI	Product Complexity Index based on HS-92 classification. Own calculations.	The Observatory of Economic Complexity
Exports	Total merchandise exports (USD value)	The Observatory of Economic Complexity
NR exports	Natural resource exports (total USD value). Own calculation based on HS section V (mineral products) covering Chapters 25-27.	The Observatory of Economic Complexity
Oil and gas exports share	Exports in oil and natural gas products divided by total merchandise exports. Own calculation based on HS products 2709, 2710, 2711, 2712 and 2713.	The Observatory of Economic Complexity
GDP per capita	GDP per capita, PPP (constant 2017 international $)	World Bank Open Data
GDP	GDP (current USD)	World Bank Open Data
Population	Total population (counts all residents regardless of legal status or citizenship)	World Bank Open Data
Natural resource rents	Total natural resources rents (% of GDP)	World Bank Open Data
Oil rents	Oil rents (% of GDP)	World Bank Open Data
Employment share in industry	Employment in industry (% total employment)	World Bank Open Data
Employment share in services	Employment in services (% total employment)	World Bank Open Data
Tertiary schooling	School enrolment, tertiary (% gross)	World Bank Open Data
Secondary schooling	School enrolment, secondary (% gross)	World Bank Open Data
HDI	Summary measure of achievement in key dimensions of human development (long and healthy life, being knowledgeable, decent standard of living)	United Nation's Human Development Reports
Crude oil prices	Global crude oil prices, measured in current US dollars per barrel	BP Statistical Review of World Energy

Table 3.A.2: Descriptive statistics, 20-year growth regression variables

Variables	N	Mean	SD	Min	Max
GDPpc growth	164	0.428	0.358	-0.406	1.609
Initial ECI	176	-1.14e-05	1.000	-3.737	2.527
Initial ECI (excl. oil)	176	-1.14e-05	1.000	-2.154	2.778
Initial GDPpc (log)	164	9.006	1.241	6.447	11.54
Increase in exports (share of initial GDP)	171	0.999	1.227	-0.211	8.402
Increase in NR exports (share of initial GDP)	171	0.301	0.693	-0.206	5.442
Exports to GDP (initial)	171	0.326	0.234	0.0123	1.291
Initial population (log)	176	15.80	1.741	11.87	20.96

Table 3.A.3: Descriptive statistics, 10-year growth regression variables

Variables	N	Mean	SD	Min	Max
GDPpc growth	332	0.198	0.222	-0.630	1.273
Initial ECI	354	1.69e-05	0.999	-3.737	2.558
Initial ECI (excl. oil)	354	2.82e-06	0.999	-2.154	2.778
Initial GDPpc (log)	332	9.134	1.223	6.447	11.68
Increase in exports (share of initial GDP)	345	0.297	0.551	-0.521	6.675
Increase in NR exports (share of initial GDP)	345	0.104	0.451	-0.389	6.286
Exports to GDP (initial)	345	0.313	0.213	0.0123	1.291
Initial population (log)	354	15.87	1.733	11.87	21.01

Table 3.A.4: Descriptive statistics, 5-year growth regression variables

Variables	N	Mean	SD	Min	Max
GDPpc growth	668	0.0882	0.120	-0.614	0.954
Initial ECI	709	8.46e-06	0.998	-3.737	2.558
Initial ECI (excl. oil)	709	1.13e-05	0.998	-2.765	2.778
Initial GDPpc (log)	668	9.183	1.219	6.447	11.69
Increase in exports (share of initial GDP)	692	0.114	0.185	-0.949	2.530
Increase in NR exports (share of initial GDP)	692	0.0361	0.128	-0.302	2.431
Exports to GDP (initial)	692	0.313	0.214	0.00955	1.513
Initial population (log)	709	15.91	1.729	11.87	21.05

Table 3.A.5: List of countries included in the analysis

Afghanistan	Djibouti	*Lesotho*	Russia (*)
Albania	Dominican Republic	Liberia	Rwanda
Algeria (*)	Ecuador (*)	Libya (*)	Sao Tome and Principe
Angola (*)	Egypt (*)	Lithuania	Saudi Arabia (*)
Argentina	El Salvador	*Luxembourg*	Senegal
Armenia	Equatorial Guinea (*)	Macau	*Serbia*
Australia	Estonia	Madagascar	Sierra Leone
Austria	*Eswatini*	Malawi	Singapore
Azerbaijan (*)	Ethiopia	Malaysia	Slovakia
Bahamas	Fiji	Maldives	Slovenia
Bahrain (*)	Finland	Mali	Solomon Islands
Bangladesh	France	Malta	Somalia
Barbados	French Polynesia	Mauritania	South Africa
Belarus	Gabon (*)	Mauritius	South Korea
Belgium	Gambia	Mexico	*South Sudan*
Belize	Georgia	Moldova	Spain
Benin	Germany	Mongolia	Sri Lanka
Bhutan	Ghana	*Montenegro*	Sudan (*)
Bolivia (*)	Greece	Morocco	Suriname
Bosnia and Herzegovina	Guatemala	Mozambique	Sweden
Botswana	Guinea	Myanmar	Switzerland
Brazil	Guinea-Bissau	*Namibia*	Syria (*)
Brunei (*)	Guyana	Nepal	Tajikistan
Bulgaria	Haiti	Netherlands	Tanzania
Burkina Faso	Honduras	New Caledonia	Thailand
Burundi	Hong Kong	New Zealand	Timor-Leste (*)
Cambodia	Hungary	Nicaragua	Togo
Cameroon (*)	Iceland	Niger	Trinidad and Tobago (*)
Canada	India	Nigeria (*)	Tunisia
Cape Verde	Indonesia	North Korea	Turkey
Central African Republic	Iran (*)	North Macedonia	Turkmenistan (*)
Chad (*)	Iraq (*)	Norway (*)	Uganda
Chile	Ireland	Oman (*)	Ukraine
China	Israel	Pakistan	United Arab Emirates (*)
Chinese Taipei	Italy	*Palestine*	United Kingdom
Colombia (*)	Jamaica	Panama	United States
Comoros	Japan	Papua New Guinea	Uruguay
Costa Rica	Jordan	Paraguay	Uzbekistan
Cote d'Ivoire	Kazakhstan (*)	Peru	Vanuatu
Croatia	Kenya	Philippines	Venezuela (*)
Cuba	Kuwait (*)	Poland	Vietnam
Cyprus	Kyrgyzstan	Portugal	Yemen (*)
Czechia	Laos	Qatar (*)	Zambia
Democratic Republic of the Congo	Latvia	Republic of the Congo (*)	Zimbabwe
Denmark	Lebanon	Romania	

Countries in *Italics* have some early years of export data, and therefore complexity variables, missing.
Countries marked with (*) are included in the oil-dependent group.

Table 3.A.6: Top and bottom products with $RCA > 1$, GCCs, 1995 and 2019, ordered by PCI

Bahrain:

1995					2019				
Product	**PCI**	**PCI Rank**	**Ubiquity**	**Export RCA**	**Product**	**PCI**	**PCI Rank**	**Ubiquity**	**Export RCA**
Aluminum foil <0.2 mm	0.872	255	21	1.1	Tungsten (wolfram)	1.44	72	17	1.16
Still image projectors	0.8	284	14	1.26	Stabilizers for rubber or plastic	1.218	120	17	12.65
Aluminum powders	0.426	428	18	140.83	Multiple-walled insulating glass	1.039	181	21	6.9
Other articles of aluminum	0.382	446	28	1.74	Time switches	0.961	219	14	2.12
Aluminum plates >0.2 mm	0.344	465	21	41.96	Parts of railway locomotives	0.856	256	23	1.79
...	...	...	...	...	...	...	...	...	...
Raw skins of sheep or lambs	-1.888	1204	55	1.24	Spices	-1.624	1141	51	1.44
Cinnamon	-1.924	1206	12	1.81	Cotton waste	-1.683	1150	32	4.26
Men's shirts, knit	-2.055	1214	46	3.34	Iron ores and concentrates	-1.726	1156	18	11.55
Crustaceans	-2.384	1230	60	1.26	Crustaceans	-2.096	1189	43	2.99
Palm oil	-2.865	1239	14	2.2	Gold	-2.836	1214	69	1.08

Kuwait:

1995					2019				
Product	**PCI**	**PCI Rank**	**Ubiquity**	**Export RCA**	**Product**	**PCI**	**PCI Rank**	**Ubiquity**	**Export RCA**
Glass fibers	0.788	289	24	1.09	Catalytic preparations	1.039	180	17	1.31
Radar	0.296	495	20	1.45	Clocks with watch movements	1.037	182	15	17.15
Hydrochloric acid	0.2	534	26	1.2	Other clocks	0.841	264	8	6.64
Paper trays and similar (base metal)	0.174	553	15	2.02	Cyclic hydrocarbons	0.796	291	17	5.35
Buses	-0.237	760	26	1.3	Scent sprays	0.778	298	10	3.97
...	...	...	...	...	...	...	...	...	...
Sulphur, crude	-1.091	1051	14	14.31	Wheat or meslin flour	-1.429	1109	60	3.02
Petroleum gases	-1.284	1099	22	7.54	Sulphur, crude	-1.574	1133	22	4.52
Petroleum oils, refined	-1.45	1132	62	15.81	Petroleum gases	-2.176	1195	34	2.05
Raw skins of sheep or lambs	-1.888	1204	55	1.79	Zirconium ore	-2.473	1203	32	2.52
Petroleum oils, crude	-2.428	1234	38	16.99	Petroleum oils, crude	-2.82	1213	39	12.11

Table 3.A.6 (continued): Top and bottom products with $RCA > 1$, GCCs, 1995 and 2019, ordered by PCI

Oman:

1995				
Product	**PCI**	**PCI Rank**	**Ubiquity**	**Export RCA**
Parts for use with hoists	1.374	116	19	1.12
Cars	1.177	167	13	1.01
Machines for testing mechanical properties	1.103	184	13	1.83
Photographic film in rolls	1.06	193	11	1.07
Plaster articles	0.808	281	17	1.06
...	...	...	...	...
Other live animals	-1.953	1208	65	31.03
Avocados, pineapples, mangos, etc.	-2.032	1211	52	1.5
Molluscs	-2.083	1217	51	1.61
Chromium ore	-2.188	1223	19	1.25
Petroleum oils, crude	-2.428	1234	38	18.52

2019				
Product	**PCI**	**PCI Rank**	**Ubiquity**	**Export RCA**
Aldehydes	1.039	179	16	1.45
Flat-rolled iron, width <600mm, not clad	0.969	216	24	1.67
Cyclic hydrocarbons	0.796	291	17	8.45
Polyacetals	0.726	315	22	2.11
Aluminum plates >0.2 mm	0.685	342	25	3.26
...	...	...	...	...
Legumes	-2.059	1186	42	1.6
Petroleum gases	-2.176	1195	34	8.66
Palm oil	-2.344	1199	27	1.19
Chromium ore	-2.473	1202	11	10.06
Petroleum oils, crude	-2.82	1213	39	8.59

Qatar:

1995				
Product	**PCI**	**PCI Rank**	**Ubiquity**	**Export RCA**
Polymers of ethylene	0.276	506	26	8.08
Buses	-0.237	760	26	1.46
Acyclic hydrocarbons	-0.264	765	19	11.21
Hot rolled bars of iron	-0.297	783	32	3.12
Gypsum	-0.451	842	27	5.35
...	...	...	...	...
Activewear	-1.397	1122	51	1.03
Petroleum oils, refined	-1.45	1132	62	3.18
Men’s shirts	-1.755	1194	65	3.91
Other raw hides and skins	-1.842	1201	66	1.31
Petroleum oils, crude	-2.428	1234	38	19.78

2019				
Product	**PCI**	**PCI Rank**	**Ubiquity**	**Export RCA**
Halogenated derivatives of hydrocarbons	1.099	156	14	6.29
Acyclic hydrocarbons	0.511	414	29	2.67
Silicon & rare gases	0.324	505	25	11.99
Ethers	0.221	544	17	4.54
Polymers of ethylene	0.213	548	23	5.84
...	...	...	...	...
Other bars of iron (no further than forged)	-1.171	1041	54	4.87
Raw skins of sheep or lambs	-1.442	1112	41	1.17
Sulphur, crude	-1.574	1133	22	22.3
Petroleum gases	-2.176	1195	34	34.09
Petroleum oils, crude	-2.82	1213	39	3.62

Table 3.A.6 (continued): Top and bottom products with $RCA > 1$, GCCs, 1995 and 2019, ordered by PCI

Saudi Arabia:

1995					2019				
Product	**PCI**	**PCI Rank**	**Ubiquity**	**Export RCA**	**Product**	**PCI**	**PCI Rank**	**Ubiquity**	**Export RCA**
Ion-exchangers based on polymers	1.671	44	6	3.75	Epoxides	1.831	16	10	5.44
Pigments, nonaqueous	0.916	231	13	1.75	Phenols, phenol-alcohols	1.621	39	14	3.75
Structures and their parts, of iron or steel	0.876	253	23	1.1	Amino-resins	1.41	77	18	2.48
Ceramic pipes	0.746	295	9	3.67	Acrylic polymers	1.38	83	16	1.03
Aluminum containers, <300 liters	0.708	318	26	1.4	Wire of stainless steel	1.319	90	16	1.04
...	...	...	...	...	...	...	...	...	...
Tanned sheepskins	-1.519	1148	31	4.14	Avocados, pineapples, mangos, etc.	-1.952	1174	54	1.21
Tanned skins of other animals	-1.905	1205	29	2.02	Tanned sheepskins	-2.071	1188	37	3.57
Avocados, pineapples, mangos, etc.	-2.032	1211	52	1.64	Petroleum gases	-2.176	1195	34	1.03
Nutmeg	-2.046	1213	28	1.57	Tanned skins of other animals	-2.182	1196	36	1.25
Petroleum oils, crude	-2.428	1234	38	18.01	Petroleum oils, crude	-2.82	1213	39	11.73

United Arab Emirates:

1995					2019				
Product	**PCI**	**PCI Rank**	**Ubiquity**	**Export RCA**	**Product**	**PCI**	**PCI Rank**	**Ubiquity**	**Export RCA**
Pigments, nonaqueous	0.916	231	13	1.88	Diamond dust	1.259	108	14	1.51
Water gas generators	0.74	298	13	2.51	Spark-ignition int. combustion engines	1.189	127	19	1.1
Aluminum containers, <300 liters	0.708	318	26	1.95	Photographic paper	1.185	130	8	3.19
Other articles of plastic	0.6	364	16	1.56	Photographic film, not developed	1.156	139	14	2.49
Aluminum tubes and pipes	0.577	368	23	8.01	Transmission apparatus for radio, phone, TV	0.867	249	13	2.24
...	...	...	...	...	...	...	...	...	...
Men’s shirts, knit	-2.055	1214	46	2.63	Legumes, dried	-2.033	1182	41	1.83
Preserved fish	-2.164	1222	53	2.36	Petroleum gases	-2.176	1195	34	1.38
Chromium ore	-2.188	1223	19	3.23	Cashew nuts & coconuts	-2.694	1210	31	1.32
Cloves	-2.259	1226	14	3.67	Petroleum oils, crude	-2.82	1213	39	4.15
Petroleum oils, crude	-2.428	1234	38	16.41	Gold	-2.836	1214	69	4.29

Figure 3.A.1: Natural resource rents (percentage of GDP), GCCs, 1995 to 2019

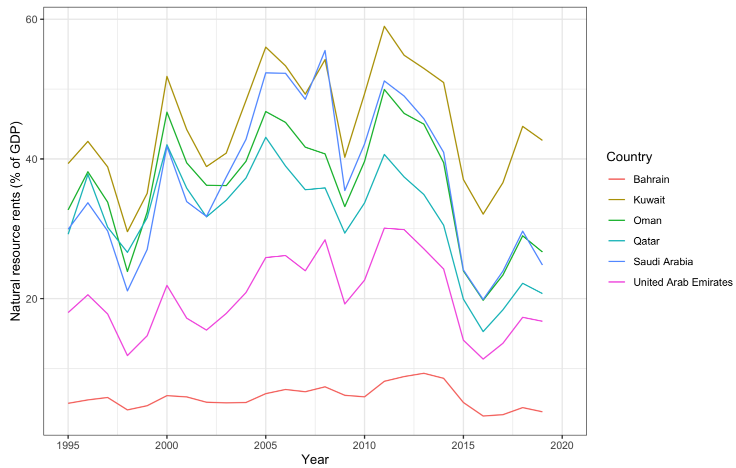


Figure 3.A.2: GDP (log, constant 2017 international $), GCCs , 1995 to 2019

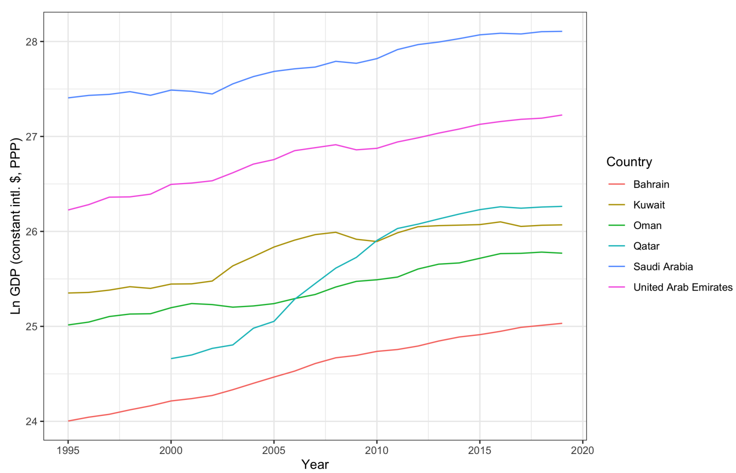


Figure 3.A.3: Population (log), GCCs, 1995 to 2019

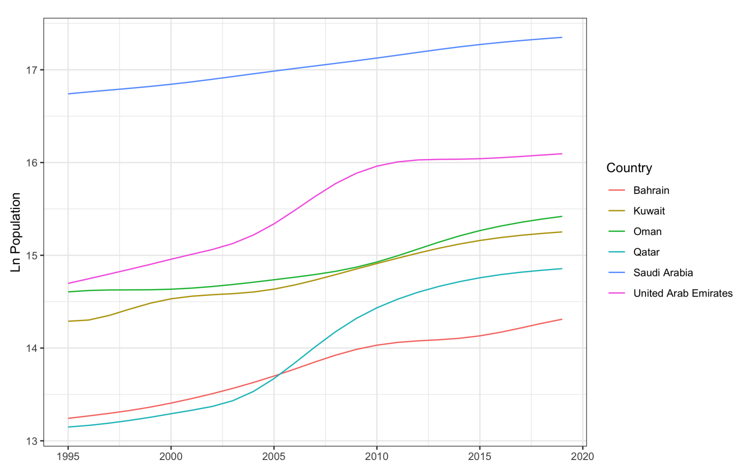


Figure 3.A.4: ECI and crude oil prices over time, GCCs, 1995 to 2019

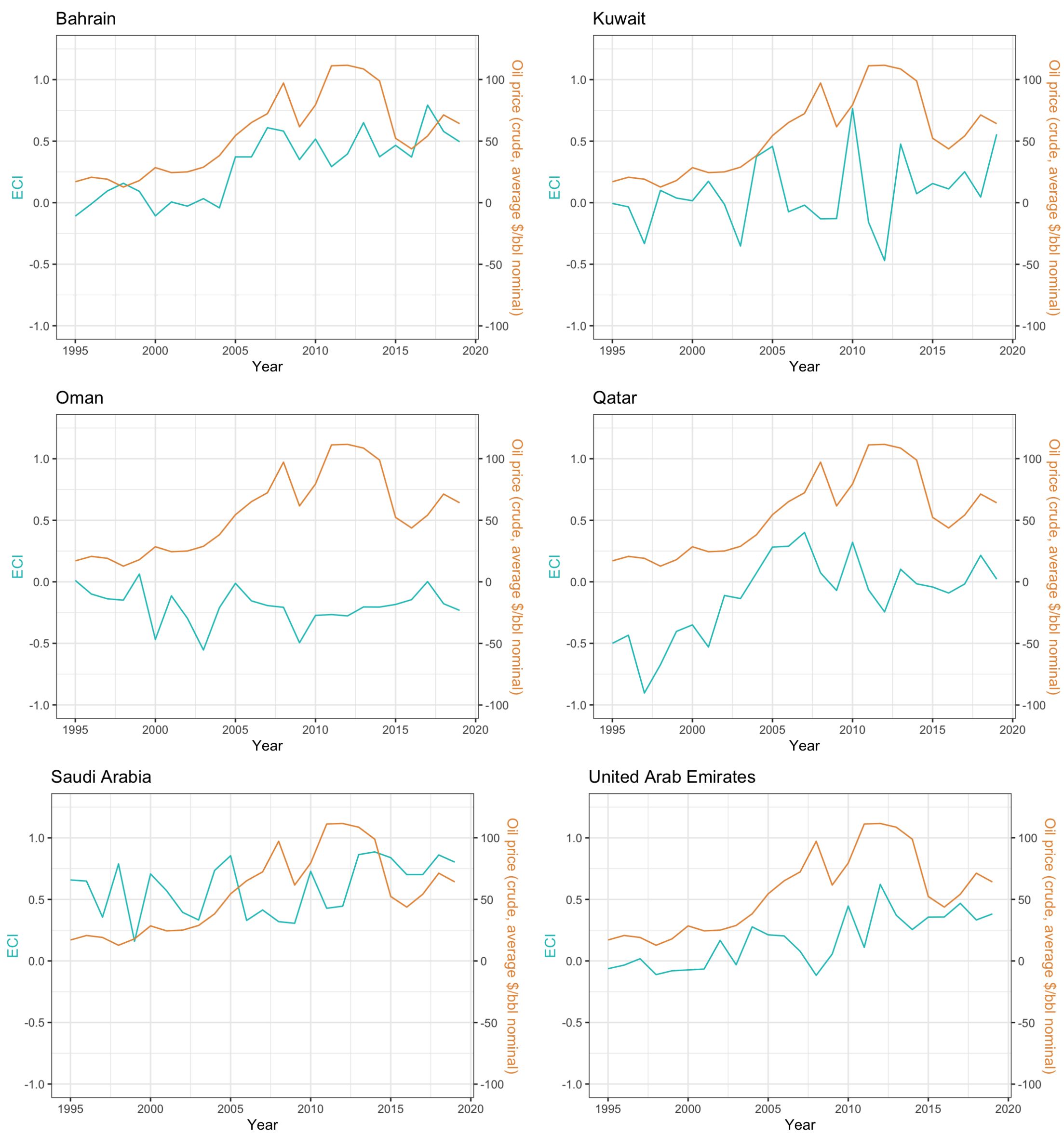
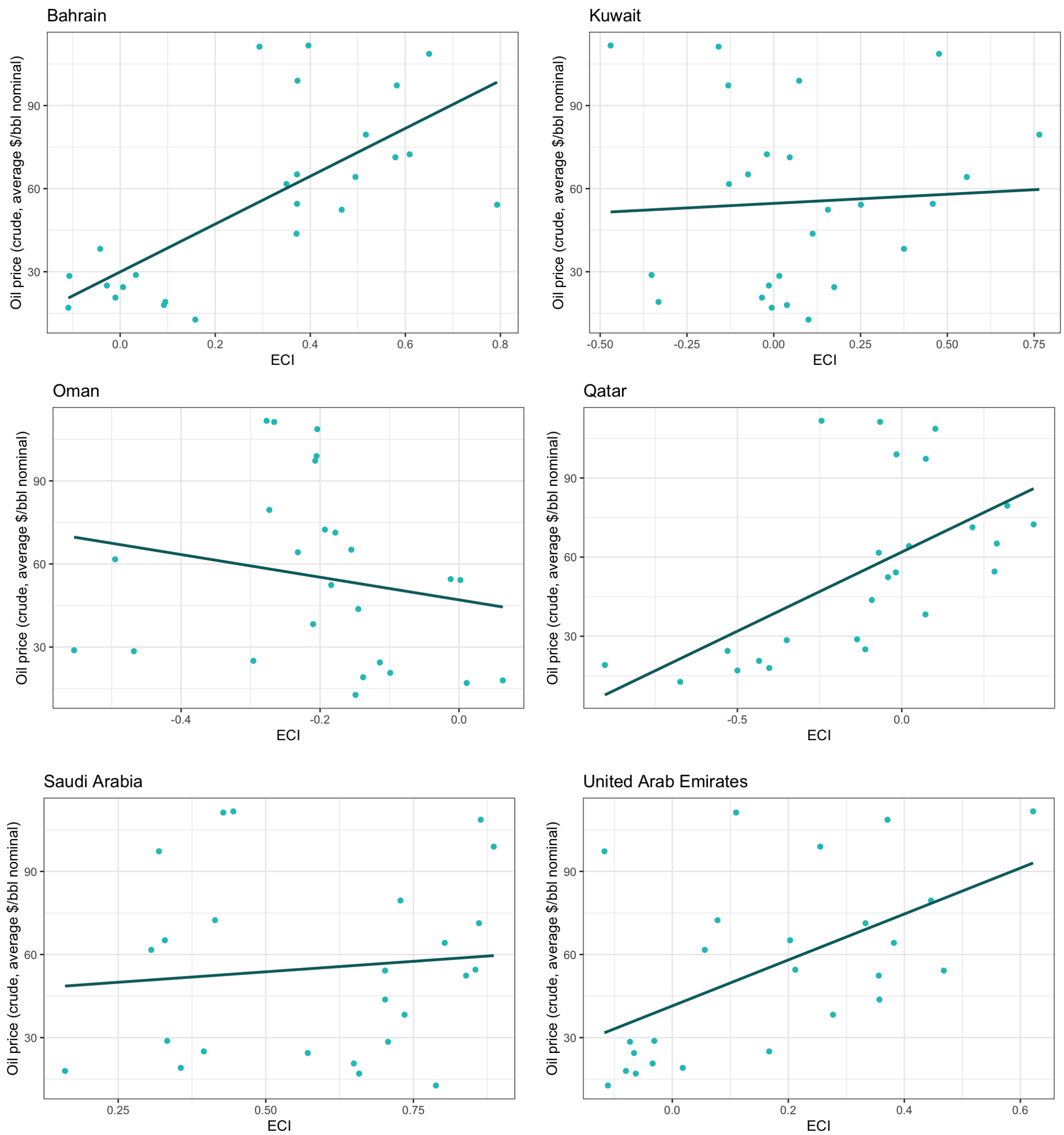


Figure 3.A.5: ECI and crude oil prices scatter plots, GCCs, 1995 to 2019



Chapter 4

Economic complexity, education and economic growth

4.1 Introduction

Throughout the previous three chapters, this thesis has investigated different key questions pertaining to the ECI as a concept and method, and its meaningfulness and applicability to different contexts and questions. A broad question that remains is whether this methodologically advanced concept adds to our existing knowledge about why some countries become more prosperous over time, while others stagnate. As discussed in Chapter 1, theoretical and empirical contributions have identified important determinants of economic development across countries. A key one is education, which is the focus of this chapter.

Education has long been recognised as a key driver of economic growth and development, in particular through key contributions in the 1960s and onward (e.g., Schultz, 1961; Becker, 1964; Nelson and Phelps, 1966). Through its contribution towards human capital and labour productivity, as well as to an economy's capacity to innovate and develop new technology, schooling can generate economic growth (e.g., Hanushek and Wossmann, 2010).

The proponents of the ECI discuss the role of education, and human capital more broadly, in economic development and argue that economic complexity can still add something to our understanding of why some countries grow more than others over time. Hausmann et al. (2014b, p. 36) argue that rather than measuring how much of the same knowledge individuals have, economic complexity can infer the level of productive knowledge, which originates not just in the classroom, but also through on-the-job training and interactions between firms and individuals, and that ultimately this type of knowledge is what drives economic prosperity. They also show that initial economic complexity levels can explain more of the variance in economic growth than traditional education measures (Hausmann et al., 2014b). Nevertheless, the empirical literature on economic complexity since then is mixed – some papers do not add education as a control variable (e.g., Stojkoski et al., 2016), other papers control for education variables and the effect of the ECI remains statistically significant (e.g., Stojkoski et al., 2023), while in other cases the ECI loses statistical significance when education is included (e.g., Fritz and Manduca, 2021).

This chapter provides a preliminary empirical analysis of the links between economic complexity, education and economic growth across countries. In particular, it explores whether economic complexity complements our existing understanding of why some countries experience economic growth, while others do not. Using the export-based economic complexity data from Chapters 1 and 3, we explore the link between economic growth, economic complexity, and well-established education measures, namely average years of education and secondary and tertiary schooling attainment from Barro and Lee (2013), for 140 countries between 1995 and 2019. We start from the baseline model, regressing economic growth on the ECI, and add key schooling variables in a horse-race exercise, which allows us to understand whether the association between the initial ECI and subsequent economic growth is robust to the inclusion of education measures, and whether the ECI does indeed explain more of the variance in economic growth than education.

This initial analysis shows that education variables, particularly average years of education, are a stronger predictor of both long- and short-term economic growth than the ECI. Moreover, once we include additional control variables, average years of education shows a more robust and consistent link with economic growth than the ECI.

The remainder of the chapter is structured as follows – Section 4.2 motivates our research question by looking at the literature on education and economic growth, and linking it with work on economic complexity. Next, Section 4.3 describes the preliminary empirical analysis of cross-country evidence carried out in this chapter, while Section 4.4 presents the results. Lastly, we provide a brief conclusion in Section 4.5.

4.2 Background and motivation

4.2.1 Education and economic development

Since the late 1980s, education and human capital more broadly have been recognised as playing an important role in economic progress. As Hanushek and Wossmann (2010) describe, there are three theoretical mechanisms behind the impact of education on economic growth. First, following the rationale in neoclassical growth theories (e.g., Solow, 1957), education increases human capital and labour productivity, leading to transitional growth towards a higher equilibrium level of output. Second, as proposed by endogenous growth models (e.g., Lucas, 1988; Romer, 1990), education increases knowledge and innovative capacity of economies, leading to new technologies, processes and products that generate economic growth. Third, education can facilitate the transmission of knowledge and new information that enables the adoption of technologies (e.g., Nelson and Phelps, 1966; Benhabib and Spiegel, 1994), while university education and research also contributes to technological advancement in industry (Rosenberg & Nelson, 1994).

According to Becker (1964), human capital captures the set of knowledge, skills, competencies and abilities that are embodied in individuals, and are acquired through a variety of ways, including not just education, but also other training, healthcare or experiences such as migration. Nevertheless, economists have mostly focused on education and its link with economic growth, with early empirical contributions drawing mostly on measures of education that capture *quantity* – at first looking simply at literacy and enrolment rates, and then moving to a more robust proxy, namely years of education of the working age population, assuming that such educational investments reflect the human capital embedded in workers (Benos & Zotou, 2014).

While the theoretical underpinning of the link between education and economic development is clear and compelling, the empirical evidence has been more mixed, in large part reflecting measurement issues (Hanushek & Wossmann, 2010). The empirical estimations consist of cross-country growth regressions where average annual growth in GDP per capita over several decades is regressed on measures of schooling and other determinants of economic growth as control variables, which is similar to the specifications used in the economic complexity literature.

Early empirical findings by Barro (1991) and Mankiw et al. (1992) pointed to a positive correlation between schooling enrolment and GDP per capita across countries. These papers were, however, criticised for having a small number of regressors and likely suffering from omitted variables bias. Levine and Renelt (1992) addressed this issue by estimating a large number of regressions to check how robust common explanatory variables were in explaining economic growth, and found that primary and secondary schooling enrolment did not consistently have positive, statistically significant impacts on economic growth. Barro and Lee (1994) were the first ones to use average years of schooling as a measure of human capital and this variable remains one of the most widely-used human capital proxies in recent contributions.

Bils and Klenow (2000) were, to the best of our knowledge, the first contribution to assess the causal link between education and economic growth. Focusing on enrolment rates, they conclude that the previous findings by Barro and others should not be interpreted as reflecting the impact of education on growth, as the link was too weak, even when considering the effect of education on the adoption of technology. Bils and Klenow (2000) emphasise in particular the potential of omitted variables that are likely to influence both initial enrolment rates and subsequent economic growth in preceding contributions.

Glewwe et al. (2014) also point to the issue of omitted variable bias, as well as endogeneity – for instance, cultural or historical aspects across countries could have effects on both economic growth and initial education levels, and thus the education variables would be correlated with the error term. Moreover, Benos and Zotou (2014) run a meta-regression analysis of the link between education and economic growth and find a strong publication selection bias towards results that find a positive impact of education on economic growth. Importantly, the authors do not find a systematic genuine impact of education on economic growth, with varying results for different education variables (Benos & Zotou, 2014).

The literature evolved towards making a distinction between educational quantity, as captured in enrolment rates or average years of education, and quality. In research using average years of education of the workforce, for instance, there is an underlying assumption that a year of education amounts to a similar outcome in terms of human capital across countries and time. Nevertheless, it is unrealistic to think that this is the case, and thus researchers have increasingly turned to measures of educational quality, such as expenditure on education, student-teacher ratios and cognitive skills test scores (Benos & Zotou, 2014). In particular, Hanushek and Kimko (2000) and Hanushek and Woessmann (2008) provided evidence that differences in the quality of schooling may be more important in explaining economic growth across countries than measures of schooling quantity. Moreover, when considering the case of Sub-Saharan Africa, Glewwe et al. (2014) make an important point for policy-making – rapid increases in school enrolment in Sub-Saharan Africa are likely to have reduced schooling quality as they led to more crowded schools and pressure on existing resources, and thus the longer-term impact on economic development may be more mixed if cognitive skills are not being improved or taken into account in empirical analysis.

More recently, Barro (2013) looks at the determinants of economic growth, including average years of school attainment and test scores, across 100 countries for 1960-1995, and finds positive and statistically significant effects of both quantity and quality on economic growth (though quality is more important quantitatively). Looking at the relationship across different groups of countries separately, Barro (2013) emphasises that, while data quality might be more heterogeneous, it is important to use a large sample that includes countries at several economic development levels, as the observed variations in policy and other variables might be too limited among a specific group of countries to make accurate inferences (e.g., looking at OECD countries only).

An important question, as posed by Hanushek (2016), is whether the linear models hold across different countries with differing skill distributions and initial levels of educational attainment. In particular, when considering the link between university education and economic growth, Hanushek (2016) argues that different technological development levels and skills demand across countries with distinct income levels can play an important role in mediating this link. Referring to other models, such as the one by Aghion et al. (2009), he argues that tertiary education is particularly important for countries near the technological frontier, as further growth for them requires new invention and innovations and, moreover, given the technological development of such countries, there will be considerable demand for the type of more advanced skills that university education tends to produce. In contrast, in lower income countries, which are likely to be further away from the technological frontier, there may be lower demand for the high-level skills produced by higher education and growth might be more contingent on the capacity to absorb or adopt existing technologies rather than to develop new technology (Hanushek, 2016). Moreover, earlier research such as Vandenbussche et al. (2006), also suggested that education facilitates the diffusion of technology, with different levels of education playing different roles – earlier years of education are more important for imitation, while higher education facilitates innovation.

Despite the mixed empirical evidence, there is a strong case, theoretically, as to why education can play a role in determining economic growth across countries. Important caveats remain empirically, not least the lack of compelling causal evidence, the plausibility of the simple economic growth models used, measurement issues, and the fact that a lot of the literature overlooks differences in schooling quality across countries (Hanushek & Wossmann, 2010; Barro, 2013). Furthermore, research has suggested that broader factors, such as institutional frameworks, are crucial in determining economic development (Acemoglu et al., 2001; Acemoglu et al., 2002), and also in mediating the effect that schooling can have in generating economic growth (North, 1990; Hanushek & Wossmann, 2010).

4.2.2 Linking economic complexity, education and economic growth

As described in detail in previous chapters, the proponents of the ECI have emphasised its predictive power for future economic growth as one of the reasons why economic complexity matters (e.g., Hausmann et al., 2014b; Hidalgo, 2023). Still, the vast literature presents mixed findings, and the link between economic complexity and economic growth might depend on the sample of countries used, as Ourens (2013) pointed out. Unlike the case of education and economic growth, which have a strong and widely-accepted theoretical foundation, the links between initial economic complexity levels and subsequent economic growth have been less developed, and economic complexity is an empirical method and concept lacking clear theoretical links with economic growth, as argued in the first chapter.

As described in Chapter 1, Hidalgo and Hausmann (2009) introduce the idea that the ECI can proxy for capabilities within countries, a term used to capture several different aspects that enable a country to export certain products competitively, including education levels and other important contributors to economic growth, such as institutional quality. Their rationale is therefore that economic complexity encompasses more than what these simpler indicators can tell us, even if taken together. It is also, they argue, why economic complexity is associated with future economic growth.

In the Atlas of Economic Complexity, Hausmann et al. (2014b) discuss more explicitly how education-based measures of human capital compare with the ECI – they argue that, rather than simply looking at the quantity of education that individuals hold, economic complexity captures the variety of productive knowledge and the interactions between individuals and firms that enable that knowledge to be used. They suggest that schooling variables capture how much of the same knowledge individuals have, while the ECI tries to capture the amount of productive knowledge that is embedded in society, which is generated by the diversity of knowledge its individuals have and is more closely related with on-the-job training than with schooling (Hausmann et al., 2014b). Importantly, Hausmann et al. (2014b) investigate how well education and economic complexity variables explain future economics growth, relying on years of schooling and secondary and tertiary school enrolment. The authors show that the ECI captures 15% of the variance in economic growth that is not explained by education, while the education variables combined only account for 3% of the variance in economic growth (Hausmann et al., 2014b, p. 38).

Still, there are important links between education and economic complexity. We expect countries with higher schooling levels, whether it be in terms of quantity or quality, to have higher economic complexity. This link occurs in both directions – on the one hand, higher education contributes to higher productive knowledge and economic complexity; on the other hand, in a country with higher economic complexity, we would expect the returns to education to be higher and thus individuals would pursue more education. In fact, Balland et al. (2020), looking at the case of the United States at the sub-national level, use average years of education of employees working within an occupation or industry as a proxy for the level of knowledge complexity within those specific activities. This is done in order to avoid using measures of complexity that are derived from spatial information, supporting this important conceptual link between complexity and education.

Addressing this link, Zhu and Li (2017) investigate the interaction between economic complexity and human capital and its impact on economic growth, focusing on 126 countries for 1995-2010. While they find a positive interaction effect of economic complexity and education on economic growth, the evidence is not entirely clear, with conflicting results for long- and short-term growth, as well as the Pooled OLS and Fixed Effects estimations. The authors do not compare the relative power of economic complexity and education in explaining future economic growth, and they only include education through its interaction with the ECI, and therefore there is margin to build on this contribution.

While the original proponents showed that the ECI could explain more of the variance in economic growth than educational variables, the empirical evidence since then has been mixed. On the one hand, some papers do not consider education as control variables at all (e.g., Stojkoski et al., 2016). On the other hand, within the papers that do consider education, some still find a statistically significant effect of the ECI on economic growth (e.g., Stojkoski et al., 2023) while in other cases the effect of economic complexity on economic growth lost all statistical significance (e.g., Fritz and Manduca, 2021, looking at United States metropolitan areas). We therefore want to explore further whether economic complexity is indeed better at explaining future economic growth than other simpler variables that researchers have long used when trying to understand economic growth across countries. We are not arguing that these links are causal – there are potential issues of reverse causality and omitted variable bias (which, as will be further discussed in the thesis conclusion, are hard or even impossible to address). Instead, we simply want to assess whether an intermediate variable like economic complexity can help us understand something about economic growth across countries beyond what education variables capture.

4.3 Methods and data

The main purpose of this chapter is to understand whether, as suggested in existing literature, the ECI does explain future economic growth more accurately than education measures. Our starting point is to run the most common specification in the economic complexity literature, following the initial contributions from Hidalgo and Hausmann (2009) and Hausmann et al. (2014b) and what we did in Chapter 3, of regressing economic growth on the initial ECI, controlling for the initial income level and the role of natural resource exports. Our main specification is as follows:

$$growth_{i,t+n} = \alpha + \beta_1 ECI_{i,t} + \beta_2 Educ_{i,t} + \beta_j \boldsymbol{X}_{j_{i,t}} + \eta_t + \epsilon_{it}$$

where $growth_{i,t+n}$ is the GDP per capita growth between t and $t+n$ for country i, calculated as $growth_{i,t+n} = log(GDPpc_{i,t+n}/GDPpc_{i,t})$.

$ECI_{i,t}$ is the initial ECI, while $Educ_{i,t}$ represents the education variables in the initial period, our main independent variables of interest. We focus on three key education variables – the average years of schooling, and secondary and tertiary schooling attainment. We add them to the model in a horse-race exercise, which allows us to understand whether the association between the initial ECI and subsequent economic growth is robust to the inclusion of initial education measures, and whether the ECI does indeed explain more of the variance in economic growth than education.

$\boldsymbol{X}_{j_{it}}$ is a vector representing the control variables – firstly, the initial GDP per capita (natural logarithm) to control for convergence across countries, and secondly, the increase in natural resource exports over the period (as a share of initial GDP) to capture the importance of natural resources. In our further analysis section, we consider three additional control variables – i) exports to GDP to account for trade openness, ii) the share of total employment in the industry sector to account for different levels of industrialisation and importance of manufacturing, and iii) the rule of law to capture institutional quality differences. Lastly, η_t and ϵ_{it} represent time fixed effects and the error term, respectively. The transformations of the variables used are in line with the aforementioned papers, for comparability. Across all models, and following the diagnostics discussed in Chapter 3, we use and report robust standard errors, clustered at the country level, to avoid violations of the OLS assumptions.

As in Chapters 1 and 3, the ECI is based on own calculations using data from the Observatory of Economic Complexity, following the detailed description in Chapter 1, and draws on a network of 179 countries and 1241 products. In terms of control variables, export data was downloaded from the OEC, while GDP and GDP per capita, the employment share in industry, and rule of law indicators are from the World Bank. Turning to education, we rely on data from Barro and Lee (2013), with a focus on the average years of schooling attained in the population over 25, and secondary and tertiary schooling attainment for this population group.[1] Given the availability of education data, we restrict our sample to 140 countries, for the period 1995 to 2019. In our further analysis section we also investigate heterogeneity across countries of different income levels, which we group based on the World Bank income classifications published in 2023.[2]

We focus on the cross-country associations between initial economic complexity and education levels, and subsequent income growth for 20-, 10- and 5-year periods between 1995 and 2019 (focusing on 2000-2019 for the long-term period). To check for within-country associations, we also estimate our model using Fixed Effects, reported in the appendix. This is in line with existing literature, including the original contribution by Hidalgo and Hausmann (2009) and more recent ones at the regional level e.g., by Mewes and Broekel (2022). As in Chapter 3, across all models, we use and report robust standard errors, clustered at the country level, to avoid violations of the Ordinary Least Squares (OLS) assumptions.

Table 4.A.1 provides the definition, source and availability of our variables, while Tables 4.A.2 to 4.A.4 show descriptive statistics for the variables and transformations used in the 20-, 10- and 5-year regressions, respectively. Table 4.A.5 provides the list of countries included in our analysis, grouped by income classification.

In our initial analysis, we also considered World Bank variables capturing public expenditure on education and looked at their link with economic complexity. Nevertheless, there are issues with using these variables from an economic development perspective as there is wide variability in the effectiveness of public spending on education and private expenditure in education is not considered, and thus these variables do not efficiently capture human capital investments. As a result, the link between government expenditure in education and the ECI changed considerably across the years and was negligible in many cases. Moreover, schooling enrolment data available from the World Bank presents too many random breaks over time for different countries, making the panel very unbalanced and limiting comparisons over time across countries. Lastly, data availability for cognitive skills and other test scores is limited, with snapshots for a single year and a limited number of countries, and thus we focus only on quantity-based measures of education in this preliminary analysis.

[1] The dataset by Barro and Lee includes the same variables for the population over 15 years old. Upon comparing the variables for the 15+ and 25+ population, we noted a very strong correlation in all cases, of 0.99 or higher. For the purpose of our analysis, we use the variables for the 25+ population, as this represents an age where educational attainment will be concluded for the vast majority of the population.

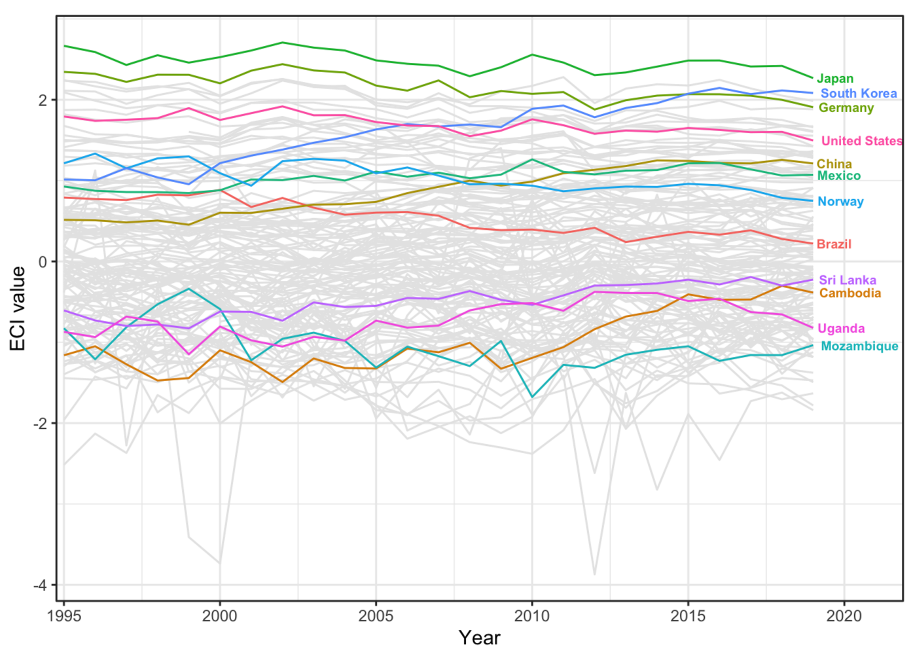
[2] World Bank Country and Lending Groups available here.

4.4 Empirical analysis

4.4.1 Economic complexity, education and growth across countries

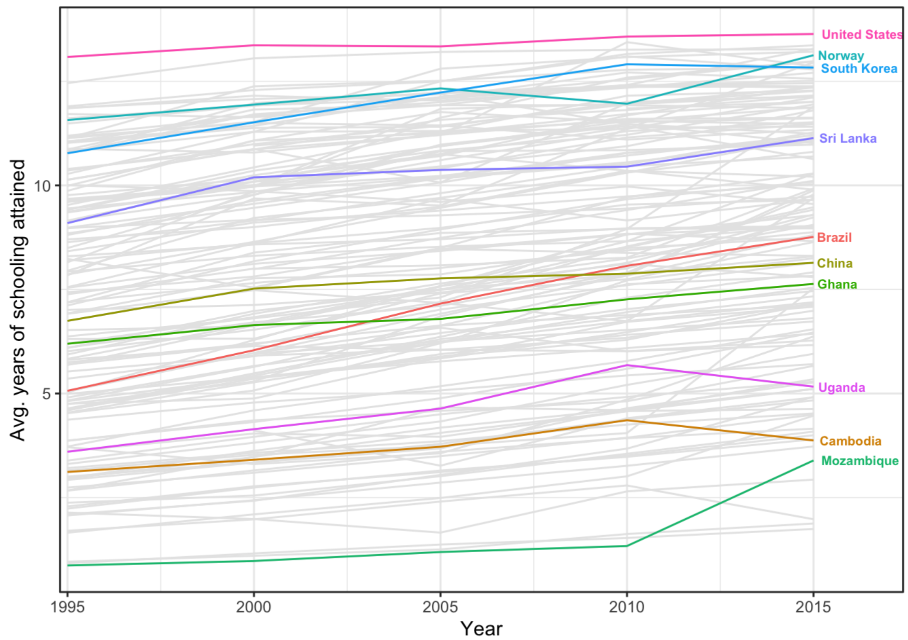
To provide context, Figure 4.1 shows changes in ECI between 1995 and 2019, highlighting countries of different complexity levels. Japan has the highest complexity levels, remaining number one in the ranking across the entire period, followed by Germany, which also remains relatively stable in terms of value and ranking. Among the countries with the highest complexity levels, South Korea experiences a steady increase over the period, both in terms of value and ranking. Still, as seen in previous chapters, the countries with lower levels of economic complexity also tend to be those that experience the sharpest year-on-year oscillations – we can see this illustrated in the case of Mozambique, for example, which experienced changes in value and relative position vis-a-vis other countries.

Figure 4.1: Economic complexity, 1995 to 2019



Turning to average years of schooling, Figure 4.2 shows its evolution over the same period (with data available every five years). As we might expect, this is more stable over time than the ECI and there is an upward tendency across all countries. For example, the United States shows the highest average years of schooling, whereas Mozambique is in the bottom most years, showing the lowest levels of complexity. In contrast with the general trend, we can see Uganda and Cambodia have experienced a decline between 2010 and 2015, whereas Brazil has seen a steady and among the highest increases over the period.

Figure 4.2: Average years of schooling, 1995 to 2015



Next, we look at correlations between the ECI and key education variables. Figure 4.3 plots the correlation between average years of schooling and the ECI, in 1995 and 2015 (the initial and final period with education data available, respectively). There is a strong positive correlation of over 0.7 in both years. In terms of patterns across income groups, we see high-income countries scoring the highest levels in terms of both ECI and average years of schooling, followed by upper-middle income countries. At lower levels of both complexity and education, we see lower-middle and low-income countries more mixed, particularly in terms of ECI level. Some outliers stand out – for example, Australia seems to have a lower level of economic complexity, given its education levels, particularly in 2015. Moreover, at lower levels of economic complexity, there appears to be a lot of variability in average years of schooling across countries.

Figure 4.4 shows similar scatter plots for 1995 and 2015, separating secondary and tertiary school attainment across countries. There is a positive correlation between the ECI and educational attainment in all plots. The top panels show secondary schooling attainment each year, with a significantly lower correlation coefficient in 2015 than in 1995. The bottom panels show tertiary schooling attainment, with similar correlation coefficients in both years, and aligned with those seen for the case of secondary schooling in 1995. In contrast with average years of education in the previous figure, when it comes to tertiary schooling attainment, there is much more variability across countries with high ECI –

e.g., the US and Germany have similar ECI levels, but very different tertiary schooling attainment percentage – than with low ECI, where we see much more similar (and generally low) levels of tertiary schooling attainment across countries.

Lastly, we consider income per capita and turn our attention to correlations between the ECI, GDP per capita and education in Figure 4.5, for 1995 (top) and 2015 (bottom). In each year, the panels under the diagonal show the scatter plots, coloured by income group, while the panels above the diagonal show the overall correlation coefficient and for each income group. When we take the full sample of countries, we see positive and fairly strong correlations between all variables. With regards to separate groups of countries, however, we see no correlation for low-income countries and, in fact, the correlations are mostly driven by the countries in the middle-income group. Besides the correlation between the two educational variables (present for all three groups as expected), the only cases where we see positive correlations for high-income countries is between the ECI and the education variables.

Figure 4.3: Average years of schooling attained and ECI, 1995 and 2015

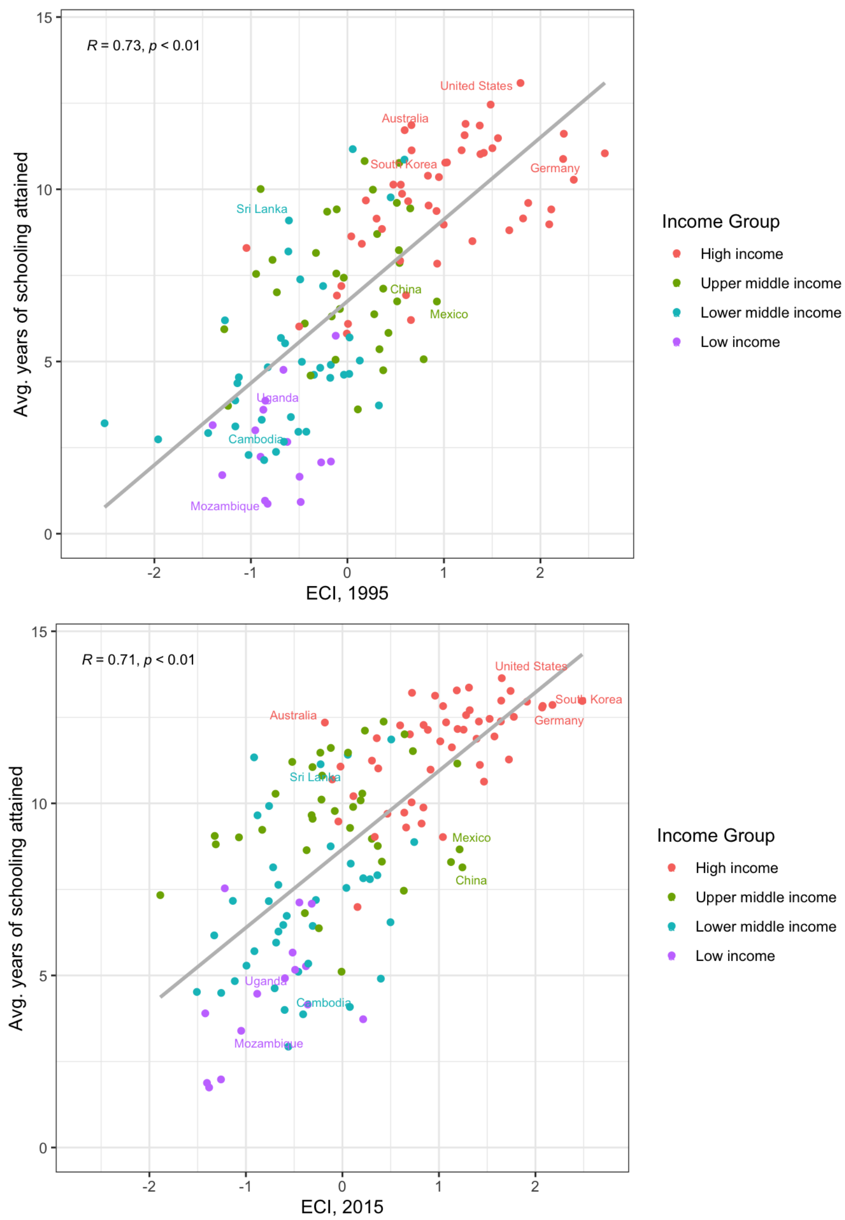


Figure 4.4: Secondary and tertiary schooling attainment and ECI, 1995 and 2015

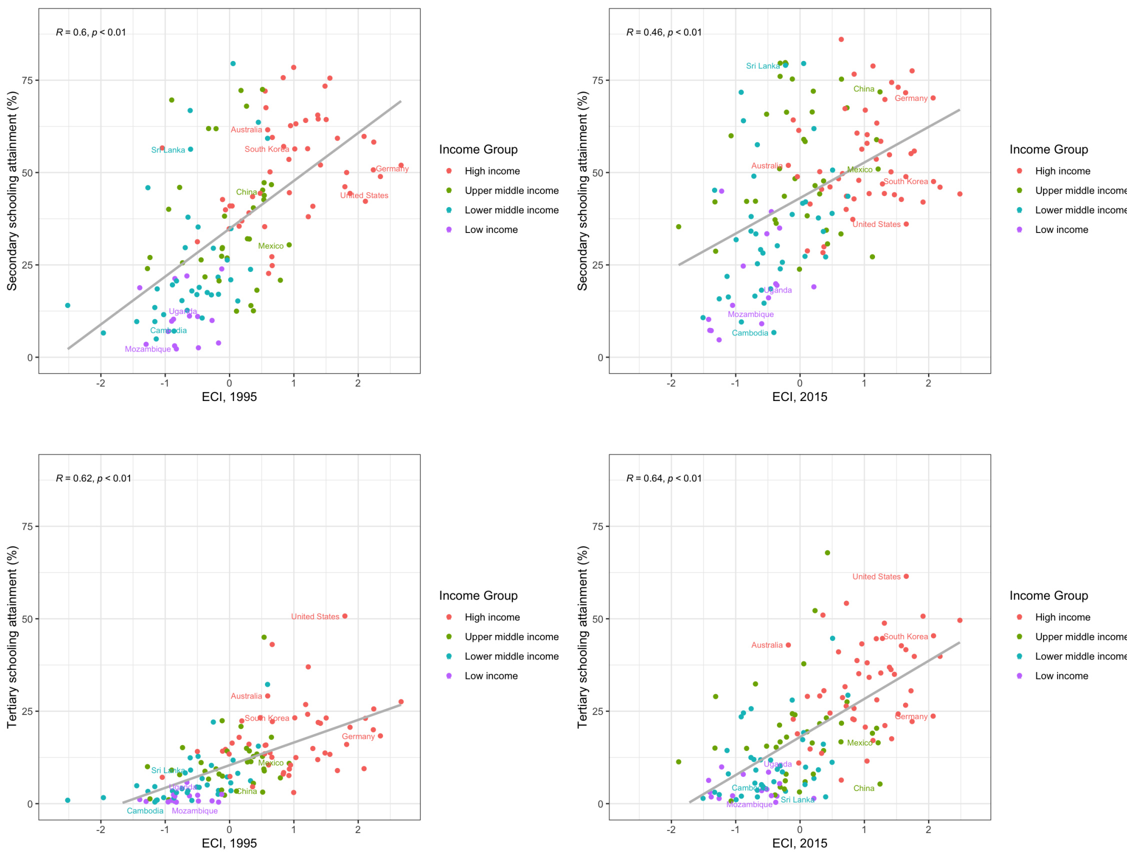
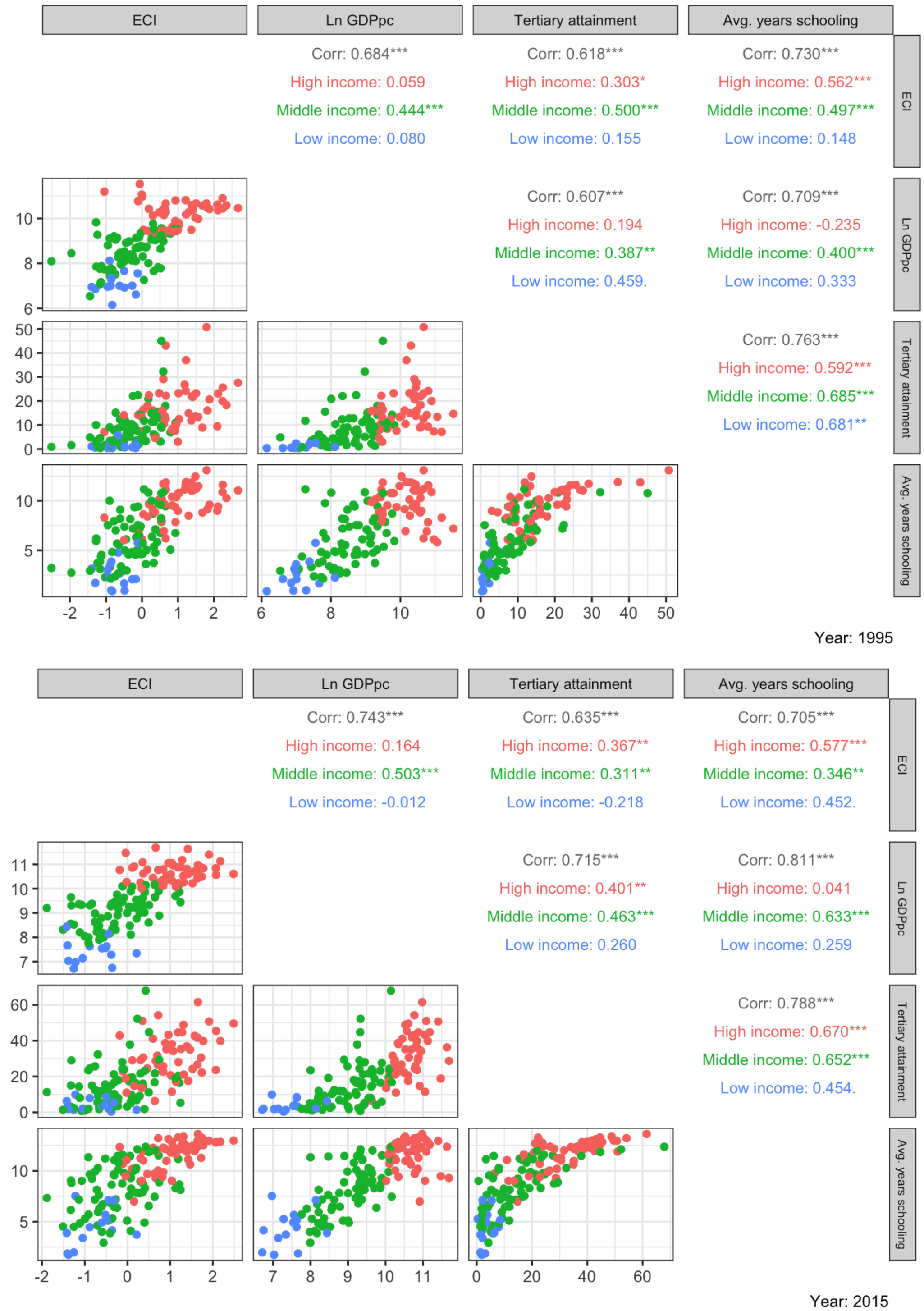


Figure 4.5: Correlations and scatter plots between selected variables, 1995 and 2015



4.4.2 Main results

Figure 4.6 shows partial regression plots, between initial ECI (first panel) and initial years of schooling (second panel), and long-term growth in GDP per capita for 2000-2019, conditional on the initial level of income per capita. There is a positive correlation in both cases, with a coefficient of 0.22 for economic complexity and a slightly stronger one of 0.3 for years of schooling.

Table 4.1 complements this, showing the cross-section regression results for the 20-year period. Here, we see the baseline model in column (1), which regresses growth in GDP per capita on the ECI, initial income per capita, and the increase in natural resources (as a share of initial GDP). The remainder of the columns add the education variables, then exclude the ECI to understand how their explanatory power compares to the baseline model with economic complexity. Adding average years of education in column (2) reduces the coefficient size and statistical significance of the ECI (though it is still positive and statistically significant at the 5% level), and it increases the explanatory power of the model by 10% (the adjusted R-squared increases from 0.202 to 0.302). Column (3) shows that including only average years of education and the two control variables leads to a higher explanatory power of income per capita across countries, than when we include only the ECI and the additional variables.[3] Turning to the other education variables, column (4) adds secondary and tertiary school attainment to the baseline model, and we see a further decrease in the ECI coefficient; the explanatory power of the model increases vis-a-vis the baseline one (though it is lower than for the case of years of schooling), and it is higher when we include only schooling attainment than when we include only the ECI (i.e., column 5 versus column 1). All education measures show a positive and statistically significant coefficient.

Turning to short-term growth, Table 4.2 shows cross-country regression results for 5-year growth in GDP per capita, mirroring the previous table. As before, adding average years of education to the estimation leads to a reduction in the coefficient for the ECI, while increasing the overall explanatory power of the model, which is higher when the ECI is removed (column 3) than in the baseline model (column 1), pointing to years of education being a better predictor of short-term economic growth than economic complexity. With regards to secondary and tertiary school attainment, they are both statistically significant, yielding a very similar adjusted R-squared to the models with years of schooling. This

[3] Moreover, when we remove the increase in natural resource exports (as a share of initial GDP) from the model, and regress economic growth on the initial GDP per capita, ECI and average years of education, the coefficient on the ECI is no longer statistically significant (available upon request).

therefore confirms the pattern seen in long-term growth. The same is true for our analysis for 10-year growth, provided in Appendix Table 4.A.6. Regarding within-country changes over time, as in Chapter 3, the ECI is not statistically significant in regressions explaining economic growth over 10- or 5-year periods across any specifications, as Appendix Tables 4.A.7 and 4.A.8 show.

Figure 4.6: Added variable plots, initial ECI, education and economic growth, 2000-2019

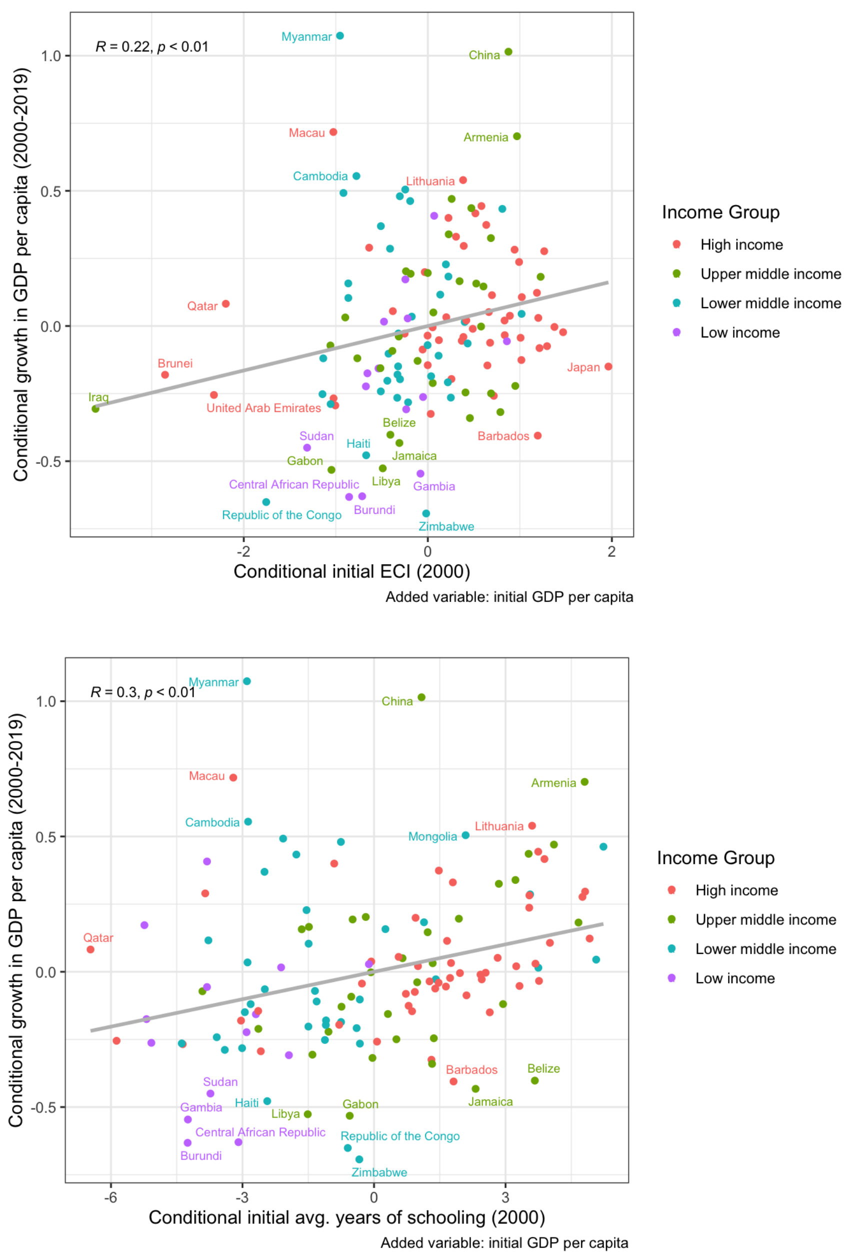


Table 4.1: Economic complexity and 20-year growth in GDP per capita, OLS, full sample

	GDPpc growth (2000-2019)				
Variables	(1)	(2)	(3)	(4)	(5)
Initial ECI	0.136***	0.0658**		0.0628*	
	(0.0340)	(0.0324)		(0.0339)	
Initial GDPpc (log)	-0.165***	-0.230***	-0.213***	-0.215***	-0.200***
	(0.0328)	(0.0348)	(0.0324)	(0.0324)	(0.0303)
Increase in NR exports	0.109***	0.0898***	0.0666**	0.0644**	0.0390
	(0.0327)	(0.0262)	(0.0262)	(0.0268)	(0.0277)
Avg. years of schooling (initial)		0.0546***	0.0638***		
		(0.0106)	(0.0105)		
Secondary school attainment (initial)				0.00577***	0.00651***
				(0.00130)	(0.00129)
Tertiary school attainment (initial)				0.00556**	0.00740***
				(0.00237)	(0.00225)
Constant	1.887***	2.075***	1.868***	2.060***	1.889***
	(0.304)	(0.289)	(0.253)	(0.287)	(0.258)
Observations	136	136	136	136	136
R-squared	0.220	0.323	0.306	0.318	0.305
Adjusted R-square	0.202	0.302	0.291	0.292	0.284

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4.2: Economic complexity and 5-year growth in GDP per capita, OLS, full sample

	GDPpc growth (1995-99, 2000-04, 2005-09, 2010-14, 2015-19)				
Variables	(1)	(2)	(3)	(4)	(5)
Initial ECI	0.0348***	0.0221**		0.0226**	
	(0.00960)	(0.00949)		(0.00947)	
Initial GDPpc (log)	-0.0359***	-0.0510***	-0.0434***	-0.0474***	-0.0396***
	(0.00888)	(0.0101)	(0.00867)	(0.00938)	(0.00785)
Increase in NR exports	0.424***	0.436***	0.397***	0.419***	0.376***
	(0.103)	(0.0962)	(0.0937)	(0.0991)	(0.0962)
Avg. years of schooling (initial)		0.0115***	0.0141***		
		(0.00285)	(0.00296)		
Secondary school attainment (initial)				0.00118***	0.00139***
				(0.000294)	(0.000311)
Tertiary school attainment (initial)				0.00108**	0.00157***
				(0.000487)	(0.000484)
Constant	0.393***	0.450***	0.366***	0.443***	0.364***
	(0.0817)	(0.0840)	(0.0655)	(0.0825)	(0.0658)
Observations	673	673	673	673	673
R-squared	0.140	0.180	0.164	0.180	0.163
Adjusted R-square	0.131	0.170	0.155	0.169	0.153
Time FE	Yes	Yes	Yes	Yes	Yes
Country FE	No	No	No	No	No

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

4.4.3 Further analysis – additional control variables

Our results thus far point to average years of education being a stronger predictor of both long-term and short-term economic growth than the ECI. In this sub-section, we expand our analysis by bringing in additional variables and assessing whether the link between the initial ECI and economic growth is robust to including additional determinants of economic growth in the model.

Omitted variable variable bias is often a concern in the economic complexity literature and in regressions looking at economic growth more broadly. As a robustness check, we investigate whether adding other key variables changes the association between economic complexity, education and economic growth. We consider three variables that are recognised as important determinants of economic development across countries and considered in existing literature seeking to explain economic growth across countries – namely, exports to GDP to capture countries' trade openness, the share of employment in industry to capture the degree of industrialisation and importance of the manufacturing sector, and rule of law to proxy for institutional quality.

Table 4.3 shows the regression results for the long-term growth period. Column (1) shows our main model with initial ECI, income per capita, increase in natural resource exports and average years of schooling, while the subsequent columns add each new variable. The initial exports to GDP and employment share in industry variables both show positive coefficients, statistically significant at the 5% level, and both make similar contributions in terms of the explanatory power they add to the model (columns 2 and 3). In contrast, rule of law (added in column 4), does not appear to be statistically significant and we see a lower adjusted R-squared than the one in column (1). Adding these variables to the model leads to very small decreases in the magnitude of average years of education, which remains statistically significant. In contrast, the ECI loses statistical significance when the employment share in industry is added to the model (column 3), while showing slightly larger coefficients when the other two variables are added. Overall, the effect of initial average years of schooling on subsequent economic growth seems more consistent and robust to the inclusion of additional variables than the ECI.

Tables 4.A.9 and 4.A.10 in the Appendix show results for the same analysis for the 10- and 5-year growth periods – in both cases, the additional control variables show no statistically significant association with economic growth, and the main variables follow the patterns seen here.[4]

[4] Due to data availability constraints, Table 4.A.10 for 5-year growth periods includes only four periods,

Table 4.3: Economic complexity and 20-year growth in GDPpc, additional variables

	GDPpc growth (2000-2019)			
Variables	(1)	(2)	(3)	(4)
Initial ECI	0.0658**	0.0771**	0.0597*	0.0679**
	(0.0324)	(0.0321)	(0.0326)	(0.0333)
Initial GDPpc (log)	-0.230***	-0.237***	-0.263***	-0.227***
	(0.0348)	(0.0350)	(0.0382)	(0.0409)
Increase in NR exports	0.0898***	0.0759***	0.0926***	0.0895***
	(0.0262)	(0.0263)	(0.0277)	(0.0264)
Avg. years of schooling (initial)	0.0546***	0.0505***	0.0525***	0.0546***
	(0.0106)	(0.0107)	(0.0107)	(0.0107)
Exports to GDP (initial)		0.220**		
		(0.0979)		
Employment share industry (initial)			0.00779**	
			(0.00384)	
Rule of law (initial)				-0.00557
				(0.0365)
Constant	2.075***	2.103***	2.237***	2.053***
	(0.289)	(0.290)	(0.297)	(0.345)
Observations	136	136	136	136
R-squared	0.323	0.340	0.339	0.323
Adjusted R-square	0.302	0.314	0.313	0.297
Time FE	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

4.5 Conclusion

This chapter investigated the links between economic complexity, education and economic growth, with a focus on understanding whether the ECI is more closely associated with future economic growth than standard education variables. Contrary to Hausmann et al. (2014b), we find that education variables are better at explaining future economic growth across countries than the ECI. While the discrepancy in results can be due to differences in the time periods, the sample of countries, or education variables (for secondary and tertiary education, this chapter focused on schooling attainment, whereas Hausmann et al. (2014b) consider enrolment rates), we would not expect the link between complexity and subsequent economic growth to differ based on this type of difference across studies. Moreover, as in Chapter 3, we do not find any statistically significant association between the ECI and subsequent economic growth within countries.

Our further analysis pointed to education being a more robust predictor of economic growth, with more consistency in terms of statistical significance when other variables

starting from 2000, rather than from 1995 as in the previous tables, thus explaining the different coefficient sizes.

are included in the analysis. Importantly, this suggests that the link between economic complexity and future economic growth might not be as robust as the literature suggests (e.g., Hidalgo, 2023), and that more traditional measures such as schooling should not be overlooked when attempting to explain differences in economic growth across countries.

Here, we presented an initial exploratory analysis. Further analysis should, firstly, seek to understand more systematically the link between economic complexity and subsequent economic growth and what the implications are – what the ECI captures, and ultimately how well it can help explain subsequent economic growth, is likely to differ significantly between countries at different development stages and income levels. Secondly, it should explore the context-specific nature of the links between education, economic complexity and economic growth, and seek to understand whether economic complexity in particular can bring new insights for economic growth, and when and where this is the case.

4.A Appendix

Table 4.A.1: Description, source and availability of variables.

Variable	**Definition**	**Source**	**Years**
ECI	Economic Complexity Index based on HS-92 classification. Own calculations	The Observatory of Economic Complexity	1995-2019
Exports	Total merchandise exports (USD value)	The Observatory of Economic Complexity	1995-2019
NR exports	Natural resource exports (total USD value). Own calculation based on HS section V (mineral products) covering Chapters 25-27	The Observatory of Economic Complexity	1995-2019
GDP per capita	GDP per capita, PPP (constant 2017 international $)	World Bank Open Data	1995-2019
GDP	GDP (current USD)	World Bank Open Data	1995-2019
Average years of schooling	Average years of schooling attained in population over 25	Barro and Lee (2013)	1995, 2000, 2005, 2010, 2015
Secondary schooling attainment	Percentage of secondary schooling attained in population over 25	Barro and Lee (2013)	1995, 2000, 2005, 2010, 2015
Tertiary schooling attainment	Percentage of tertiary schooling attained in population over 25	Barro and Lee (2013)	1995, 2000, 2005, 2010, 2015
Employment share in industry	Employment in industry (% total employment)	World Bank Open Data	1995-2019
Rule of law	Perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence	World Bank Worldwide Governance Indicators	1996-2019

Table 4.A.2: Descriptive statistics, 20-year growth regression variables

Variables	N	Mean	SD	Min	Max
GDPpc growth	136	0.419	0.342	-0.406	1.609
Initial ECI	139	0.0915	0.985	-3.737	2.527
Initial GDPpc (log)	136	9.145	1.244	6.447	11.54
Increase in NR exports (share of initial GDP)	139	0.256	0.621	-0.139	5.442
Avg. years of schooling (initial)	139	7.582	3.162	0.976	13.37
Secondary school attainment (%, initial)	139	39.54	21.76	2.420	86.64
Tertiary school attainment (%, initial)	139	12.81	11.10	0.150	56.83
Exports to GDP (initial)	139	0.330	0.220	0.0266	1.291
Employment share in industry (initial)	139	20.86	8.549	2.260	39.95
Rule of law (initial)	139	0.0320	0.994	-1.905	1.985

Table 4.A.3: Descriptive statistics, 10-year growth regression variables

Variables	N	Mean	SD	Min	Max
GDPpc growth	273	0.194	0.196	-0.590	0.980
Initial ECI	279	0.0933	0.995	-3.737	2.558
Initial GDPpc (log)	273	9.269	1.220	6.447	11.68
Increase in NR exports (share of initial GDP)	279	0.0695	0.189	-0.178	1.556
Avg. years of schooling (initial)	279	8.092	3.186	0.976	13.58
Secondary school attainment (%, initial)	279	41.98	21.55	2.420	92.44
Tertiary school attainment (%, initial)	279	14.70	12.67	0.0200	67.24
Exports to GDP (initial)	279	0.321	0.205	0.0266	1.291
Employment share in industry (initial)	279	20.58	8.250	2.260	56.33
Rule of law (initial)	279	0.0277	0.999	-1.905	1.985

Table 4.A.4: Descriptive statistics, 5-year growth regression variables

Variables	N	Mean	SD	Min	Max
GDPpc growth	673	0.0849	0.109	-0.614	0.634
Initial ECI	691	0.116	0.973	-3.737	2.667
Initial GDPpc (log)	673	9.264	1.219	6.151	11.69
Increase in NR exports (share of initial GDP)	690	0.0274	0.0746	-0.257	0.713
Avg. years of schooling (initial)	691	8.078	3.198	0.869	13.64
Secondary school attainment (%, initial)	691	41.45	21.15	2.250	92.44
Tertiary school attainment (%, initial)	691	15.00	12.94	0.0200	67.85
Exports to GDP (initial)	690	0.307	0.200	0.0160	1.401
Employment share in industry (initial)	691	20.76	8.272	2.260	56.33
Rule of law (initial)	558	0.0307	1.004	-2.032	2.089

Table 4.A.5: List of countries included in the analysis, grouped by income classification

Low income (15):	Tanzania	Barbados
Burundi	Ukraine	Brunei
Central African Republic	Vietnam	Canada
Democratic Republic of the Congo	Zimbabwe	Switzerland
Gambia	Lesotho	Chile
Mali	Eswatini	Cyprus
Mozambique		Czechia
Malawi	***Upper middle income (36):***	Germany
Niger	Albania	Denmark
Rwanda	Argentina	Spain
Sudan	Armenia	Estonia
Sierra Leone	Bulgaria	Finland
Syria	Belize	France
Togo	Brazil	United Kingdom
Uganda	China	Greece
Zambia	Colombia	Hong Kong
	Costa Rica	Croatia
Lower middle income (38):	Cuba	Hungary
Benin	Dominican Republic	Ireland
Bangladesh	Ecuador	Iceland
Bolivia	Fiji	Israel
Cote d'Ivoire	Gabon	Italy
Cameroon	Guatemala	Japan
Republic of the Congo	Guyana	South Korea
Algeria	Iraq	Kuwait
Egypt	Jamaica	Lithuania
Ghana	Jordan	Latvia
Honduras	Kazakhstan	Macau
Haiti	Libya	Malta
Indonesia	Moldova	Netherlands
India	Maldives	Norway
Iran	Mexico	New Zealand
Kenya	Mauritius	Panama
Kyrgyzstan	Malaysia	Poland
Cambodia	Peru	Portugal
Laos	Paraguay	Qatar
Sri Lanka	Russia	Romania
Morocco	Thailand	Saudi Arabia
Myanmar	Turkey	Singapore
Mongolia	Venezuela	Slovakia
Mauritania	South Africa	Slovenia
Nicaragua	Botswana	Sweden
Nepal	Namibia	Trinidad and Tobago
Pakistan	Serbia	Uruguay
Philippines		United States
Papua New Guinea	***High income (51):***	Belgium
Senegal	United Arab Emirates	Luxembourg
El Salvador	Australia	
Tajikistan	Austria	
Tunisia	Bahrain	

Table 4.A.6: Economic complexity and 10-year growth in GDPpc, OLS, full sample

	GDPpc growth (2000-09, 2010-19)				
Variables	(1)	(2)	(3)	(4)	(5)
Initial ECI	0.0738***	0.0398*		0.0378*	
	(0.0201)	(0.0201)		(0.0206)	
Initial GDPpc (log)	-0.0866***	-0.123***	-0.109***	-0.115***	-0.102***
	(0.0200)	(0.0213)	(0.0176)	(0.0195)	(0.0158)
Increase in NR exports	0.235***	0.220***	0.172***	0.192***	0.144**
	(0.0720)	(0.0586)	(0.0573)	(0.0595)	(0.0582)
Secondary school attainment (initial)				0.00290***	0.00326***
				(0.000658)	(0.000665)
Tertiary school attainment (initial)				0.00321***	0.00414***
				(0.00110)	(0.00107)
Avg. years of schooling (initial)		0.0289***	0.0337***		
		(0.00563)	(0.00573)		
Constant	0.968***	1.088***	0.933***	1.081***	0.949***
	(0.183)	(0.179)	(0.134)	(0.173)	(0.133)
Observations	273	273	273	273	273
R-squared	0.159	0.240	0.224	0.240	0.227
Adjusted R-square	0.146	0.226	0.213	0.223	0.212
Time FE	Yes	Yes	Yes	Yes	Yes
Country FE	No	No	No	No	No

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4.A.7: Economic complexity and 10-year growth in GDPpc, Fixed Effects, full sample

	GDPpc growth (2000-09, 2010-19)				
Variables	(1)	(2)	(3)	(4)	(5)
Initial ECI	-0.00140	-0.00132		-2.89e-05	
	(0.0380)	(0.0384)		(0.0393)	
Initial GDPpc (log)	-0.591***	-0.591***	-0.591***	-0.587***	-0.587***
	(0.0806)	(0.0807)	(0.0792)	(0.0804)	(0.0787)
Increase in NR exports	0.227***	0.227***	0.227***	0.230***	0.230***
	(0.0703)	(0.0704)	(0.0681)	(0.0704)	(0.0681)
Avg. years of schooling (initial)		-0.00131	-0.00134		
		(0.0194)	(0.0192)		
Secondary school attainment (initial)				-0.00127	-0.00127
				(0.00176)	(0.00174)
Tertiary school attainment (initial)				-0.00172	-0.00172
				(0.00341)	(0.00332)
Constant	5.592***	5.602***	5.600***	5.624***	5.623***
	(0.736)	(0.748)	(0.739)	(0.738)	(0.725)
Observations	273	273	273	273	273
R-squared	0.512	0.512	0.512	0.514	0.514
Adjusted R-square	0.504	0.503	0.504	0.503	0.505
Number of country	137	137	137	137	137
Time FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4.A.8: Economic complexity and 5-year growth in GDPpc, Fixed Effects, full sample

	GDPpc growth (1995-99, 2000-04, 2005-09, 2010-14, 2015-19)				
Variables	(1)	(2)	(3)	(4)	(5)
Initial ECI	0.0128	0.0124		0.0126	
	(0.0146)	(0.0146)		(0.0145)	
Initial GDPpc (log)	-0.149***	-0.148***	-0.147***	-0.149***	-0.148***
	(0.0228)	(0.0230)	(0.0231)	(0.0229)	(0.0230)
Increase in NR exports	0.588***	0.593***	0.599***	0.591***	0.597***
	(0.0994)	(0.0976)	(0.102)	(0.0993)	(0.103)
Avg. years of schooling (initial)		0.00705	0.00734		
		(0.00633)	(0.00634)		
Secondary school attainment (initial)				0.000421	0.000448
				(0.000613)	(0.000625)
Tertiary school attainment (initial)				6.69e-05	9.65e-05
				(0.000966)	(0.000979)
Constant	1.408***	1.355***	1.344***	1.396***	1.386***
	(0.208)	(0.217)	(0.218)	(0.208)	(0.210)
Observations	673	673	673	673	673
R-squared	0.261	0.262	0.261	0.262	0.260
Adjusted R-square	0.253	0.254	0.253	0.252	0.251
Number of country	137	137	137	137	137
Time FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4.A.9: Economic complexity and 10-year growth in GDPpc, full sample, additional variables

	GDPpc growth (2000-09, 2010-19)			
Variables	(1)	(2)	(3)	(4)
Initial ECI	0.0398*	0.0414**	0.0383*	0.0391*
	(0.0201)	(0.0197)	(0.0200)	(0.0207)
Initial GDPpc (log)	-0.123***	-0.126***	-0.134***	-0.124***
	(0.0213)	(0.0217)	(0.0236)	(0.0239)
Increase in NR exports	0.220***	0.203***	0.220***	0.221***
	(0.0586)	(0.0598)	(0.0605)	(0.0619)
Avg. years of schooling (initial)	0.0289***	0.0280***	0.0284***	0.0289***
	(0.00563)	(0.00565)	(0.00570)	(0.00565)
Exports to GDP (initial)		0.0740		
		(0.0591)		
Employment share in industry (initial)			0.00272	
			(0.00179)	
Rule of law (initial)				0.00180
				(0.0198)
Constant	1.088***	1.096***	1.136***	1.095***
	(0.179)	(0.180)	(0.186)	(0.199)
Observations	273	273	273	273
R-squared	0.240	0.245	0.247	0.240
Adjusted R-square	0.226	0.228	0.230	0.223
Time FE	Yes	Yes	Yes	Yes
Country FE	No	No	No	No

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4.A.10: Economic complexity and 5-year growth in GDPpc, full sample, additional variables

	GDPpc growth (2000-04, 2005-09, 2010-14, 2015-19)			
Variables	(1)	(2)	(3)	(4)
Initial ECI	0.0195*	0.0198*	0.0189*	0.0190*
	(0.0102)	(0.0101)	(0.0101)	(0.0102)
Initial GDPpc (log)	-0.0561***	-0.0569***	-0.0590***	-0.0568***
	(0.0124)	(0.0127)	(0.0135)	(0.0141)
Increase in NR exports	0.403***	0.393***	0.400***	0.405***
	(0.0918)	(0.0927)	(0.0929)	(0.0972)
Avg. years of schooling (initial)	0.0130***	0.0128***	0.0128***	0.0130***
	(0.00328)	(0.00327)	(0.00329)	(0.00330)
Exports to GDP (initial)		0.0203		
		(0.0319)		
Employment share in industry (initial)			0.000774	
			(0.000935)	
Rule of law (initial)				0.00145
				(0.00975)
Constant	0.502***	0.503***	0.513***	0.507***
	(0.0974)	(0.0984)	(0.101)	(0.111)
Observations	546	546	546	546
R-squared	0.190	0.191	0.192	0.190
Adjusted R-square	0.179	0.179	0.180	0.178
Time FE	Yes	Yes	Yes	Yes
Country FE	No	No	No	No

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4.A.11: Economic complexity and 10-year growth in GDPpc, income group effects

	GDPpc growth (2000-09, 2010-19)		
Variables	(1)	(2)	(3)
Initial ECI	0.0738***	0.0320	0.00463
	(0.0201)	(0.0277)	(0.0272)
Initial GDPpc (log)	-0.0866***	-0.167***	-0.180***
	(0.0200)	(0.0262)	(0.0263)
Increase in NR exports (share of initial GDP)	0.235***	0.243***	0.228***
	(0.0720)	(0.0677)	(0.0601)
Avg. years of schooling (initial)			0.0210***
			(0.00573)
Income group *(ref: High Income)*			
Middle Income		-0.134***	-0.118**
		(0.0471)	(0.0459)
Low Income		-0.495***	-0.417***
		(0.111)	(0.110)
Income group * Initial ECI *(ref: High Income)*			
Middle Income * Initial ECI		0.0885**	0.0901***
		(0.0350)	(0.0339)
Low Income * Initial ECI		-0.0149	0.00135
		(0.0619)	(0.0619)
Constant	0.968***	1.840***	1.791***
	(0.183)	(0.262)	(0.257)
Observations	273	273	273
R-squared	0.159	0.292	0.331
Adjusted R-square	0.146	0.271	0.309
Time FE	Yes	Yes	Yes
Country FE	No	No	No

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4.A.12: Economic complexity and 5-year growth in GDPpc, income group effects

	GDPpc growth (1995-99 to 2015-19)		
Variables	(1)	(2)	(3)
Initial ECI	0.0348***	0.0179*	0.00820
	(0.00960)	(0.0101)	(0.00946)
Initial GDPpc (log)	-0.0359***	-0.0718***	-0.0777***
	(0.00888)	(0.0118)	(0.0127)
Increase in NR exports (share of initial GDP)	0.424***	0.452***	0.457***
	(0.103)	(0.0995)	(0.0955)
Avg. years of schooling (initial)			0.00824***
			(0.00294)
Income group *(ref: High Income)*			
Middle Income		-0.0613***	-0.0547***
		(0.0203)	(0.0199)
Low Income		-0.221***	-0.193***
		(0.0465)	(0.0463)
Income group * Initial ECI *(ref: High Income)*			
Middle Income * Initial ECI		0.0398***	0.0399***
		(0.0146)	(0.0145)
Low Income * Initial ECI		-0.0201	-0.0154
		(0.0304)	(0.0317)
Constant	0.393***	0.778***	0.769***
	(0.0817)	(0.117)	(0.119)
Observations	673	673	673
R-squared	0.140	0.215	0.234
Adjusted R-square	0.131	0.202	0.220
Time FE	Yes	Yes	Yes
Country FE	No	No	No

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Conclusion

Over these four chapters, this thesis has provided a critical evaluation of the ECI, focusing both on theoretical and conceptual issues, as well as specific empirical questions. Here, we reflect on the evidence provided, discuss some limitations of our analysis, and suggest areas for further research.

Chapter 1 presented an extensive overview of the theoretical and empirical foundations of the ECI, at both national and sub-national levels, and argued that while the ECI represents a new and advanced method, it still presents important drawbacks, particularly around theoretical grounding, data used and empirical estimation challenges. The remainder of the thesis has explored key empirical questions, with a broad focus on understanding what economic complexity captures in different settings and applications.

Chapter 2 focused on the sub-national level using employment data for the case of Portugal and compared occupation and industry-based complexity measures. We explored the complexity levels of regions and activities and how they evolve over time, as well as their association with employment growth. We found a negative association between both economic complexity measures and employment growth, though the association is not statistically significant when we control for initial employment and education levels, suggesting that it was capturing convergence across regions. Moreover, the chapter linked this literature with the one on the task-based approach to labour markets. In this particular case, the ECI might reflect specialisation patterns more strongly than it reflects underlying 'capabilities' per se.

In Chapter 3, we returned to the original geographical level in which economic complexity was first introduced, and looked at exports-based ECI across the world, with a focus on countries that are very heavily dependent on natural resources. Our results, based on the GCCs and their dependence on oil, a highly volatile natural resource, pointed to the possibility that the ECI in those countries reflects shifts in the price and demand for oil, rather than reflecting underlying capabilities within those countries. Importantly, while we found a cross-country association between the initial ECI and subsequent economic growth, the same was not true for the Fixed Effects specifications, suggesting that the ECI cannot help us understand economic growth within countries.

In Chapter 4, we took a step further to ask whether the export-based ECI is indeed more closely associated with future economic growth across countries than simpler variables that have long been recognised as important for economic development, focusing on education. In contrast with the original proponents' results, we found that education variables can explain more of the variance in future economic growth than the ECI. Once again, however, this is only a cross-country phenomenon and in our Fixed Effects specifications the coefficient on economic complexity was not statistically significant.

Overall, this thesis presented a critical survey of the economic complexity literature, both in terms of its sub-national and country-level theoretical foundations and empirical findings, it went on to explore pressing empirical questions at different geographical levels and considering different types of data, and finished with an exploratory analysis of the links between economic complexity, education and economic growth. This research has raised questions over the power of the ECI in explaining future income growth and, importantly, in guiding policy in meaningful ways across all settings and geographical levels. It should also caution those using the ECI in policy settings not to disregard traditional variables, such as education measures or simpler measures of diversification, and to consider more carefully the local context and specificity.

Limitations

While these four chapters contribute to our existing understanding of complexity and raise important questions and discussions within this literature, there are important limitations to discuss. This thesis focused mostly on the method of reflections, proposed by the initial contributors. As discussed, the literature has since moved on to consider different interpretations of this method, alternative methods such as the Fitness Complexity Index (Tacchella et al., 2012), as well as different measures such as Structural Diversity (Broekel, 2019). These might help address some of the issues identified with the ECI, but may also come with different conceptual challenges, which were not discussed at length here. Focusing on the method of reflections also allows us to maintain some consistency in terms of methodology throughout the different chapters. The sub-national application to employment data and for the specific case of Portugal is of course different from the data used in the remaining chapters, and therefore the main implications derived there are specific to that type of data and possibly to the geographical context considered too.

This research also focused mainly on export data, as did the original application of the

ECI, as well as on employment data for the second chapter. As discussed at length, there are different drawbacks to these different types of data. Focusing on exports allowed us to analyse the ECI using the data and methodology originally proposed, in order to assess the case of natural resource dependent countries, as well as to explore the implications of this indicator. In the second chapter, we explored one additional type of data, allowing us to compare directly the application of the ECI methodology to both industries and occupations and see how they fare in terms of reflecting economic activity and employment growth outcomes across Portuguese regions. Here too, we could have explored other types of data. Patent data has drawn particular attention, initially used at the sub-national level (Balland & Rigby, 2017) and, more recently, across countries too (Stojkoski et al., 2023). Such applications focus on technological capabilities and, particularly at the sub-national level, tend to be well grounded in theory and provide insightful analyses. Still, on the one hand, they have been widely applied to the European context while, on the other, they do not cover 'lagging' regions, which tend to have low patent counts (as discussed in Chapter 2). As a result, we focused on exports and employment throughout the thesis.

What does the ECI capture?

The question of what exactly the ECI captures remains. On the one hand, in some sub-national contexts, spatial measures might be too highly influenced by the typical distribution of activity across space, which we already know a lot about within urban and regional economics research. On the other hand, there is a question of outliers and cases where the specialisation area of a country might affect the ECI measure itself and render it useless for policy-making or for explaining outcomes across countries. Lastly, the evidence on the capacity of the ECI for predicting outcomes such as economic growth might be more mixed than what initial findings suggested.

Moreover, the idea that the ECI can capture productive knowledge and its variety across countries is, of course, based on the implicit assumption that a country's exports allow us to infer this accurately and disregards that, through intermediate trade, countries export some goods despite not having the full productive knowledge embedded in that product. The same is inevitably true for sub-national regions across the world.

As described, the vast amount of literature that emerged over the past years associated the ECI with several different outcomes. At the same time, more recent papers have started to look at the determinants of economic complexity, focusing instead on the ECI as the main dependent variable. The broader and more important question then is what this seemingly intermediate variable can tell us that we did not know before and how it can help us with development policy across countries and regions.

Economic complexity and economic growth – estimation challenges

An important issue, particularly in the face of this vast number of papers using economic complexity methods, is the question of causality. While a recent review paper by Hidalgo (2021) claimed that the link between economic complexity and subsequent economic growth is highly robust, and that the direction of causality has to be from economic complexity to economic growth, there are important estimation challenges that remain, and causality is far from being confirmed.

On the one hand, the estimation methods used in the economic complexity literature are not robust enough to be considered as evidence of causal links; moreover, given that the ECI is a composite indicator, it is very unlikely that advanced causal identification methods such as instrumental variables or randomised experiments will be convincing enough in establishing causal links between the ECI and relevant outcomes.

On the other hand, there is the issue of omitted variable bias – just like the literature on education suggested, there might be country or region characteristics that impact both economic growth and economic complexity. In fact, given the vast amount of aspects the ECI might reflect (for instance, thinking back to how Hidalgo and Hausmann, 2009 introduced their all-encompassing idea of capabilities), it is likely that what we are capturing in this association is not a link between economic complexity per se and economic growth, but the underlying aspects that make countries able to export certain products competitively. In that sense, economic complexity can be thought of as an intermediate variable between different country or regional characteristics and economic growth or other desirable outcomes.

If this is an intermediate variable, and given the difficulty in establishing a clear causal link (as is often the case in social sciences), economic complexity should instead be used as a descriptive tool that can help us understand the relative position of a country or region vis-a-vis their counterparts, as well as identify clear stylised facts. Even for descriptive purposes, the reference for comparison should be carefully considered – for instance, it may be more sensible to draw comparisons between countries or regions of comparable development levels. Moreover, while it is not unique to the ECI, there are extensive challenges with data availability and quality in lower income countries, where the sub-national dimension is rarely available even for standard variables.

Finally, we should consider where the ECI can and cannot help us understand economic development and guide policy. We should be cautious not to overlook simpler and well-established indicators such as a education, institutional quality or other diversity and exports measures, as well as consider local contexts carefully. When it comes to policy-making, several papers suggest national or regional governments should aim to increase economic complexity. However, it is not clear how to increase economic complexity, particularly given the wide-range of characteristics the ECI might reflect, which likely differ across countries with different income and development levels. Lastly, it might be easier to find causality between educational expenditure or attainment and subsequent economic development, for example through quasi-natural experiments, than it is for the case of economic complexity, due to its methodology pooling countries against each other and the yearly oscillations observed.

Further research

This leads us to future research in this area. Here, I propose three main avenues related to our conceptual understanding of economic complexity, new types of data and exploring the transition to green activities.

First, there are areas to explore concerning our conceptual understanding of economic complexity. We need to better understand when and where it is useful to rely on the ECI. In particular, what it captures in different geographic and socioeconomic contexts, whether it has the same meaning and broader conceptualisation at the sub-national and national levels, and whether there are specific contexts in which we should avoid using the ECI for policy guidance – for instance, in economies with high dependence on natural resources or other areas of specialisation that might affect measurement, countries with very low income levels or high poverty rates, or those that are highly dependent on external aid, along with several other circumstances that might inhibit accurate and appropriate analysis based on the ECI.

Moreover, there is still a gap in our understanding of how institutional contexts differ across countries and sub-national regions and how this interacts with economic complexity measurements, as well as how institutions can be leveraged in order to implement policy to achieve improved outcomes. While the importance of institutions has been acknowledged in the literature, and there is some existing research considering the links between inequality, complexity and institutions across countries (e.g., Hartmann et al., 2017), more can be done in this area to capture the role of institutions and actors in generating different economic complexity outcomes.

Second, there is a trend in the broader economic geography literature towards methods like web scraping and text analysis in order to collect and analyse data, thanks to fast improving computing power and software. New datasets are emerging that allow us to better understand, for example, activities that different firms are performing, or what kind of skills employers are seeking. This can allow us to track industrial changes and other regional dynamics much more quickly than in the past, when we were limited by industrial classifications, which are less flexible and take time to update, giving us a much more lagged view of economic activity. There are many opportunities here, and some existing examples – The Data City are using Artificial Intelligence to identify emerging sectors, clusters and companies across the UK and their data is being used by Raquel Ortega-Argilés and other researchers at the University of Manchester to measure relatedness, and their Real Time Industrial Classification could likely be used to measure economic complexity across the UK. As this technology and type of data evolve, we will likely see new datasets for different economic activities and locations, and thus there is a possibility that economic complexity and network-based methods can help us make the most of this data.

Third, the green transition has become a crucial area of concern for countries and regions across the world, as places need to evolve and adapt to the challenge of reducing emissions while still wanting to remain prosperous and at the forefront of new technological developments. This is an interesting area that will certainly draw a lot of attention from economic geographers and other social scientists. The ECI methodology has already been applied to a Green Complexity Index, and has been used to understand which countries and regions will likely be able to transition towards greener economic activities more smoothly. This already includes work by Mealy and Teytelboym (2020) and, more recently, by Andres et al. (2023) and Stojkoski et al. (2023), among others. There are several ways in which this might be continued, for instance by considering different types of data and regional contexts as well as developing tools that can be accessed by policymakers and other stakeholders.

Bibliography

Abramovitz, M. (1986). Catching up, forging ahead, and falling behind. *The Journal of Economic History*, *46*(2), 385–406. https://doi.org/10.1017/S0022050700046209

Abramovitz, M. (1994a). Catch-up and convergence in the postwar growth boom and after. In W. J. Baumol, R. R. Nelson & E. N. Wolf (Eds.), *Convergence of Productivity: Cross-national studies and historical evidence* (pp. 86–125). Oxford University Press.

Abramovitz, M. (1994b). The origins of the postwar catch-up and convergence boom. In J. Fagerberg, B. Verspagen & N. von Tunzelmann (Eds.), *The Dynamics of Technology, Trade and Growth* (pp. 21–52). Edward Elgar.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *The Quarterly Journal of Economics*, *117*(4), 1231–1294. https://doi.org/10.1162/003355302320935025

Acemoglu, D., & Autor, D. H. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of Labor Economics* (pp. 1043–1171, Vol. 4). Elsevier.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, *91*(5), 1369–1401. https://doi.org/10.1257/aer.91.5.1369

Acemoglu, D., & Zilibotti, F. (1999). Information accumulation in development. *Journal of Economic Growth*, *4*(1), 5–38.

Aghion, P., Boustan, L., Hoxby, C., & Vandenbussche, J. (2009). The causal impact of education on economic growth: Evidence from U.S. *Brookings Papers on Economic Activity*, *1*(1), 1–73.

Ajide, K. B. (2022). Is natural resource curse thesis an empirical regularity for economic complexity in Africa? *Resources Policy*, *76*, 102755. https://doi.org/10.1016/j.resourpol.2022.102755

Albeaik, S., Kaltenberg, M., Alsaleh, M., & Hidalgo, C. A. (2017). *729 new measures of economic complexity (Addendum to Improving the Economic Complexity Index)*. http://arxiv.org/abs/1708.04107

Alexeev, M., & Conrad, R. (2009). The elusive curse of oil. *The Review of Economics and Statistics*, *91*(3), 586–598. https://doi.org/10.1162/rest.91.3.586

Andres, P., Mealy, P., Handler, N., & Fankhauser, S. (2023). Stranded nations? Transition risks and opportunities towards a clean economy. *Environmental Research Letters*, *18*(4), 045004.

Antonelli, C. (1998). Localized technological change, new information technology and the knowledge-based economy: The European evidence. *Journal of Evolutionary Economics*, *8*(2), 177–198. https://doi.org/10.1007/s001910050061

Antonelli, C. (2008a). The dynamic legacies: Smith, Marx, Marshall and Schumpeter. In *Localised technological change: Towards the economics of complexity* (pp. 17–46). Routledge.

Antonelli, C. (2008b). Localised technological change: A critical assessment. In *Localised technological change: Towards the economics of complexity* (pp. 47–62). Routledge.

Antonelli, C. (2009). The economics of innovation: From the classical legacies to the economics of complexity. *Economics of Innovation and New Technology*, *18*(7), 611–646. https://doi.org/10.1080/10438590802564543

Antonelli, C., Crespi, F., Mongeau Ospina, C. A., & Scellato, G. (2017). Knowledge composition, Jacobs externalities and innovation performance in European regions. *Regional Studies*, *51*(11), 1708–1720. https://doi.org/10.1080/00343404.2016.1217405

Antonelli, C., Crespi, F., & Quatraro, F. (2020). Knowledge complexity and the mechanisms of knowledge generation and exploitation: The European evidence. *Research Policy*, 104081. https://doi.org/10.1016/j.respol.2020.104081

Antonietti, R., & Franco, C. (2021). From FDI to economic complexity: A panel Granger causality analysis. *Structural Change and Economic Dynamics*, *56*, 225–239. https://doi.org/10.1016/j.strueco.2020.11.001

Apergis, N., & Payne, J. E. (2014). The oil curse, institutional quality, and growth in MENA countries: Evidence from time-varying cointegration. *Energy Economics*, *46*, 1–9. https://doi.org/10.1016/j.eneco.2014.08.026

Arman, H., Iammarino, S., Ibarra-Olivo, J. E., & Lee, N. (2021a). *Breaking out of the innovation trap? Towards promoting private R&D investment in Kuwait*. London School of Economics Middle East Centre Kuwait Programme Paper Series. http://eprints.lse.ac.uk/109010/

Arman, H., Iammarino, S., Ibarra-Olivo, J. E., & Lee, N. (2021b). Systems of innovation, diversification, and the R&D trap: A case study of Kuwait. *Science and Public Policy*, 1–12.

Asheim, B. T. (1996). Industrial districts as 'learning regions': A condition for prosperity. *European Planning Studies*, *4*(4), 379–400. https://doi.org/10.1080/09654319608720354

Asheim, B. T., & Gertler, M. S. (2005). The geography of innovation: Regional innovation systems. In J. Fagerberg, D. C. Mowery & R. R. Nelson (Eds.), *The Oxford handbook of innovation* (pp. 219–317). Oxford University Press.

Asheim, B. T., & Isaksen, A. (2002). Regional innovation systems: The integration of local 'sticky' and global 'ubiquitous' knowledge. *Journal of Technology Transfer*, *27*, 77–83.

Atkinson, A. B., & Stiglitz, J. E. (1969). A new view of technological change. *The Economic Journal*, *79*(315), 573. https://doi.org/10.2307/2230384

Audretsch, D. B. (1997). Technological regimes, industrial demography and the evolution of industrial structures. *Industrial and Corporate Change*, *6*(1), 49–82. https://doi.org/10.1093/icc/6.1.49

Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *The American Economic Review*, *86*(3), 630–640. https://www.jstor.org/stable/2118216

Audretsch, D. B., & Feldman, M. P. (2004). Knowledge spillovers and the geography of innovation. In J. Henderson & J. Thisse (Eds.), *Handbook of Regional and Urban Economics* (pp. 2713–2739). Elsevier Inc. http://dx.doi.org/10.1016/S1574-0080(04)80018-X

Autor, D. H. (2013). *The task approach to labor markets: An overview*. NBER Working Paper 18711. http://www.nber.org/papers/w18711

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, *118*(4), 1279–1333. https://doi.org/10.1162/003355303322552801

Auty, R. (1993, April). *Sustaining development in mineral economies: The resource curse thesis*. Routledge. https://doi.org/10.4324/9780203422595

Avom, D., Keneck-Massil, J., Njangang, H., & Nvuh-Njoya, Y. (2022). Why are some resource-rich countries more sophisticated than others? The role of the regime type and political ideology. *Resources Policy*, *79*, 103067. https://doi.org/10.1016/j.resourpol.2022.103067

Badeeb, R. A., Lean, H. H., & Clark, J. (2017). The evolution of the natural resource curse thesis: A critical literature survey. *Resources Policy*, *51*(October 2016), 123–134. https://doi.org/10.1016/j.resourpol.2016.10.015

Bahar, D., Rapoport, H., & Turati, R. (2020). Birthplace diversity and economics complexity: Cross-country evidence. *Research Policy*, 103991. https://doi.org/10.1016/j.respol.2020.103991

Balassa, B. (1965). Trade liberalisation and "revealed" comparative advantage. *The Manchester School*, *33*(2), 99–123.

Balland, P. A. (2017). *Economic Geography in R: Introduction to the EconGeo Package*. SSRN. https://dx.doi.org/10.2139/ssrn.2962146

Balland, P. A., & Boschma, R. (2021). Complementary interregional linkages and Smart Specialisation: An empirical study on European regions. *Regional Studies*, 1–12. https://doi.org/10.1080/00343404.2020.1861240

Balland, P. A., Boschma, R., Crespo, J., & Rigby, D. (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., Broekel, T., Diodato, D., Giuliani, E., Hausmann, R., O'Clery, N., & Rigby, D. (2022). The new paradigm of economic complexity. *Research Policy*, *51*(3), 104450. https://doi.org/10.1016/j.respol.2021.104450

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

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

Barro, R. J. (1991). Economic growth in a cross section of countries. *The Quarterly Journal of Economics*, *106*(2), 407. https://doi.org/10.2307/2937943

Barro, R. J. (2013). Education and economic growth. *Annals of Economics and Finance*, *142*, 301–328.

Barro, R. J., & Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics*, *104*, 184–198. https://doi.org/10.1016/j.jdeveco.2012.10.001

Barro, R. J., & Lee, J.-W. (1994). Sources of economic growth. *Carnegie-Rochester Conference Series on Public Policy*, *40*, 1–46. https://doi.org/10.1016/0167-2231(94)90002-7

Basile, R., Cicerone, G., & Iapadre, L. (2019). *Economic complexity and regional labor productivity distribution: Evidence from Italy.*

Beblawi, H. E. (2011). Gulf industrialization in perspective. In J. F. Seznec & M. Kirk (Eds.), *Industrialization in the Gulf: A Socioeconomic Revolution* (pp. 185–97). London: Center for Contemporary Arab Studies. Georgetown University/Routledge.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education.* Columbia University Press.

Bell, M. (2009). *Innovation capabilities and directions of development.* STEPS Working Paper 33. Brighton.

Bell, M., & Pavitt, K. (1995). The development of technological capabilities. In *Trade, technology and international competitiveness* (pp. 69–101, Vol. 22).

Benhabib, J., & Spiegel, M. M. (1994). The role of human capital in economic development evidence from aggregate cross-country data. *Journal of Monetary Economics*, *34*(2), 143–173. https://doi.org/10.1016/0304-3932(94)90047-7

Benos, N., & Zotou, S. (2014). Education and economic growth: A meta-regression analysis. *World Development*, *64*, 669–689. https://doi.org/10.1016/j.worlddev.2014.06.034

Bils, M., & Klenow, P. J. (2000). Does schooling cause growth? *The American Economic Review*, *90*(5), 1160–1183. https://www.jstor.org/stable/2677846

Bishop, A., Dobre, A. M., & Mateos-Garcia, J. (2018). Economic complexity and the emergence of new ideas. *SSRN Electronic Journal*, 1–33. https://doi.org/10.2139/ssrn.3242960

Bishop, A., & Mateos-Garcia, J. (2019). Exploring the link between economic complexity and emergent economic activities. *National Institute Economic Review*, *249*(1), R47–R58. https://doi.org/10.1177/002795011924900114

Bjorvatn, K., Farzanegan, M. R., & Schneider, F. (2012). Resource Curse and Power Balance: Evidence from Oil-Rich Countries. *World Development*, *40*(7), 1308–1316. https://doi.org/10.1016/j.worlddev.2012.03.003

Bodea, C., Higashijima, M., & Singh, R. J. (2016). Oil and civil conflict: Can public spending have a mitigation effect? *World Development*, *78*, 1–12. https://doi.org/10.1016/j.worlddev.2015.09.004

Boleti, E., Garas, A., Kyriakou, A., & Lapatinas, A. (2021). Economic complexity and environmental performance: Evidence from a world sample. *Environmental Modeling & Assessment*, *26*, 251–270.

Boschini, A. D., Pettersson, J., & Roine, J. (2007). Resource curse or not: A question of appropriability. *Scandinavian Journal of Economics*, *109*(3), 593–617. https://doi.org/10.1111/j.1467-9442.2007.00509.x

Boschma, R. (2017). Relatedness as driver of regional diversification: A research agenda. *Regional Studies*, *51*(3), 351–364. https://doi.org/10.1080/00343404.2016.1254767

Boschma, R. (2022). Global value chains from an evolutionary economic geography perspective: A research agenda. *Area Development and Policy*, *7*(2), 123–146. https://doi.org/10.1080/23792949.2022.2040371

Boschma, R., Balland, P. A., & Kogler, D. F. (2014). 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

Boschma, R., & Capone, G. (2016). Relatedness and diversification in the European Union (EU-27) and European Neighbourhood Policy countries. *Environment and Planning C: Government and Policy*, *34*(4), 617–637. https://doi.org/10.1177/0263774X15614729

Boschma, R., Coenen, L., Frenken, K., & Truffer, B. (2017). Towards a theory of regional diversification: Combining insights from Evolutionary Economic Geography and Transition Studies. *Regional Studies*, *51*(1), 31–45. https://doi.org/10.1080/00343404.2016.1258460

Boschma, R., & Frenken, K. (2011). The emerging empirics of evolutionary economic geography. *Journal of Economic Geography*, *11*(2), 295–307. https://doi.org/10.1093/jeg/lbq053

Boschma, R., & Iammarino, S. (2009). Related variety, trade linkages and regional growth in Italy. *Economic Geography*, *85*(3), 289–311. https://www.jstor.org/stable/40377307

Boschma, R., & Lambooy, J. G. (1999). Evolutionary economics and economic geography. *Journal of Evolutionary Economics*, *9*(4), 411–429. https://doi.org/10.1007/s001910050089

Breschi, S., Lissoni, F., & Malerba, F. (2003). Knowledge relatedness in firm technological diversification. *Research Policy*, *32*(1), 69–87.

Breschi, S., Malerba, F., & Orsenigo, L. (2000). Technological regimes and Schumpeterian patterns of innovation. *The Economic Journal*, *110*(463), 388–410. https://www.jstor.org/stable/2566240

Brinkley, I., Hutton, W., Schneider, P., & Ulichsen, K. (2012). *Kuwait and the Knowledge Economy*. London School of Economics.

Broekel, T. (2019). Using structural diversity to measure the complexity of technologies. *PLoS ONE*, *14*(5), 1–23. https://doi.org/10.1371/journal.pone.0216856

Broekel, T., Fitjar, R. D., & Haus-Reve, S. (2021). *The roles of diversity, complexity, and relatedness in regional development: What does the occupational perspective add?* Utrecht University Papers in Evolutionary Economic Geography, 21.35.

Brunnschweiler, C. N. (2008). Cursing the blessings? Natural resource abundance, institutions, and economic growth. *World Development*, *36*(3), 399–419. https://doi.org/10.1016/j.worlddev.2007.03.004

Brunnschweiler, C. N., & Bulte, E. H. (2008). The resource curse revisited and revised: A tale of paradoxes and red herrings. *Journal of Environmental Economics and Management*, *55*(3), 248–264. https://doi.org/10.1016/j.jeem.2007.08.004

Buckley, P. J., Pass, C. L., & Prescott, K. (1988). Measures of international competitiveness: A critical survey. *Journal of Marketing Management*, *4*(2), 175–200. https://doi.org/10.1080/0267257X.1988.9964068

Caines, C., Hoffmann, F., & Kambourov, G. (2017). Complex-task biased technological change and the labor market. *Review of Economic Dynamics*, *25*, 298–319. https://doi.org/10.1016/j.red.2017.01.008

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

Canh, N. P., Schinckus, C., & Thanh, S. D. (2020). The natural resources rents: Is economic complexity a solution for resource curse? *Resources Policy*, *69*(July), 101800. https://doi.org/10.1016/j.resourpol.2020.101800

Cantwell, J. (1989). *Technological innovation and multinational corporations*. Blackwell.

Cantwell, J., & Iammarino, S. (2003). Regional systems of innovation in Europe and the globalisation of technology. In *Multinational Corporations and European Regional Systems of Innovation* (pp. 1–21). Routledge.

Castaldi, C. (2009). The relative weight of manufacturing and services in Europe: An innovation perspective. *Technological Forecasting and Social Change*, *76*(6), 709–722. https://doi.org/10.1016/j.techfore.2008.08.001

Chamberlain, T. W., & Kalaitzi, A. S. (2020). Fuel-mining exports and economic growth: Evidence from the UAE. *International Advances in Economic Research*, *26*(1), 119–121. https://doi.org/10.1007/s11294-020-09766-4

Chenery, H. B. (1960). Patterns of industrial growth. *The American Economic Review*, *50*(4), 624–654.

Chu, L. K. (2021). Economic structure and environmental Kuznets curve hypothesis: New evidence from economic complexity. *Applied Economics Letters*, *28*(7), 612–616. https://doi.org/10.1080/13504851.2020.1767280

Coe, N. M., Dicken, P., & Hess, M. (2008). Global production networks: Realizing the potential. *Journal of Economic Geography*, *8*(3), 271–295. https://doi.org/10.1093/jeg/lbn002

Coe, N. M., & Yeung, H. W. C. (2019). Global production networks: Mapping recent conceptual developments. *Journal of Economic Geography*, *19*(4), 775–801. https://doi.org/10.1093/jeg/lbz018

Coe, N. M., & Yeung, H. W.-C. (2015). *Global production networks: Theorizing economic development in an interconnected world* (First edition). Oxford University Press.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, *35*(1), 128–152.

Colozza, F., Boschma, R., Morrison, A., & Pietrobelli, C. (2021, December). *The importance of global value chains and regional capabilities for the economic complexity of EU-regions*. MERIT Working Papers.

Content, J., & Frenken, K. (2016). Related variety and economic development: A literature review. *European Planning Studies*, *24*(12), 2097–2112. https://doi.org/10.1080/09654313.2016.1246517

Cooke, P. (2001). Regional innovation systems, clusters, and the knowledge economy. *Industrial and Corporate Change*, *10*(4), 945–974. https://doi.org/10.1093/icc/10.4.945

Cooke, P. (2016). Four minutes to four years: The advantage of recombinant over specialized innovation – RIS3 versus 'smartspec'. *European Planning Studies*, *24*(8), 1494–1510. https://doi.org/10.1080/09654313.2016.1151482

Corden, W. M. (1984). Booming sector and Dutch disease economics: Survey and consolidation. *Oxford Economic Papers*, *36*(3), 359–380. http://www.jstor.org/stable/2662669

Corden, W. M., & Neary, J. P. (1982). Booming sector and de-industrialisation in a small open economy. *The Economic Journal*, *92*(368), 825–848. https://doi.org/10.2307/2232670

Cristelli, M., Gabrielli, A., Tacchella, A., Caldarelli, G., & Pietronero, L. (2013). Measuring the intangibles: A metrics for the economic complexity of countries and products. *PLoS ONE*, *8*(8). https://doi.org/10.1371/journal.pone.0070726

Dam, A. V., & Frenken, K. (2022). Variety, complexity and economic development. *Research Policy*, *51*(8), 103949. https://doi.org/10.1016/j.respol.2020.103949

David, P. A. (1985). Clio and the economics of QWERTY. *The American Economic Review*, *75*(2), 332–337.

Davies, B., & Maré, D. C. (2020). Relatedness, complexity and local growth. *Regional Studies*, 1–16. https://doi.org/10.1080/00343404.2020.1802418

Davis, G. A., & Tilton, J. E. (2005). The resource curse. *Natural Resources Forum*, *29*, 233–242. https://doi.org/10.1111/j.1477-8947.2005.00133.x

De Benedictis, L., & Tamberi, M. (2001). *A note on the Balassa Index of Revealed Comparative Advantage*.

Deegan, J., Broekel, T., & Fitjar, R. D. (2021). Searching through the haystack: The relatedness and complexity of priorities in smart specialisation strategies. *Economic Geography*, 1–24. https://doi.org/10.1080/00130095.2021.1967739

Demir, F. (2019). IMF conditionality, trade structure and economic complexity: What did structural adjustment programs really adjust. *University of Oklahoma*.

Dietz, S., Neumayer, E., & De Soysa, I. (2007). Corruption, the resource curse and genuine saving. *Environment and Development Economics*, *12*(1), 33–53. https://doi.org/10.1017/S1355770X06003378

Ding, N., & Field, B. C. (2005). Natural resource abundance and economic growth. *Land Economics*, *81*(4), 496–502.

Döring, T., & Schnellenbach, J. (2006). What do we know about geographical knowledge spillovers and regional growth?: A survey of the literature. *Regional Studies*, *40*(3), 375–395. https://doi.org/10.1080/00343400600632739

Dosi, G., & Nelson, R. R. (1994). An introduction to evolutionary theories in economics. *Journal of Evolutionary Economics*, *4*, 153–172. https://doi.org/10.5860/choice.29-5098

Duranton, G., & Puga, D. (2002). *From sectoral to functional urban specialisation*. NBER Working Paper 9112. http://www.nber.org/papers/w9112

Ederer, P., Nedelkoska, L., Patt, A., & Castellazzi, S. (2015). What do employers pay for employees' complex problem solving skills? *International Journal of Lifelong Education*, *34*(4), 430–447.

Edquist, C. (1997). *Systems of innovation: Technologies, institutions and organizations*. Pinter Publishers.

Ennis, C. A. (2015). Between trend and necessity: Top-down entrepreneurship promotion in Oman and Qatar. *The Muslim World*, *105*(1), 116–138. https://doi.org/10.1111/muwo.12083

Evangelista, R. (2000). Sectoral patterns of technological change in services. *Economics of Innovation and New Technology*, *9*(3), 183–222. https://doi.org/10.1080/10438590000000008

Fagerberg, J. (1988). International competitiveness. *The Economic Journal*, *98*(391), 355–374. https://www.jstor.org/stable/2233372

Fagerberg, J. (1996). Technology and competitiveness. *Oxford Review of Economic Policy*, *12*(3), 39–51. https://www.jstor.org/stable/23606440

Fagerberg, J., & Srholec, M. (2008). National innovation systems, capabilities and economic development. *Research Policy*, *37*, 1417–1435. https://doi.org/10.1016/j.respol.2008.06.003

Fagerberg, J., & Srholec, M. (2017). Capabilities, economic development, sustainability. *Cambridge Journal of Economics*, *41*, 905–926. https://doi.org/10.1093/cje/bew061

Faria, A. P., Barbosa, N., & Bastos, J. (2020). Portuguese regional innovation systems efficiency in the European Union context. *European Planning Studies*, *28*(8), 1599–1618. https://doi.org/10.1080/09654313.2019.1680611

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

Ferraz, D., Moralles, H. F., Campoli, J. S., Oliveira, F. C. R. d., & Rebelatto, D. A. d. N. (2018). Economic complexity and human development: DEA performance measurement in Asia and Latin America. *Gestão & Produção*, *25*(4), 839–853. https://doi.org/10.1590/0104-530x3925-18

Fleming, L., & Sorenson, O. (2001). Technology as a complex adaptive system: Evidence from patent data. *Research Policy*, *30*(7), 1019–1039. https://doi.org/10.1016/S0048-7333(00)00135-9

Fonseca, T., Lima, F., & Pereira, S. C. (2018a). Job polarization, technological change and routinization: Evidence for Portugal. *Labour Economics*, *51*, 317–339. https://doi.org/10.1016/j.labeco.2018.02.003

Fonseca, T., Lima, F., & Pereira, S. C. (2018b). Understanding productivity dynamics: A task taxonomy approach. *Research Policy*, *47*(1), 289–304. https://doi.org/10.1016/j.respol.2017.11.004

Foster, J. (1993). Economics and the self-organisation approach: Alfred Marshall revisited? *The Economic Journal*, *103*(419), 975–991.

Foster, J., & Metcalfe, J. S. (2009). Evolution and economic complexity: An overview. *Economics of Innovation and New Technology*, *18*(7), 607–610. https://doi.org/10.1080/10438590802564477

Frankel, J. A. (2010). *The natural resource curse: A survey*. NBER Working Paper Series. http://www.nber.org/papers/w15836

Frankel, J. A. (2012). *The natural resource curse: A survey of diagnoses and some presctiptions*. John F. Kennedy School of Government, Harvard University. http://nrs.harvard.edu/urn-3:HUL.InstRepos:8694932

Freeman, C. (1987). *Technology policy and economic performance: Lessons from Japan*. Pinter Publishers.

Freeman, C., & Lundvall, B. Å. (1988). *Small countries facing the technological revolution*. Pinter Publishers.

French, S. (2017). Revealed comparative advantage: What is it good for? *Journal of International Economics*, *106*, 83–103. https://doi.org/10.1016/j.jinteco.2017.02.002

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. S. L., & 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

Galego, A., & Pereira, J. (2014). Decomposition of regional wage differences along the wage distribution in Portugal: The importance of covariates. *Environment and Planning A*, *46*(10), 2514–2532. https://doi.org/10.1068/a130055p

Gelb, A. H. (1988). *Oil windfalls: Blessing or curse?* Oxford University Press.

Gertler, M. S. (2003). Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there). *Journal of Economic Geography*, *3*(1), 75–99. https://doi.org/10.1093/jeg/3.1.75

Glaeser, E. L. (1999). Learning in cities. *Journal of Urban Economics*, *46*(2), 254–277. https://doi.org/10.1006/juec.1998.2121

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities. *Journal of political economy*, *100*(6), 1126–1152.

Glewwe, P., Maïga, E., & Zheng, H. (2014). The contribution of education to economic growth: A Review of the evidence, with special attention and an application to

Sub-Saharan Africa. *World Development*, *59*, 379–393. https://doi.org/10.1016/j.worlddev.2014.01.021

Griliches, Z. (1998). Patent statistics as economic indicators: A survey. In *R&D and productivity: The econometric evidence* (pp. 287–343). University of Chicago Press. http://www.nber.org/chapters/c8351

Guevara, M. R., Hartmann, D., Aristarán, M., Mendoza, M., & Hidalgo, C. A. (2016). The research space: Using career paths to predict the evolution of the research output of individuals, institutions, and nations. *Scientometrics*, *109*, 1695–1709. https://doi.org/10.1007/s11192-016-2125-9

Gylfason, T. (2001). Natural resources, education, and economic development. *European Economic Review*, *45*(4-6), 847–859. https://doi.org/10.1016/S0014-2921(01)00127-1

Gylfason, T., & Zoega, G. (2006). Natural resources and economic growth: The role of investment. *The World Economy*, *29*(8), 1091–1115. https://doi.org/10.1111/j.1467-9701.2006.00807.x

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

Hanushek, E. A., & Wossmann, L. (2010). Education and economic growth. In *International Encyclopedia of Education* (pp. 245–252, Vol. 2). Elsevier.

Hanushek, E. A. (2016). Will more higher education improve economic growth? *Oxford Review of Economic Policy*, *32*(4), 538–552. https://doi.org/10.1093/oxrep/grw025

Hanushek, E. A., & Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *The American Economic Review*, *90*(5), 1184–1208. https://www.jstor.org/stable/2677847

Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, *46*(3), 607–668. https://doi.org/10.1257/jel.46.3.607

Hardy, W., Keister, R., & Lewandowski, P. (2018). Educational upgrading, structural change and the task composition of jobs in Europe. *Economics of Transition*, *26*(2), 201–231. https://doi.org/10.1111/ecot.12145

Hartmann, D., Bezerra, M., & Pinheiro, F. L. (2019). *Identifying smart strategies for economic diversification and inclusive growth in developing economies: The case of*

Paraguay. Hohenheim Discussion Papers in Business, Economics and Social Sciences, 04-2019. Stuttgart.

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

Hartmann, D., Jara-Figueroa, C., Guevara, M. R., Simoes, A., & Hidalgo, C. A. (2016). The structural constraints of income inequality in Latin America. *Integration and Trade Journal*, (40), 70–85. http://arxiv.org/abs/1701.03770

Hassink, R. (2005). How to unlock regional economies from path dependency? From learning region to learning cluster. *European Planning Studies*, *13*(4), 521–535. https://doi.org/10.1080/09654310500107134

Hausmann, R., & Chauvin, J. (2015). *Moving to the adjacent possible: Discovering paths for export diversification in Rwanda*. CID Working Paper No. 294.

Hausmann, R., Cunningham, B., Matovu, J., Osire, R., & Wyett, K. (2014a). *How should Uganda grow?* CID Working Paper No. 275.

Hausmann, R., & Hidalgo, C. A. (2010). *Country diversification, product ubiquity, and economic divergence*. CID Working Paper No. 201.

Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., Simoes, A., & Yildirim, M. A. (2014b). *The atlas of economic complexity: Mapping paths to prosperity*. MIT Press. https://growthlab.cid.harvard.edu/files/growthlab/files/atlas_2013_part1.pdf

Hausmann, R., & Klinger, B. (2007). *The structure of the product space and the evolution of comparative advantage*. CID Working Paper No. 146.

Hausmann, R., & Rodrik, D. (2003). Economic development as self-discovery. *Journal of Development Economics*, *72*(2), 603–633. https://doi.org/10.1016/S0304-3878(03)00124-X

Hausmann, R., Santos, M. A., & Obach, J. (2017). *Appraising the economic potential of Panama: Policy recommendations for sustainable and inclusive growth*. CID Working Paper No. 334.

Hayek, F. A. (1937). Economics and knowledge. *Economica*, (13), 33–54.

Hayek, F. A. (1945). The use of knowledge in society. *The American Economic Review*, *35*, 519–530.

Heimeriks, G., Li, D., Lamers, W., Meijer, I., & Yegros, A. (2019). Scientific knowledge production in European regions: Patterns of growth, diversity and complexity. *European Planning Studies*, *27*(11), 2123–2143. https://doi.org/10.1080/09654313.2019.1645814

Henderson, J., Dicken, P., Hess, M., Coe, N. M., & Yeung, H. W. C. (2002). Global Production Networks and the analysis of economic development. *Review of International Political Economy*, *9*(3), 436–464. https://doi.org/10.1080/09692290210150842

Henderson, V. (1999). *Marshall's economies*. NBER Working Paper 7358. http://www.nber.org/papers/w7358

Henderson, V. (2010). Cities and development. *Journal of Regional Science*, *50*(1), 515–540. https://doi.org/10.1111/j.1467-9787.2009.00636.x

Hidalgo, C. A. (2021). Economic complexity theory and applications. *Nature Reviews Physics*. 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., Balland, P. A., Boschma, R., Delgado, M., Feldman, M. P., Frenken, K., Glaeser, E. L., He, C., Kogler, D. F., Morrison, A., Neffke, F., Rigby, D., Stern, S., Zheng, S., & Zhu, S. (2018). The principle of relatedness. *Springer Nature Switzerland AG*, 451–457. https://doi.org/10.1007/978-3-319-96661-8

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

Hidalgo, C. A., Klinger, B., Barabási, A., & Hausmann, R. (2007). The product space conditions the development of nations. *Science*, *317*(5837), 482–487.

Howells, J. (1999). Regional systems of innovation? In D. Archibugi, J. Howells & J. Michie (Eds.), *Innovation Policy in a Global Economy*. Cambridge University Press.

Humphreys, M., Sachs, J. D., Stiglitz, J. E., Humphreys, M., & Soros, G. (2007). *Escaping the resource curse*. Columbia University Press.

Hvidt, M. (2011). Economic and Institutional Reforms in the Arab Gulf Countries. *Middle East Journal*, *65*(1), 85–102. Retrieved August 9, 2022, from http://www.jstor.org/stable/23012095

Hvidt, M. (2013). *Economic diversification in GCC countries: Past record and future trends*. Kuwait Programme on Development, Governance and Globalisation in the Gulf States.

Iammarino, S. (2005). An evolutionary integrated view of regional systems of innovation: Concepts, measures and historical perspectives. *European Planning Studies*, *13*(4), 497–519. https://doi.org/10.1080/09654310500107084

Iammarino, S., Rodríguez-Pose, A., & Storper, M. (2019). Regional inequality in Europe: Evidence, theory and policy implications. *Journal of Economic Geography*, *19*(2), 273–298. https://doi.org/10.1093/jeg/lby021

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidencced by patent citations. *The Quarterly Journal of Economics*, *108*(3), 577–598. https://doi.org/10.1111/j.1467-8276.2008.01144.x

Jouini, J., & Harrathi, N. (2014). Revisiting the shock and volatility transmissions among GCC stock and oil markets: A further investigation. *Economic Modelling*, *38*, 486–494. https://doi.org/10.1016/j.econmod.2014.02.001

Kalaitzi, A. S., & Chamberlain, T. W. (2019). Further evidence on export-led growth in the United Arab Emirates: Are non-oil exports or re-exports the key to economic growth? *Review of Middle East Economics and Finance*, *15*(2), 1–16. https://doi.org/10.1515/rmeef-2019-0007

Kalaitzi, A. S., & Chamberlain, T. W. (2020). Merchandise exports and economic growth: Multivariate time series analysis for the United Arab Emirates. *Journal of Applied Economics*, *23*(1), 163–182. https://doi.org/10.1080/15140326.2020.1722384

Kalaitzi, A. S., & Chamberlain, T. W. (2021). The validity of the export-led growth hypothesis: Some evidence from the GCC. *Journal of International Trade and Economic Development*, *30*(2), 224–245. https://doi.org/10.1080/09638199.2020.1813191

Kalaitzi, A. S., & Cleeve, E. (2018). Export-led growth in the UAE: Multivariate causality between primary exports, manufactured exports and economic growth. *Eurasian Business Review*, *8*(3), 341–365. https://doi.org/10.1007/s40821-017-0089-1

Kemeny, T., & Storper, M. (2015). Is specialization good for regional economic development? *Regional Studies*, *49*(6), 1003–1018. https://doi.org/10.1080/00343404.2014.899691

Kemp-Benedict, E. (2014). *An interpretation and critique of the method of reflections*. MPRA Paper No. 60705. https://mpra.ub.uni-muenchen.de/60705/

Kim, L. (1997). *Imitation to innovation: The dynamics of Korea's technological learning*. Harvard Business School Press.

Koch, P. (2021). Economic complexity and growth: Can value-added exports better explain the link? *Economics Letters*, *198*, 109682. https://doi.org/10.1016/j.econlet.2020.109682

Kogler, D. F. (2017). Relatedness as driver of regional diversification: A research agenda - A commentary. *Regional Studies*, *51*(3), 365–369. https://doi.org/10.1080/00343404.2016.1276282

Kogler, D. F., Rigby, D., & 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

Kogut, B., & Zander, U. (1992). Knowledge of the firm, combinative capabilities and the replication of technology. *Organization Science*, *3*(3), 383–397. https://doi.org/10.1287/orsc.3.3.383

Koren, M., & Tenreyro, S. (2010). *Volatility, diversification and development in the Gulf Cooperation Council countries*. Research Paper, Kuwait Program on Development, Governance, and Globalization in the Gulf States, London School of Economics.

Kumar, K. B., & van Welsum, D. (2013). *Knowledge-based economies and basing economies on knowledge: Skills a missing link in GCC countries*. Rand Corporation.

Lall, S. (1992). Technological capabilities and industrialization. *World Development*, *20*(2), 165–186.

Lall, S., Weiss, J., & Zhang, J. (2006). The "sophistication" of exports: A new trade measure. *World Development*, *34*(2), 222–237. https://doi.org/10.1016/j.worlddev.2005.09.002

Lambooy, J. G., & Boschma, R. (2001). Evolutionary economics and regional policy. *Annals of Regional Science*, *35*(1), 113–131. https://doi.org/10.1007/s001680000033

Lapatinas, A. (2019). The effect of the Internet on economic sophistication: An empirical analysis. *Economics Letters*, *174*, 35–38. https://doi.org/10.1016/j.econlet.2018.10.013

Lapatinas, A., Kyriakou, A., & Garas, A. (2019). Taxation and economic sophistication: Evidence from OECD countries. *PLoS ONE*, *14*(3), 1–21. https://doi.org/10.1371/journal.pone.0213498

Laranja, M., Edwards, J., Pinto, H., & Foray, D. (2020). *Implementation of Smart Specialisation Strategies in Portugal: An assessment*. European Commission Joint Research Centre Technical Report.

Lawson, C., & Lorenz, E. (1999). Collective learning, tacit knowledge and regional innovative capacity. *Regional Studies*, *33*(4), 305–317. https://doi.org/10.1080/713693555

Leamer, E. E. (1985). *Sources of international comparative advantage: Theory and evidence*. MIT Press.

Lee, K. K., & Vu, T. V. (2020). Economic complexity, human capital and income inequality: A cross-country analysis. *Japanese Economic Review*, *71*(4), 695–718. https://doi.org/10.1007/s42973-019-00026-7

Levine, R., & Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. *The American Economic Review*, *82*(4), 942–963. https://www.jstor.org/stable/2117352

Lo Turco, A., & Maggioni, D. (2020). The knowledge and skill content of production complexity. *Research Policy*, 104059. https://doi.org/10.1016/j.respol.2020.104059

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, *22*(1), 3–42. https://doi.org/10.1016/0304-3932(88)90168-7

Lundberg, L. (1988). Technology, factor proportions and competitiveness. *The Scandinavian Journal of Economics*, *90*(2), 173–188. https://www.jstor.org/stable/3440097

Lundvall, B. Å. (1992). *National systems of innovation: Towards a theory of innovation and interactive learning*. Pinter Publishers.

Malerba, F., & Orsenigo, L. (1997). Technological regimes and sectoral patterns of innovative activities. *Industrial and Corporate Change*, *6*(1), 83–118. https://doi.org/10.1093/icc/6.1.83

Malerba, F., & Orsenigo, L. (1993). Technological regimes and firm bebavior. *Industrial and corporate change*, *2*(1), 45–71.

Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, *107*(2), 407–437. https://doi.org/10.2307/2118477

Marrocu, E., Paci, R., Usai, S., & Rigby, D. (2020). *Smart Specialization Strategy: Any relatedness between theory and practice?* CRENoS Working Paper 2020/04. https://crenos.unica.it/crenos/publications/smart-specialization-strategy-any-relatedness-between-theory-and-practice

Marshall, A. (1890). *Principles of economics*. Macmillan & Co.

Martin, R., & Sunley, P. (2006). Path dependence and regional economic evolution. *Journal of Economic Geography*, *6*(4), 395–437. https://doi.org/10.1093/jeg/lbl012

Maskell, P., & Malmberg, A. (1999). Localised learning and industrial competitiveness. *Cambridge Journal of Economics*, *23*, 167–185. https://doi.org/10.1093/cje/23.2.167

McCann, P., & Ortega-Argilés, R. (2015). Smart specialization, regional growth and applications to European Union cohesion policy. *Regional Studies*, *49*(8), 1291–1302. https://doi.org/10.1080/00343404.2013.799769

Mealy, P., & Coyle, D. (2019). *Economic complexity analysis*. Technical report for the research on Innovation & Global Competitiveness, Bennett Institute for Public Policy, University of Cambridge.

Mealy, P., & Coyle, D. (2022). To them that hath: Economic complexity and local industrial strategy in the UK. *International Tax and Public Finance*, *29*(2), 358–377. https://doi.org/10.1007/s10797-021-09667-0

Mealy, P., Farmer, J. D., & Hausmann, R. (2018a). Determining the differences that matter: Development and divergence in US states over 1850-2010. *SSRN Electronic Journal*, 1–44. https://doi.org/10.2139/ssrn.3235193

Mealy, P., Farmer, J. D., & Teytelboym, A. (2018b). *A new interpretation of the economic complexity index*. INET Oxford Working Paper 2018-04.

Mealy, P., Farmer, J. D., & Teytelboym, A. (2019). Interpreting economic complexity. *Science Advances*, *5*, 1–8. https://doi.org/10.1126/sciadv.aau1705

Mealy, P., & Teytelboym, A. (2020). Economic complexity and the green economy. *Research Policy*, (September 2019), 103948. https://doi.org/10.1016/j.respol.2020.103948

Mehlum, H., Moene, K., & Torvik, R. (2006). Institutions and the resource curse. *The Economic Journal*, *116*, 1–20. https://doi.org/10.1111/j.1468-0297.2006.01045.x

Mehrara, M. (2009). Reconsidering the resource curse in oil-exporting countries. *Energy Policy*, *37*(3), 1165–1169. https://doi.org/10.1016/j.enpol.2008.10.040

Metcalfe, J. S. (1994). Evolutionary economics and technology policy. *The Economic Journal*, *104*(425), 931–944. http://www.jstor.com/stable/2234988

Mewes, L., & Broekel, T. (2022). Technological complexity and economic growth of regions. *Research Policy*, *51*(8), 104156. https://doi.org/10.1016/j.respol.2020.104156

Mishra, S., Tewari, I., & Toosi, S. (2020). Economic complexity and the globalization of services. *Structural Change and Economic Dynamics*, *53*, 267–280. https://doi.org/10.1016/j.strueco.2020.03.002

Morgan, K. (2007). The learning region: Institutions, innovation and regional renewal. *Regional Studies*, *31*(5), 491–503. https://doi.org/10.1080/00343400701232322

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), 1–11. https://doi.org/10.1038/s41598-017-14603-6

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). https://doi.org/10.3390/su11020497

Neary, J. P., & van Wijnbergen, S. (1986). *Natural resources and the macroeconomy*. MIT Press.

Neffke, F., & Henning, M. (2013). Skill relatedness and firm diversification. *Strategic Management Journal*, *34*, 297–316. https://doi.org/10.1002/smj.2014

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

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

Nelson, R. R., & Phelps, E. S. (1966). Investment in Humans, Technological Diffusion, and Economic Growth. *The American Economic Review*, *56*(1/2), 69–75. Retrieved July 17, 2023, from https://www.jstor.org/stable/1821269

Nelson, R. R., & Winter, S. G. (1973). Towards an evolutionary theory of economic capabilities. *The American Economic Review*, *63*(2), 440–449. https://www.jstor.org/stable/1817107

Nelson, R. R., & Winter, S. G. (1982). *An evolutionary theory of economic change*. Belknap Press of Harvard University Press.

North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge University Press.

North, D. C. (1991). Institutions. *Journal of Economic Perspectives*, *5*(1), 97–112. https://doi.org/10.1257/jep.5.1.97

O'Clery, N. (2015). A tale of two clusters: The evolution of Ireland's economic complexity since 1995. *Journal of the Statistical and Social Inquiry Society of Ireland*, *45*(2009), 16–66.

Ourens, G. (2013). *Can the Method of Reflections help predict future growth?* Université Catholique de Louvain, Institut de Recherches Economiques et Sociales (IRES), Discussion Paper No. 2013008. https://dial.uclouvain.be/pr/boreal/en/object/boreal%3A127055

Parrilli, M. D., Nadvi, K., & Yeung, H. W. C. (2013). Local and regional development in Global Value Chains, Production Networks and Innovation Networks: A comparative review and the challenges for future research. *European Planning Studies*, *21*(7), 967–988. https://doi.org/10.1080/09654313.2013.733849

Patel, P., & Pavitt, K. (1987). The elements of British technological competitiveness. *National Institute Economic Review*, *122*(1), 72–83. https://doi.org/10.1177/002795018712200107

Patelli, A., Pietronero, L., & Zaccaria, A. (2022). Integrated database for economic complexity. *Scientific Data*, *9*(1), 628. https://doi.org/10.1038/s41597-022-01732-5

Pavitt, K. (1984). Sectoral patterns of technical change: Towards a taxonomy and a theory. *Research Policy*, *13*, 343–373. https://doi.org/10.1016/0048-7333(84)90018-0

Pereira, J., & Galego, A. (2011). Regional wage differentials in Portugal: Static and dynamic approaches. *Papers in Regional Science*, *90*(3), 529–548. https://doi.org/10.1111/j.1435-5957.2010.00328.x

Pereira, J., & Galego, A. (2015). Intra-regional wage inequality in Portugal. *Spatial Economic Analysis*, *10*(1), 79–101. https://doi.org/10.1080/17421772.2014.992360

Pilat, D., & Wölfl, A. (2005). *Measuring the interaction between manufacturing and services*. OECD STI Working Paper 2005/5.

Pinheiro, F. L., Balland, P.-A., Boschma, R., & Hartmann, D. (2022). The dark side of the geography of innovation: Relatedness, complexity and regional inequality in Europe. *Regional Studies*, 1–16. https://doi.org/10.1080/00343404.2022.2106362

Pintar, N., & Scherngell, T. (2022). The complex nature of regional knowledge production: Evidence on European regions. *Research Policy*, *51*(8), 104170. https://doi.org/10.1016/j.respol.2020.104170

Poncet, S., & Starosta de Waldemar, F. (2013). Export upgrading and growth: The prerequisite of domestic embeddedness. *World Development*, *51*, 104–118. https://doi.org/10.1016/j.worlddev.2013.05.010

Rego, C., Ramos, I. J., Lucas, M. R., Baltazar, M. d. S., & Dionísio, A. (2021). New geography in old territories: A multivariate approach based on Portuguese regions. *Regional Studies, Regional Science*, *8*(1), 25–50. https://doi.org/10.1080/21681376.2020.1860807

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

Rigby, D., & Essletzbichler, J. (1997). Evolution, process variety, and regional trajectories of technological change in U.S. manufacturing. *Economic Geography*, *73*(3), 269–284. https://doi.org/10.2307/144484

Rigby, D., Roesler, C., Kogler, D. F., Boschma, R., & Balland, P. A. (2019). *Do EU regions benefit from smart specialization?* Papers in Evolutionary Economic Geography 19.31.

Rodríguez-Pose, A. (2013). Do institutions matter for regional development? *Regional Studies*, *47*(7), 1034–1047. https://doi.org/10.1080/00343404.2012.748978

Rodrik, D., Subramanian, A., & Trebbi, F. (2004). Institutions rule: The primacy of institutions over geography and integration in economic development. *Journal of*

Economic Growth, *9*(2), 131–165. https://doi.org/10.1023/B:JOEG.0000031425.72248.85

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, *98*(5), S71–S102.

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

Rosenberg, N., & Nelson, R. R. (1994). American universities and technical advance in industry. *Research policy*, *23*(3), 323–348.

Rosser, A. (2009). Natural resource wealth, development and social policy: Evidence and issues. In K. Hujo & S. McClanahan (Eds.), *Financing Social Policy* (pp. 165–182). Palgrave Macmillan UK. https://doi.org/10.1057/9780230244337_7

Rostow, W. (1961). *The stages of economic growth: A non-communist manifesto*. Cambridge University Press.

Sachs, J. D. (2007). How to handle the macroeconomics of oil wealth. In *Escaping the Resource Curse*. Columbia University Press.

Sachs, J. D., & Warner, A. M. (1995). *Natural resource abundance and economic growth*. National Bureau of Economic Research.

Sachs, J. D., & Warner, A. M. (1999). The big push, natural resource booms and growth. *Journal of Development Economics*, *59*(1), 43–76. https://doi.org/10.1016/S0304-3878(99)00005-X

Sachs, J. D., & Warner, A. M. (2001). The curse of natural resources. *European Economic Review*, *45*, 827–838.

Sala-i-Martin, X., & Subramanian, A. (2013). Addressing the natural resource curse: An illustration from Nigeria. *Journal of African Economies*, *22*(4), 570–615. https://doi.org/10.1093/jae/ejs033

Samans, R., & Zahidi, S. (2017). *The future of jobs and skills in the Middle East and North Africa: Preparing the region for the Fourth Industrial Revolution*. World Economic Forum.

Santos, D. (2000). Innovation and territory: Which strategies promote regional innovation systems in Portugal? *European Urban and Regional Studies*, *7*(2), 147–157. https://doi.org/10.1177/096977640000700204

Santos, D., & Simões, M. J. (2014). Regional innovation systems in Portugal: A critical evaluation. *Investigaciones Regionales*, *2014*(28), 37–56.

Schultz, T. W. (1961). Investment in human capital. *The American Economic Review*, *51*(1), 1–17. https://www.jstor.org/stable/1818907

Schumpeter, J. A. (1942). Creative destruction. In *Capitalism, socialism, and democracy* (Vol. 825). Taylor & Francis Group.

Schumpeter, J. A. (1947). The creative response in economic history. *Journal of Economic History*, *7*, 149–159. https://doi.org/https://doi.org/10.1017/S0022050700054279

Sciarra, C., Chiarotti, G., Ridolfi, L., & Laio, F. (2020). Reconciling contrasting views on economic complexity. *Nature Communications*, *11*(3352), 1–10. https://doi.org/10.1038/s41467-020-16992-1

Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*. Penguin Books.

Solow, R. M. (1957). Technical change and the aggregate production function. *The Review of Economics and Statistics*, *39*(3), 312–320.

Stijns, J.-P. (2006). Natural resource abundance and human capital accumulation. *World Development*, *34*(6), 1060–1083. https://doi.org/10.1016/j.worlddev.2005.11.005

Stojkoski, V., Koch, P., & Hidalgo, C. A. (2023). Multidimensional economic complexity and inclusive green growth. *Communications Earth & Environment*, *4*(130), 1–12. 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), 1–29. https://doi.org/10.1371/journal.pone.0161633

Storper, M. (1993). Regional 'worlds' of production: Learning and innovation in the technology districts of France, Italy and the USA. *Regional Studies*, *27*(5), 433–455. https://doi.org/10.1080/00343409312331347675

Storper, M. (1997). *The regional world: Territorial development in a global economy*. Guilford press.

Sutton, J., & Trefler, D. (2016). Capabilities, wealth and trade. *Journal of Political Economy*, *124*(3), 826–878.

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

Tabash, M. I., Mesagan, E. P., & Farooq, U. (2022). Dynamic linkage between natural resources, economic complexity, and economic growth: Empirical evidence from Africa. *Resources Policy*, *78*, 102865. https://doi.org/10.1016/j.resourpol.2022.102865

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

Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., & Pietronero, L. (2012). A new metrics for countries' fitness and products' complexity. *Scientific Reports*, *2*(723). https://doi.org/10.1038/srep00723

Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., & Pietronero, L. (2013). Economic complexity: Conceptual grounding of a new metrics for global competitiveness. *Journal of Economic Dynamics and Control*, *37*(8), 1683–1691. https://doi.org/10.1016/j.jedc.2013.04.006

Timmer, M. P., Miroudot, S., & de Vries, G. (2019). Functional specialisation in trade. *Journal of Economic Geography*, *19*(1), 1–30. https://doi.org/10.1093/jeg/lby056

Valenta, M., & Jakobsen, J. (2016). Moving to the Gulf: An empirical analysis of the patterns and drivers of migration to the GCC countries, 1960–2013. *Labor History*, *57*(5), 627–648. https://doi.org/10.1080/0023656X.2016.1239885

Van der Ploeg, F. (2011). Natural resources: Curse or blessing? *Journal of Economic Literature*, *49*(2), 366–420. http://www.jstor.org/stable/23071620

Van der Ploeg, F., & Poelhekke, S. (2009). Volatility and the natural resource curse. *Oxford Economic Papers*, *61*(4), 727–760. https://doi.org/10.1093/oep/gpp027

Vandenbussche, J., Aghion, P., & Meghir, C. (2006). Growth, distance to frontier and composition of human capital. *Journal of Economic Growth*, *11*(2), 97–127. https://doi.org/10.1007/s10887-006-9002-y

Venables, A. J. (2016). Using natural resources for development: Why has it proven so difficult? *Journal of Economic Perspectives*, *30*(1), 161–184. https://doi.org/10.1257/jep.30.1.161

von Tunzelmann, N., & Wang, Q. (2007). Capabilities and production theory. *Structural Change and Economic Dynamics*, *18*(2), 192–211. https://doi.org/10.1016/j.strueco.2006.11.002

Vu, T. V. (2019). *Does institutional quality foster economic complexity?* University of Otago Economics Discussion Paper No. 1909.

Vu, T. V. (2020). Economic complexity and health outcomes: A global perspective. *Social Science & Medicine*. https://doi.org/10.1016/j.socscimed.2020.113480

Weaver, W. (1948). Science and complexity. *American Scientist*, *36*(4), 536–544. https://www.jstor.org/stable/27826254

Weber, A. (1929). *Theory of the location of industries*. Chicago University Press.

Weitzman, M. L. (1998). Recombinant growth. *The Quarterly Journal of Economics*, *113*(2), 331–360. https://doi.org/10.21648/arthavij/2013/v55/i3/111235

Whittle, A. (2019). Local and non-local knowledge typologies: Technological complexity in the Irish knowledge space. *European Planning Studies*, *27*(4), 661–677. https://doi.org/10.1080/09654313.2019.1567695

Winter, S. G. (1984). Schumpeterian competition in alternative technological regimes. *Journal of Economic Behavior & Organization*, *5*(3-4), 287–320. https://doi.org/10.1016/0167-2681(84)90004-0

World Bank. (2022). World Bank country and lending groups. Retrieved August 11, 2022, from https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

Yalta, A. Y., & Yalta, T. (2021). Determinants of economic complexity in MENA countries. *Journal of Emerging Economies and Policy*, *6*(1), 5–16.

Yeung, H. W.-c. (2021). Regional worlds: From related variety in regional diversification to strategic coupling in global production networks. *Regional Studies*, *55*(6), 989–1010. https://doi.org/10.1080/00343404.2020.1857719

Zaccaria, A., Cristelli, M., Kupers, R., Tacchella, A., & Pietronero, L. (2016). A case study for a new metrics for economic complexity: The Netherlands. *Journal of Economic Interaction and Coordination*, *11*(1), 151–169. https://doi.org/10.1007/s11403-015-0145-9

Zhu, S., Hey, C., & Zhou, Y. (2017). How to jump further and catch up? Path-breaking in an uneven industry space. *Journal of Economic Geography*, *17*(3), 521–545. https://doi.org/10.1093/jeg/lbw047

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

Zhu, S., & Li, R. (2017). Economic complexity, human capital and economic growth: Empirical research based on cross-country panel data. *Applied Economics*, *49*(38), 3815–3828. https://doi.org/10.1080/00036846.2016.1270413