

**Towards a Post-Oil Economy:
A Complexity Approach to Understanding
Natural Resource Dependency and
Economic Diversification in Ecuador**



Pablo A. Astudillo-Estevez

St Edmund Hall

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2021

This thesis is dedicated to my parents and brother:
Pablo, Marcela and Ivan.

Acknowledgements

Undertaking this DPhil programme has been a truly life-changing experience for me and it would not have been possible without the support and guidance that I received from many people and organisations.

I would like to first say a very big thank you to my supervisor Prof. Gordon L. Clark for all the support and encouragement he provided throughout this programme. Without his guidance and feedback, this DPhil would not have been completed. I would like to further extend my sincere gratitude to the University of Oxford for providing all the support and the environment needed to push my intellectual boundaries during my doctoral studies.

Many thanks also to Prof. Cesar Hidalgo and the former Collective Learning Group at the Massachusetts Institute of Technology for their support and friendship. They helped me to explore new ways to “let the data talk” and to embrace innovative technological tools for data visualisation and analyses. Also, many thanks to Dr. Frank Neffke and the Growth Lab at Harvard University, who were key to help me expand my knowledge about development, politics and economics.

I gratefully acknowledge the funding received for my DPhil from the Secretaría Nacional de Educación Superior, Ciencia, Tecnología e Innovación from the Republic of Ecuador. In addition, I would like to thank the Central Bank of Ecuador and the Center for Fiscal Studies for providing the data to conduct this research.

I greatly appreciate the academic and personal support received from Xi Hu, whose critical eye, constant feedback and emotional support helped me, not only improving my research but my life too. My time at Oxford was made enjoyable in large part due to the many friends that became an important part of my life. I am particularly grateful for time spent with my really good friends: Gonzalo Griebenow, Homero Paltán, Raghav Pant, Gustavo Quino, Filipa Soares, Yayoi Teramoto, Chico Camargo, Pablo Escudero and Felipe Ledesma. Definitely, our conversations shaped the direction of this thesis to some extent, and they

vastly improved my mental well-being. Many thanks to my friends from the Oxford Latin Combo, perhaps one of the most important group of people whose company helped me to relax. Playing music with them was key for preserving my mental sanity during this arduous and challenging process. Finally, many thanks to my best friend Esteban that accompanied at a distance through this journey.

Lastly, I would like to thank my family for all their love and encouragement, specially my parents who raised me with a love for science and supported me in all my pursuits.

Abstract

Oil and mineral export-dependency has been a long-standing issue for several countries, especially developing ones. Resource-rich countries not only are highly vulnerable to macroeconomic shocks, due to the volatility of international commodity prices, but also tend to under-perform economically when compared to resource-poor ones. To overcome the problems associated with the overreliance on natural resources, it has been argued that countries must diversify their productive base and significantly reduce their dependency on primary sectors. Notwithstanding the importance of this topic, surprisingly little is known about how this can be achieved.

To improve our understanding about this structural dependency, this thesis argues that one of the fundamental constraints for diversification in oil exporting economies is linked with the ways in which their productive structures are configured. Understanding these productive structures could help explain the limited capacity of such economies to generate new productive activities, based on the productive knowledge they generate and accumulate, and provide insights to guide policy making and investment decisions for diversification pathways.

This thesis proposes an experimental network-based methodology to understand how an oil exporting economy is configured and uses Ecuador as a study case. This methodology enables assessing the importance of different industries in a country, as measured by how connected they are with the rest of the economy and their relative importance for promoting growth. In addition, this thesis tests the impact that the 1972 oil boom had on the productive structure of Ecuador, as a way to assess the importance of the oil industry in terms of its capacity to stimulate other economic sectors. Finally, this thesis examines how the current concentration of different economic activities (and respective production networks) in specific locations influences the emergence of pioneering activities in those places, serving as key components to future productive diversification paths for Ecuador.

In this research, it was observed that the industries that generate more income for Ecuador (i.e., extraction or direct use of natural resources such as agriculture, mining or oil) are the ones that have the lowest centralities. This means that the country relies mostly on industries which are relatively isolated from the rest of the economy and are not big consumers of intermediate inputs, with the lowest capacity to generate knowledge spillovers, thus potentially acting as one of the main constraints for growth in Ecuador. In line with previous findings, it was also observed that there were changes on different industrial centralities after the 1972 oil boom. However, the methodology proposed was not useful to explain if the changes observed were directly attributed to the discovery of crude oil. Finally, this study also shows that the existing productive structure of a country, the location and concentration of productivity and how production networks are linked are good predictors of the emergence of pioneering activities in a given place, which is coherent with the existing literature.

A general conclusion of this thesis is that the productive structure of a country is crucial for its economic development and could limit its diversification potential. In the case of the Ecuadorian economy, it relies mostly on industries with the lowest degree of connectivity, which do not have the capacity to stimulate other sectors effectively. Furthermore, we observed that the productive configuration of an economy could limit the productive knowledge that it can acquire or develop. Its future development paths will depend on this configuration. If oil-exporting economies seek to diversify their productivity, they first need to look at how their economies are configured in-depth. Only then they are in a position to develop strategies towards their aspired development and goals. Diversification processes are heavily path- and place- dependent.

Contents

1	Introduction	1
1.1	Context	1
1.2	Research Objectives and Goals	3
1.3	Ecuador's oil dependency as a study case	6
1.4	Research Significance	8
1.5	Thesis Outline	9
2	Literature Review	13
2.1	Development and Economic Growth: an overview	14
2.2	Productive Structures and Growth	18
2.2.1	Structuralism, import substitution industrialisation and economic growth	18
2.2.2	New structuralism and economic growth	20
2.2.3	Structuralism and economic complexity	22
2.3	Knowledge, Technology and Economic Growth	24
2.4	Spatial characteristics and dynamics of knowledge and innovation	27
2.5	Literature Summary	30
3	Assessing the Productive Structure of an Oil Exporting Country	32
3.1	Introduction	32
3.1.1	Objectives	33
3.2	Literature Review	35
3.2.1	Economic structures, economic growth and development	35
3.2.2	Understanding productive structures	36
3.2.3	Key economic sectors and growth	38
3.2.4	Identifying key sectors and products	39
3.3	Methodology and Data	40
3.3.1	Data	40
3.3.2	Methodology	43

3.4	Results and Discussion	46
3.4.1	Key Products by Centrality	46
3.4.2	Key Sectors by Centrality	47
3.4.3	Comparisons between aggregated and network measurements	51
3.5	Conclusions	52
3.5.1	Methodological conclusions	53
3.5.2	Research conclusions	54
4	Structural Productive Impact of a Resource Boom in Ecuador	56
4.1	Introduction	56
4.2	The Resource Curse and Dutch Disease	57
4.3	Evolution of the productive structure of Ecuador	59
4.3.1	Colonial times to the 20th century	59
4.3.2	Oil boom and Import Substitution Industrialisation (1960 - 1980) . .	60
4.3.3	Structural Adjustment Period and Post-Neoliberal reforms (1980-2010)	62
4.3.4	Current situation	63
4.4	Methodology and Data	65
4.4.1	Synthetic Control Method	66
4.4.2	Application of the method to the Ecuadorian case	67
4.4.3	Method cautions	70
4.5	Results and Discussion	71
4.5.1	Evolution of the Economic Structure of Ecuador	71
4.5.2	Oil boom effects in the Ecuadorian productive structure	76
4.6	Conclusions	92
4.6.1	Methodological conclusions	93
4.6.2	Research conclusions	94
5	Regional Economic Diversification and Industrial Relatedness	96
5.1	Introduction	96
5.2	Literature Review	98
5.2.1	Understanding Economic Diversification in Oil Exporting Economies	98
5.2.2	Regional Productive Diversification and Industrial Relatedness . . .	101
5.2.3	Literature Gaps	103
5.3	Data and Methodology	104
5.3.1	Data	104
5.3.2	Methodology	111
5.3.3	Empirical Model	117

5.4	Findings: Relatedness and Pioneering Industries' Emergence	121
5.4.1	Crude Oil and Mining Industry Emergence vs Manufacturing	124
5.5	Discussion	126
5.6	Conclusions	129
6	Conclusions and Implications	131
6.1	Thesis Overview	131
6.2	Guiding questions and relevance to the literature	132
6.3	Main findings and conclusions	134
6.4	Implication for future research	137
6.5	Implication for public policy	138
6.6	Limitations of this research	140
	Appendix A	143
	Bipartite Network Projections	143
	Appendix B	149
	Comparison Between Models: Logistic, Probit and Linear Probability (OLS) . .	149
	Appendix C	151
	Comparison of Regression Models: Agriculture, Manufacture and Commerce . .	151
	Appendix D	153
	Geographical Agglomeration and Firm Output	153
	Bibliography	155

List of Figures

3.1	Use table as a bipartite network	44
3.2	Products importance, based on degree centrality	47
3.3	Products importance, based on weighted centrality	48
3.4	Products importance, based on eigenvector centrality	48
3.5	Sector relevance, based on degree centrality	49
3.6	Sector importance, based on weighted centrality	50
3.7	Sector importance, based on eigenvector centrality	50
3.8	Comparison between standard (total demand) and network (degree central- ity) measurements	51
3.9	Comparison between network (degree centrality) and standard (total de- mand) measurements	52
4.1	OPEC Oil Production in selected countries (% of OPEC total production) .	61
4.2	Ecuador' Exports from 1972 to 1982	62
4.3	Ecuadorian exports share 1968 - 2014	63
4.4	Ecuadorian exports 1968 - 2014	65
4.5	Evolution of the Ecuadorian Economy (1965-1985)	72
4.6	Evolution of the Ecuadorian Economy (1965-1985)	73
4.7	Evolution of the Ecuadorian Economy (1965-1985)	74
4.8	Evolution of the Ecuadorian Economy (1965-1985)	75
4.9	Evolution of the Ecuadorian Economy (1965-1985)	77
4.10	Banana, coffee and cocoa beans	78
4.11	Other agricultural products	78
4.12	Animal products	79
4.13	Forestry products	79
4.14	Hunting and fishing products	80
4.15	Oil and natural gas	80
4.16	Oil refining products	81

4.17 Other mining products	81
4.18 Meat and processed fish	82
4.19 Cereals and bakery	82
4.20 Sugar	83
4.21 Various foodstuffs	83
4.22 Drinks	84
4.23 Textiles, clothing and leather products	85
4.24 Wood	85
4.25 Paper and printing	86
4.26 Chemicals, plastics and rubber products	86
4.27 Basic mineral products	87
4.28 Machinery, transport equipment	87
4.29 Other manufactured products	88
4.30 Electricity, gas and water	89
4.31 Construction	89
4.32 Commerce	90
4.33 Transport	90
4.34 Communication	91
4.35 Financial services	91
4.36 Rental housing	92
4.37 Business services	92
4.38 Hotels, bars and restaurants	93
5.1 Crude oil wealth flow in Ecuador (2015)	100
5.2 Total Output by Industry in Ecuador, by ISIC (2015)	106
5.3 Total Number of Entities by Industry (2015)	107
5.4 Economic Activities by Province (2015)	108
5.5 Output Shared Generated by Province (2015)	109
5.6 Economic Activities by Industry per Province (2015)	110
5.7 Entities' Output Distribution (2015)	111
5.8 Number of Industry-Location Per Year	117
5.9 Total class subscription	118
5.10 Class subscription by year	118
5.11 Confusion Matrix	123
5.12 Area Under the Curve	124
5.13 Confusion Matrix	126

5.14	Area Under the Curve	126
1	Bipartite to Unipartite Projection	143
2	Weighed Projection	146
3	Collaboration Projection	147
4	Jaccard Projection	147
5	Products relevance, based on weighted centrality	148

Chapter 1

Introduction

1.1 Context

The last decade marked the end of a commodity ‘supercycle’¹ with detrimental impacts on economies that rely mostly on the oil industry (Canuto, 2014; Brenes et al., 2016; Gruss, 2014; Helm, 2016). For instance, Venezuela, Saudi Arabia and Russia experienced severe economic contractions due to plummeting oil prices between 2014 and 2016. According to some estimations, Venezuela, whose economy is 95 per cent reliant on crude oil exports, lost around 60% of its income in that period (BBC, 2016). Similarly, oil income in Saudi Arabia and Russia fell by 23% and 50%, respectively (BBC, 2016). In fact, Gelb (2011) explains that 35 countries² are constantly exposed to the risks of falling oil prices due to their dependency on oil and to new producers coming on stream.

Oil and mineral export-dependency has been a long-standing issue for several countries, especially developing ones. Research has shown that these countries are highly vulnerable to macroeconomic shocks, which are mainly caused by the volatility of international commodity prices. These countries usually rely on a few primary exports, which makes them less resilient to variation in their prices. Moreover, many scholars have studied the tendency of resource-rich countries to under-perform economically when compared to resource-poor ones. This body of research is often discussed under the auspices of the Resource Curse literature, according to which natural resource dependency hinders economic growth and is

¹A supercycle is a fluctuation of a broad range of commodity prices that tends to span a long period of time with upswings of 10–35 years, generating 20–70 years complete cycles (Erten and Ocampo, 2013). In the latest supercycle, the world economy experienced a high uprising of commodity prices in 2010 – 2012 and their collapse in the following years, which can be seen as a continuation of the 2004 – mid-2008 boom (Erten and Ocampo, 2013).

²Saudi Arabia, Russia, Iraq, United States, Canada, United Arab Emirates, Kuwait, Nigeria, Qatar, Angola, Kazakhstan, Mexico, Venezuela, Norway, Oman, Azerbaijan, Algeria, Iran, Colombia, United Kingdom, Brazil, Ecuador, Libya, Indonesia, Equatorial Guinea, Australia, South Sudan, Republic of the Congo, Malaysia, Gabon, Vietnam, Yemen, Denmark, Bahrain, Syria.

correlated with weak institutions and a lack of human capital, among other issues (Sachs and Warner, 1995, 2001; Gelb and Grasmann, 2010; Gelb, 2011; Bornhorst et al., 2008; Stevens et al., 2015; Papyrakis, 2017; Lederman and Maloney, 2006; Sy et al., 2011; Humphreys et al., 2007; Auty, 1994, 2001, 2002; Auty and Pontara, 2008; Murshed and Serino, 2011; Ahmadov, 2014).

Among the issues associated with the Resource Curse, existing literature argues that countries' overreliance on a few natural resources prevents them from diversifying their productive base for two main reasons. Firstly, the Dutch Disease³, which is the economic dimension of the Resource Curse, has the potential to deindustrialise the non-resource traded sectors (Corden and Neary, 1982; Déés et al., 2008; Sachs and Warner, 2001; Gelb, 2011), because the influx of cheaper imports, due to local currency appreciation, could make the domestic manufacturing industry less competitive (Humphreys et al., 2007; Auzer, 2017; Papyrakis and Pellegrini, 2019). These effects not only affect existing industries, but also impact the countries' capacity to develop new productive sectors. Secondly, research has shown that countries' productive structures are a reflection of the knowledge and know-how available and accumulated in an economy over time (Milanovic, 2013; Hidalgo et al., 2007; Hidalgo and Hausmann, 2009; Hausmann et al., 2014; Hidalgo, 2015). Thus, the existing productive structure of a country could limit their ability to produce new and diverse goods, unless new knowledge is acquired (Hausmann et al., 2014, 2010; Hidalgo et al., 2007). In the case of countries whose knowledge is limited to the extraction of few natural resources, the variety of goods that can be produced is constrained by the accumulated know-how and this usually limits them to move beyond the production of raw materials. Research has also shown that these countries' productive diversity and respective growth would often be lower than economies with a more diversified knowledge base (Hartmann, 2014; Hausmann et al., 2014; Hidalgo, 2015).

To overcome the problems of overreliance on natural resources and as attempts to avoid the resource curse, scholars have argued that countries must diversify their productivity and significantly reduce their primary sectors' dependency. Studies have shown that there are many benefits of having a more diverse economy. For example, Auty (2001); Collier and Page (2009); Lederman and Maloney (2006); Ross (2013) explain that having a bigger variety of products to export protects these countries against the effects of commodity prices' volatility (specially for the case of crude oil). Similarly, productive diversity could help their economies to create new jobs less dependent on the resource sector, so employment

³The Dutch Disease occurs when a country experiences a sudden windfall of income from an extractive industry and subsequently its non-traded sector is crowded out by the growth in the natural resource sector (Corden and Neary, 1982; Humphreys et al., 2007; Auzer, 2017; Papyrakis and Pellegrini, 2019)

opportunities will not rely mainly on the rents created by natural resource extraction. In the long run, Caldecott et al. (2014) argue that a more diverse productive base can prepare oil exporting economies for future resource depletion and/or even reduce the risks associated with carbon stranded assets. Notwithstanding the importance of this topic, surprisingly little is known about how this can be done, as Ross (2019) notes. Most of the existing literature related to oil-dependence has placed a disproportionate amount of attention to the macroeconomic effects that a windfall of crude oil revenues generates, at the expense of the mechanisms that limit the diversification potential of those economies. Furthermore, little is known about the regional dynamics that could boost their diversification and could guide policy making to accomplish these goals. This research aims to fill in this gap.

1.2 Research Objectives and Goals

This thesis argues that one of the fundamental constraints for diversification in oil exporting economies is linked with the ways in which their economic structures are configured, as hypothesised by existing literature. More specifically, this dissertation argues that understanding the productive structures of countries could help us explain why they have a limited capacity to generate new economic activities, which is based on the productive knowledge they generate and accumulate. Furthermore, improving the understanding of these structures could provide insights to guide policy making and investment decisions for diversification pathways.

To explore the main arguments of this thesis and shed new light on crude oil export dependency and economic diversification, this dissertation is guided by and organised around three main general questions⁴:

- a. What is the productive structure of an oil exporting economy?
- b. How is this productive structure affected by a resource boom?
- c. How can this resource dependency be overcome?

To address the first question, this thesis proposes a novel methodology that would allow us to extract the structural information of input-output (I/O) data of a given country. This methodology helps us to assess the importance of different industries of a country, as measured by industrial centrality, based on their inter-industrial (I/O) relationships (including oil), which provides an approximation of its productive structure⁵.

⁴The questions complement each other, but they will be dealt separately in the individual empirical chapters (3, 4 and 5).

⁵A way to explore the internal economic structure of a country is to understand the role that individual sectors and products play, as well as their corresponding relationships and dynamics over time.

Traditionally, Input-Output analysis is a good tool for identifying key sectors based on the linkages between industries. However, this kind of identification can not be performed when data are incomplete, especially in cases where only supply or use tables are available.⁶ The proposed methodology attempts to overcome this limitation by using the tools developed by network science. Specifically it calculates the network centrality for different industries in cases when such cases are prevalent.

Despite the recent scholarly advances, much of the research that focused on productive structures use international macroeconomic and trade data, for generalisability. In addition, previous studies have not looked at how resource-rich economies are configured internally nor investigated the interactions that exist between industries and products. Further, current research and their methodologies have focused on places where data are rich and easily accessible and excluded countries (mostly developing ones) where historical data is difficult to find or is incomplete.

The second question attempts to explore the structural effect that a resource boom has on the productive structure of a country over time, assessing industrial centralities' changes for a given time period, as per calculated in the first part of the thesis. In particular, I hypothesise that the economic shock generated by windfalls of oil resources has minimal impacts on the productive structure of an oil-dependent country. To test this hypothesis, I will use a synthetic control method to identify a possible 'structural' crowding-out effect of the non-traded sectors on the economy (i.e, the Dutch Disease 'structural' effect). Both, the first and second questions attempts to contribute to the resource curse literature by taking a structural perspective, which will be explain in detail in the respective chapters.

To explore the solution to overcome resource dependency, as posed in the third question, I analyse how current concentrations of industries in a given place influences the emergence of "pioneering" economic activities (i.e., the "Principle of Relatedness"⁷, which signals future paths for diversification within a country). This analysis takes a holistic point of view, encompassing all economic sectors as opposed to the usual focus on manufacturing. Furthermore, how production and consumer networks influence the emergence of pioneering activities are examined.

Noguera-Méndez et al. (2016) argues that changes on these inter-industrial relationships, measured by their centralities, could affect the economic system and modify its structural characteristics.

⁶The methodologies used to study input-output relationships before the 1980s were not the I/O tables that are known today. According to the data examined for this thesis, there were only use tables, which were incomplete to build a current I/O table

⁷The strong relationship between how related economic activities in a place are and the emergence of new economic activities in related sectors, is known as the "Principle of Relatedness". This Principle has been reported extensively in the economic geography literature, and has been summarised by Hidalgo et al. (2018).

To address the research questions set out above, this thesis uses Ecuador as a case study. This country was selected for empirical investigation because of its long economic dependency on crude oil exports and few other primary goods. Relatively good data accessibility and quality were also one of the main reasons for this selection. Ecuador has consistent time series information on its National Accounts (including Input/Output tables) since the late sixties, which allow us to track its inter-industrial relationships over time and track the effects of the oil boom. These data are key to address the first two questions. To address the last question on diversification, a proprietary⁸ unique and large dataset⁹ provided by the Internal Revenues Service (SRI) of Ecuador under a non-disclosure agreement, is used. This data set comprises fine-grained information of all the legal (firms) and natural persons¹⁰ registered in the country from 2007 to 2015 (approximately 150 000 entities per year). This data includes records of every transaction between all these entities¹¹ (approximately 7 million observations), their locations at the municipality level, their industrial classification (ISIC) at a 4 digits level and more. This unique dataset is key for understanding how real production networks are structured and will be key to understand the commodity dependence issues in-depth and will be the foundation for the diversification analysis in this academic piece of work.

The three research questions deal with specific components to support the main argument of this thesis, in that understanding the economic structures is crucial for a resource dependent economy to grow and diversify. The first two questions help us to understand the structural dependency of an oil-exporting country and how its economic structure could constrain the development of new productive activities. In particular, the first question examines how an economy is configured by extracting structural information out of I/O tables (Use Tables in the case of Ecuador), which allow us to study how each industry is linked with the rest of the economy and its relative importance for promoting growth. The second question explores the effects that a resource boom has on the productive structure that was assessed previously. As explained before, I hypothesise that the structural impact of this boom is minimal because the oil industry operates as an enclave¹², which is weakly linked with the rest of the economy and its knowledge spill-overs are limited. The last question is built

⁸The dataset is not publicly available and it is not commercial.

⁹As far as the author is aware, there are a few countries that have similar records, which includes Japan (Tokyo Shoko Research (TSR), LTD), South Korea (KOSIS Korean Statistical Information Service.), Switzerland (Federal Tax Administration), Chile (Internal Taxes Service)

¹⁰In jurisprudence, a natural person is "a living human" or "one single person" (in legal meaning, i.e., one who has its own legal personality) as opposed to a legal person, which may be a private or public organisation (CLS, 2020; CD, 2020).

¹¹Natural and Legal Persons.

¹²This is one of the main conclusions of Chapter 3 of this thesis, which is supported by existing academic literature.

upon the previous analyses and deals with the challenges of diversification. It begins by analysing of the current productive structure of the case country and subsequently assesses the likelihood of new industries to emerge in the country based on the relatedness principle. Altogether, these questions help us understand how the productive structure of a country could limit its capacity to generate new economic activities, and considering it could set a path for future diversification. Further, these questions are a step forward to address the research gaps identified by Michael Ross (2019)¹³ in relation with the literature on productive diversification of oil-exporting economies.

1.3 Ecuador’s oil dependency as a study case

In 1967, crude oil was discovered in the Ecuadorian Amazon. This discovery significantly boosted economic growth and, in many ways, changed the social structure of the country (Correa, 2012). Ecuador, once one of the poorest, became one of the middle-income countries in South America. In the early 1970s, real GDP grew at an average rate of almost 15 per cent every year, of which more than 60 percent was attributable to crude oil exports. The value of exports (at constant prices) grew by 60 percent per year (Belli, 1984). Due to oil exports, Ecuador’s terms of trade turned strongly in its favour (Belli, 1984).

By the 1970s, non-oil exports were mostly composed of bananas, coffee and cacao beans, which enjoyed favourable international prices and received subsidised credit, as well as other fiscal benefits. However, they were affected by overvalued exchange rates during the oil boom period, and as a result, they experienced slow growth (Larrea, 1992). With the important exception of shrimp, non-oil exports grew slowly in volume after 1974 and fell by a third in 1983 (Belli, 1984; Acosta, 2006). Several authors argue that these effects were caused by the “Dutch Disease” (Acosta, 2009b; Larrea, 2013).

During this period (1972-1982), the manufacturing sector grew but its economic performance was weak (Belli, 1984). In fact, during the oil bonanza, structural diversification of manufacturing was minimal (Belli, 1984) and the attempts of industrialisation were mostly short-lived –. A small expansion was observed in protected sectors (mainly agriculture)(Belli, 1984). Moreover, Vos (1987) and Belli (1984) analysed the changes in the internal structure of manufacturing and concluded that domestic linkages were low and changed little during the oil boom and afterwards.

¹³Ross (2019) argues that despite the importance of diversification, little is known about how these economies can achieve it. Existing literature has focused on the challenges that oil exporting economies face, having failed to shed light on how to overcome their dependency on few commodities. He argues that scarce, missing or unreliable data regarding these countries is one of the reasons of the limited research on this topic (Ross, 2013).

Currently, Ecuador still depends heavily on crude oil exports. These represent approximately fifty percent of its total exports and, on average, 30% of the total income of the state budget in the past six years ¹⁴. However, it has been stated that these oil reserves have already reached the peak and will be exhausted within the next 30-40 years (Espinoza et al., 2019; Mateo and Garcia, 2014). In fact, in several scenarios, Ecuador could become a net oil importer between 2024 and 2035 and economic oil trade balance could be seriously affected, decreasing from a current positive value of around 2 billion USD to incur deficits of 0.6–16.7 billion USD in 2035 (Espinoza et al., 2019).

Finally, the overall economic growth in Ecuador after crude oil discovery was poor, with an average growth in per capita income (1.38% per annum), which was lower than the pre-oil period (2.42% per annum), with minimal productive diversification. In addition, only shrimps and flowers have been added to the export basket in the past four decades (Larrea, 2013). Manufacturing has not significantly increased its share in its GDP. The banana sector, which is Ecuador’s biggest export after oil¹⁵, is facing serious challenges. It is possible that a fungus (*Fusarium R4T*), having been detected in Colombia, may affect Ecuador’s production as well, with the potential to diminish the country’s plantations (Perez, 2019). As such, it is not only important to understand its commodity dependence in more detail but also finding diversification pathways.

As it can be observed, Ecuador has been reliant on oil exports for almost half of a century and it has not been able to broaden its product base despite some unsuccessful and short-lived attempts by different governments. However, this is a norm rather than an exception for most oil-exporting economies. For instance, Michael Ross (2019) found that oil producers have moved sharply toward greater export concentration during the recent price boom, much as they had in the 1970s. Moreover, only eight states¹⁶ managed to diversify between 1998 and 2010. In this thesis, I test the methodologies proposed using Ecuador as a case study, which could be applied in other contexts. However, the insights obtained from this analysis must be interpreted with caution. Countries are inherently different. Ecuador, for example, is an upper-middle income country according to the World Bank (2020). It has a democratic government and has no control over its monetary policy¹⁷. In contrast, several oil exporting economies do not have democratic governments – Saudi Arabia or the United Arab Emirates, for instance. In the case of Venezuela, it has an authoritarian

¹⁴According to Ecuador’s Central Bank statistics and information of capital flows and income of central government budget (BCE, 2019).

¹⁵Banana exports generate \$US 2.6 billion and 2.5 million jobs in the country could be seriously affected in the coming years (Perez, 2019).

¹⁶Four of them did so due to resource depletion or economic sanctions.

¹⁷Ecuador changed its currency to US Dollars in 2000.

regime. On the other hand, some oil exporting countries are either high-income countries with diversified economies, such as Canada or Russia, whereas Nigeria, is lower-middle income country (World Bank, 2020). Further, fossil fuel reserves in there countries as well as their geopolitical context could yield findings which are different to our analysis using the same methodologies. As such, the Ecuadorian case results may or may not be generalizable to other contexts.

1.4 Research Significance

As previously discussed, overcoming the reliance on oil exports has been a long-standing challenge for many countries. Although diversification is generally considered to be one of the solutions to overcoming oil dependency, little is known about how it can actually be achieved (Ross, 2019). This project advances our knowledge on oil dependency at a structural level and how it may be overcome.

In particular, this study uses an experimental approach to look at the productive structure of an oil exporting country, Ecuador, using its national Input/Output tables. This research develops an innovative network science-based methodology to overcome data limitations on incomplete I/O tables using only supply¹⁸ or use tables¹⁹ in order to extract structural information and are able to assess the relative importance of different industries, measured by centralities, based on the linkages between them.

I subsequently test if there is a crowding-out effect on the non-traded sectors due to a resource boom (the Dutch Disease). To this end, I use a synthetic control method to assess how the industrial centralities of Ecuador changed before and after an oil boom (1972-1973) for the period between 1965 and 1985, and whether this boom caused any re-configuration of the productivity of different sectors in the country. Furthermore, it allows us to identify which sectors were the most affected and/or the potential beneficiaries of the oil boom.

Most importantly, this thesis deepens our understanding of how we may overcome oil dependency by productive diversification. Based on the existing literature on regional diversification, this thesis explores and expands the well-known "Principle of Relatedness" or the "Relatedness Based on Co-location" as proposed by recognised economic geographers²⁰. This principle argues that the probability that a region will develop a new activity (i.e.,

¹⁸A supply table describes the supply of goods and services, which are either produced in the domestic industry or imported(UNSTAT, 2018).

¹⁹A use table shows where and how goods and services are used in the economy(UNSTAT, 2018).

²⁰The following scholars recognised this principle in the cited article: Ron Boschma, Mercedes Delgado, Maryann Feldman, Pierre-Alexandre Balland, César Hidalgo, Koen Frenken, Edward Glaeser, Dieter F. Kogler, Andrea Morrison, Frank Neffke, David Rigby, Scott Stern, Siqi Zheng, Shengjun Zhu.

exporting a new product, creating a new industry, or developing a new technology) increases with the number of related activities present in a location, or with the presence of that activity in neighbouring locations (Alshamsi et al., 2018; Hidalgo et al., 2018; Boschma, 2017). In this thesis, I test this hypothesis using a unique dataset at a granular level from Ecuador and explore the power of the principle to explain the emergence of pioneering industries in given places. More importantly, this research goes beyond the existing literature because it quantifies “actual relatedness” based on the production and customer networks of different economic sectors, as opposed to the principle of relatedness which is on co-location alone. Doing this helps us to better understand the role of production and customer networks in economic diversification, which is relevant to identify potential diversification paths both at country or smaller industrial and/or firm scales. Also, this research makes a comparison between the extractive industry (oil and mining sector) and manufacturing, to find key differences for their diversification processes.

Overall, this dissertation attempts to link two bodies of literature that, as far as I am aware, have not done before: the resource curse and regional diversification. It argues in particular that the literature on regional diversification has much to offer to the problems of the resource curse. Methodologically, this thesis adopts an innovative and experimental network science approach which enables us to study the inter-dependencies of economic sectors, which traditional economic analyses fail to do. It is important to recall that the network approach to study production networks is still under development and this thesis contributes to this bigger academic effort. Also, this thesis has generated new data that has been made for public use. For instance, in the first part of the thesis, all the Input/Output records of Ecuador between 1965 and 1995 were digitised. I also cleaned and systematised the most comprehensive dataset on economic activities of all the firms and individuals in the country. Several open source data visualisation tools have been developed to explore interactively all the data used in this thesis. Finally, but not less relevant, this thesis attempts to shed light on one of the most pressing political and economic challenges of the country: the need to diversify its economy and reduce its reliance on the oil industry.

1.5 Thesis Outline

This thesis consists of six chapters. Chapter 2 offers an overview of the existing literature and explains how this dissertation fits in the broad academic debates. Specifically, this section covers discussions on economic development in the context of resource dependency. Later, the chapter focuses on the structuralist tradition of economic growth, in particular with references to work done by Hans Singer, Raul Prebisch and Albert Hirschman and

structural transformation. This literature is useful because it provides a framework, which will be discussed more extensively in Chapter 3, to understand how the productivity of countries are configured and how these configurations influence economic growth. Subsequently, the ideas of Kenneth Arrow, Paul Romer and Robert Lucas on endogenous technological change are connected to the aforementioned literature, which is relevant because the productive structure of a country is strongly linked with the knowledge and technology that is endogenously created or acquired in a given place. Both structuralists and technological change are key for understanding oil-export dependency. Finally, this chapter supplements the previous discussions with existing debates on regional diversification and geography of knowledge, which are important to overcoming resource dependency. At this point, it is important to stress that each substantial chapter (Chapter 3, 4 and 5) has its own and in-depth literature review, which also makes clear of their contributions to specific literature. Similarly, each chapter has its own methodology section.

Chapter 3 addresses the first research objective, as explained in Section 2.1. It will propose a network-based methodology that allows the study of the productive structure of a resource-rich economy. More specifically, it assesses the importance of different industries of a country based on their inter-industrial relationships, which also allows us to examine what is the structural role of the crude oil industry. Using this methodology, the chapter proposes a new measurement that captures the structural role of an industry measured by its centrality. This methodology is tested using Ecuador as a case study. As explained previously, the country was selected given the relatively low complexity of its economy as well as the availability and accessibility of consistent Input/Output (Use Tables) data for 30 consecutive years (1965-1995). This data was compiled and validated by the United Nations Development Program and the Central Bank of Ecuador.

This chapter starts by exploring the academic discussions about structuralism and economic growth in detail, the role of key industries in achieving economic development, as well as summarising the literature gaps that this chapter attempts to fill. The next section of the chapter describes the data used for the empirical part and explains the methodology proposed. In the results section, the findings will be contrasted with traditional aggregate measurements and a discussion about how the proposed approach may or may not complement the ways in which economic systems are analysed. Finally, several conclusions are stated. It is important to note that this methodology will be the foundation for the analysis to be developed in the next chapter.

Chapter 4 deals with the second research objective of the thesis and allow us to assess the structural impact of the windfall of oil resources (oil boom) in a country. More specifically,

this chapter aims to analyse how the Ecuadorian productive structure - measured by the centrality of their industries - has changed between 1965 and 1985, and identifying the effects caused by the 1972 oil boom with the help of a synthetic control method developed by Abadie et al. (2010, 2015). In other words, this chapter assesses if the Dutch Disease's effects can be observed on a structural level²¹ by identifying the crowding-out effect of the non-traded sectors. The data used are industrial centralities calculated in the previous chapter.

This chapter will present a detailed literature review, which will summarise the existing debates around the Resource Curse, the Dutch Disease and its effects, as well as the specific literature gaps that the chapter attempts to cover. Second, the chapter examines the evolution of the Ecuadorian economy with a focus on the main commodities, as well as the political conditions that surrounded the oil boom. Next, the methodology used will be described, including a discussion about its benefits and limitations of the proposed approach. Finally, the results will be presented, followed by the corresponding discussion and relevant conclusions.

In Chapters 5, the scale of the analysis moves from a macroeconomic perspective (as in previous chapters) towards a firm level. By looking at firm and city level data, which were provided by the Internal Revenue Service of Ecuador, and using logistic regression and machine learning, this chapter explores if the Principles of Relatedness holds true in a developing country context. Furthermore, this chapter contributes to the literature further by exploring the role of supplier-based relatedness (supply chains) and consumer-based relatedness for developing pioneering industries in a given location. This chapter addresses the third research objective, as explained in Section 2.1.

This chapter starts with an overview of the existing literature on economic diversification in oil exporting economies (Section 5.1), the role that relatedness plays in regional diversification (Section 5.2), and identifies the research gaps to be filled. Subsequently, a description of the data used will be made, accompanied by descriptive statistics, and the econometric and machine learning models proposed for the analysis are explained. In the last section of the chapters, the main findings will be reported and key takeaways will be drawn.

Finally, the concluding chapter (Chapter 6) brings together the analyses undertaken in the preceding chapters and reflects upon the empirical contributions of the thesis as a whole. The chapter starts by providing an overview of this thesis and brings together the analyses undertaken in the preceding chapters and discuss the contributions of the thesis as

²¹Changes in the industrial centralities calculated in the previous chapter.

a whole. The chapter begins by summarising the motivation and objectives of this thesis. The following section is used to reflect upon the main findings of this research in relation to the three guiding research questions posted in this introductory chapter. Based on these findings, it will consider the policy implication, focusing mainly on diversification issues, and suggest pathways for future research. In addition, the appendixes will summarise some preliminary ongoing and future work.

Chapter 2

Literature Review

Development is an intrinsically complex and contested concept, so it is possible to find hundreds of different definitions (Staudt, 1991). These various definitions, and debates surrounding it, are messy. For instance, Cowen and Shenton (2010) explain that there is ‘good’ and ‘bad’, ‘under’ and ‘over’ development. Likewise, they argue that “development” could be ‘autonomous’, ‘appropriate’, ‘gender-conscious’, ‘sustainable’, or the opposite of all these (Cowen and Shenton, 2010). Moreover, it is important to recall that the terms used to define ‘development’ depend upon the views and perspectives of those in positions of authority in national governments, international agencies, universities, among others (Cowen and Shenton, 2010). Therefore, the concept of development itself is a political concept.

There is, however, a general understanding that it relates to the “betterment” of humankind through the alleviation of poverty and the realisation of human potential” (Cowen and Shenton, 2010, p. 3). For many countries, achieving this idea is one of their main priorities and a vast amount of international and domestic efforts and resources have been invested in order to accomplish this goal. “Development” has been named the central organising concept of our time (Cowen and Shenton, 2010) but reaching it has been elusive for many “developing” countries.

This thesis does not attempt to engage with any of the previous mentioned debates. However, for practical terms, it subscribes to the idea that economic growth plays a central role in the development debates and it is one of the main components to achieve the betterment of humankind, poverty alleviation and the realisation of human potential. Although I am very aware that more elements are needed to accomplish any development goal (i.e., the capabilities considerations proposed by Amartya Sen (1990), the institutional approach proposed by Douglass North (1995)) and a more comprehensive approach is needed, the main

purpose of this thesis gravitates around the previous idea. Therefore, this literature review is focused on some of the different approaches and debates related with economic growth, notably the structuralist tradition and Endogenous Growth Theory. These are foundational literatures to understand the debates around the main objectives of this dissertation outlined in the introduction. However, each substantial chapter has its particular literature review and will specify the contributions it makes to the literature.

This chapter is divided in three sections. It begins with a brief overview of the relation that exists between economic growth and development. This section (2.1) summarises how the ideas of material progress and economic growth became central in the development debates, starting with the theories proposed by Adam Smith and his contemporaries. Section 2.2 explores the ideas proposed by Hans Singer, Raul Prebisch and Albert Hirschman, or new structuralist thinkers such as Lin (2017); Diao et al. (2017), who argued that productive structures have specific characteristics that constrain the economic performance of countries or regions (Chenery, 1975; Prebisch, 1950; Singer, 1950a,b). This will be crucial to engage with conversations regarding the economic structures of oil exporting economies and their role as providers of raw materials for the global economy. This section is the foundation to address the research objectives 1 and 2. The next section (2.3) explores the ideas proposed by Kenneth Arrow, Paul Romer and others, in relation to "Endogenous Growth Theory". Their main arguments focus on the significant role of human capital, innovation and knowledge for economic growth. Their ideas are particularly useful to understand the constraints that have limited growth in oil exporting countries and its relation with their limited production of knowledge and technology. This literature also provides the basis for discussions related with the economic diversification of these countries and allows to create a link with the literature on regional diversification in Economic Geography. This section will help us address the third objective of this thesis.

2.1 Development and Economic Growth: an overview

In his book "Theories of Development", Jorge Larraín (2013) explains that the concept of development first emerged with the advent of capitalism in the 18th century, which encouraged the evolution of production forces, allowing the ideas of material progress to gain prominence in society. During that period, Adam Smith, John Locke and their contemporary thinkers proposed the idea that individual self-interest indirectly promotes welfare and societal progress (Smith, 2010). Thomas Hobbes argued that rational self-interest was morally and philosophically legit for the emerging capitalist system (Levy, 1954). He also suggested that individuals have the right to preserve their own lives and subsist on products

from the Earth (land and other material resources) by appropriation and exercising control rights over them through labour (Locke, 2016).

For these thinkers, amongst others, who set the foundation of modern economic thought, material progress was the ultimate goal of society. For instance, David Ricardo (1891) argued that countries could generate economic growth and, therefore, progress by producing commodities according to their own geographical characteristics, which could be exchanged with other countries. Ricardo (1891) assumed that all trade partners would benefit from an increase in total production in accordance with their comparative advantage. Even Karl Marx, a fierce critic of the basic ideas of capitalism, believed that progress could be achieved by material means by the generation of value through labour. He believed that social relations that combined labour with means of production determinate the quality and quantity of productive development (Marx, 2019; Peet and Hartwick, 2015). Since the 18th century, there was thus a sort of consensus that physical capital accumulation could be used for the ‘betterment’ of society.

Since then, many new economic ideas and a variety of approaches have been proposed to deal with the questions of economic development. For instance, some economists, most notably John Maynard Keynes, argued that the State should have direct involvement into development by managing interest rates and fiscal and monetary policies. For Keynes, this could shift the economy from one equilibrium level to another, therefore generating more employment (Moggridge, 1980). Keynesian economics argues that fiscal policy is an effective measure to increase social investment (health, education and so on). Furthermore, it supports the idea that general government spending is crucial for the economy (Keynes, 2018).

Robert Solow (1956) and many others argued against Keynesian economics, stating that market mechanisms (instead of state intervention) could bring the economy to an equilibrium, as well as warranted economic growth. In particular, Solow showed that the neoclassical approach to growth is stable, stating that its self-adjusting mechanism guarantees a return to equilibrium (Solow, 1956). For him, the main source of economic growth is the increase in capital and labour, and technology is an important component to increase the supply of labour (Solow, 1957). Many of economists contended to his line of thinking and argued that, if left unimpeded (i.e., limited role of the state), market processes will result in the optimum levels of production and allocation, leading to economic development (Peet and Hartwick, 2015).

Contemporary economist such as Daron Acemoglu argues that although there is disagreement amongst economists about how to approach short-run macroeconomic phenomena,

there is a broad agreement on the workhorse models of dynamic macroeconomic growth, which includes the Solow's growth model, the neoclassical growth model, the overlapping generations model and models of technological change and technology adoption (Acemoglu, 2009). This common agreement explains that there are 'approximate causes of economic growth and economic success', such as physical capital, human capital and technology, and that the main differences between countries' growth are associated with how these "capitals" are generated and/or distributed. The latter is, according to Acemoglu (2009), the fundamental question regarding economic growth.

At the risk of oversimplifying complex phenomena of economic growth, Acemoglu (2009) explains that we can think of the following list of its potential fundamental causes for achieving growth:

- Luck (or multiple equilibria), which are a set of fundamental causes that lead to divergent paths of economic performance among societies with identical opportunities, preferences and market structures. This can happen because of some small uncertainty between societies which can lead them to different choices with far-reaching consequences. For example, Acemoglu argues that the importance of luck can be made by emphasising the role of leaders. He explains that "perhaps it was Mao who held back China, and his death and the identity, beliefs and policies of his successor were at the root of its subsequent growth". In other words, the identify of a leader can be viewed as a stochastic event that could shape economic performance.
- Geographical differences that affect the environment in which individuals live and that influence the productivity of agriculture, the availability of natural resources, certain constraints on individual behaviour, or even individual attitudes. For example, soil quality, which can affect agricultural productivity can influence the country's growth processes. Similarly, mineral resources such as copper or iron could facilitate the development of associated industries. Or crude oil, which directly contributes to the wealth of a nation, could also affect the development of other industries, as it will be explored further in this thesis. Even, where a country is located could have influenced historically the accessibility to trading routes.
- Institutional differences that affect the laws and regulations under which individuals and firms function and thus shape the incentives they have for capital accumulation, investment and trade. As the author explains, these differences refer to rules, regulations, laws and policies that effect economic incentives, which have a direct affect on the accumulation of other production factors as technology, physical capital and

human capital. These institutional differences differ from *luck* and *geography* because they are *social choices*.

- Cultural differences that determine individuals' values, preferences and beliefs have a significant influence on individual economic behaviour. For example, religious identities across societies are one of the clearest examples of how cultural differences affect economic behaviour. A study done by Keister (2003) found that different religious affiliations can shape action indirectly by altering fertility and marriage behaviour, educational attainment, work behaviour, and other behaviours, which have an influence on wealth accumulation and ownership. The study points out that Jews, for instance, enjoy tremendous gains in wealth ownership, while conservative Protestants accumulate relatively little wealth. In contrast, Protestants and Catholics are indistinguishable from each other and from the general population.

These different realities can be observed, for instance, in the case of the Americas. Armendáriz and Larraín (2017) explain that there are clear differences between the former British and Spanish/Portuguese colonies in terms of 'development' and there are many factors that contributed to these difference. From a geographic perspective, for instance, there are different weather patterns that influence agricultural practices; the tropical regions of the continent are not "blessed with predictable water from defrosting and low exposure to the sun" (Armendáriz and Larraín, 2017, p. 5). Also, there is a higher incidence of tropical diseases and the health standards are lower compared to those in the north. Therefore, labour productivity due to poor health reasons is comparatively lower in Central and South America (Armendáriz and Larraín, 2017). Moreover, the legal systems and institutional (colonial) legacy have significantly influenced, the authors argued, how private ownership was conceived. These are just few examples on how the previous items identified by Acemoglu (2009) mark clear differences on countries' economic development.

In the case of Ecuador, some of these aspects can be observed. From luck and institutional perspectives, the country lived a period of high political instability from the mid-90s to the mid-2000s. During this period, the country had eight presidents¹, two coups and after a big financial crisis in 2000, the US Dollar was adopted as its national currency (Acosta, 2006). This instability affected negatively the country's growth and have repealed foreign direct investment until nowadays (Correa, 2012). From the cultural perspective, Hurtado (2018) argues that corrupt practices are rooted in the society, which diminish trust. People

¹ Abdala Bucaram (August 10, 1996- February 6, 1997), Rosalia Arteaga February 6, 1997 - February 11, 1997), Fabián Alarcon (February 11, 1997 - August 10, 1998), Jamil Mahuad (August 10, 1998 - January 21, 2000), Gustavo Noboa (January 22, 2000 - January 15, 2003), Lucio Gutierrez (January 15, 2003 - April 20, 2005), Alfredo Palacio (April 20, 2005 - January 15, 2007)

are afraid of being deceived or taken advantage of by strangers or even elected authorities, which led to the creation of unnecessary and somehow ineffective corruption controls (both public and private), resulting in a waste of time and resources. Finally, from a geographical perspective, which is the main focus of this thesis, Ecuador has a long-standing dependence on natural resource exports, mainly crude oil, which has helped the country to develop but it has also affected its capacity to create new industries. This question will be explored in detailed in the empirical chapters of this thesis.

As previously stated, there are long discussions regarding the ways to achieve economic development and myriad theories to back those discussions. A comprehensive overview of these goes beyond the scope of the thesis. Below, I focus on the two main economic theories mentioned above, which provide the theoretical framework to address the research questions and upon which the empirical chapters are grounded.

2.2 Productive Structures and Growth

2.2.1 Structuralism, import substitution industrialisation and economic growth

One way to study the unique features of an economy is through the lens of structuralism in Economics, which was originally developed in the late 1940s in Latin America. This approach emphasises the importance of "structures" that affect the economic development of countries, especially the developing ones (Dutt, 2019). Economists such as Hans Singer, Raul Prebisch and Albert Hirschman argued that the countries' productive structures (along with the relationships between developed and developing countries) could constrain their economic performance (Chenery, 1975; Prebisch, 1950; Singer, 1950a,b). These authors understood "economic structures" as the composition of productive activities associated with some influences, such as "the specialisation pattern in international trade; the technological capacities of the economy, including the educational level of the labour force; the property structure of the production factors; the nature and development basis of institutions; the level of development and restrictions under which certain markets operate, among other" (Missio et al., 2015, p. 248). This structural tradition contrasts with the neo-classical approach in Economics, which arguably views different economies as being similar, with markets generally allocating resources efficiently (Dutt, 2019).

These economists stated that the aforementioned productive structures could explain phenomena such as balance of payments disequilibrium, the limited ability of economies to absorb workforce, their vulnerability to external shocks, or even unequal income distribution (Chenery, 1975). Furthermore, they argued that "if a developing country wants to be

as rich as a high-income country, it needs to have the same high level of productivity as the high-income country, which in turn requires the country to develop the same advanced capital-intensive industries” (Lin, 2019). In other words, structuralism advocates promoted the viewpoint that developing countries should take, to a certain extent, the industrial structure of high-income countries as their development goal (Lin and Wang, 2020). Thus, nations should progressively change their imports and internal production structures (i.e., structural change) (Prebisch, 1950; Singer, 1950a,b; Harvey et al., 2010a), which would also allow them to end dependency on primary product exports (Love, 2005).

Structuralists promoted the idea of ”Import Substitution Industrialisation² (ISI)” and the development of regional-wide markets to ”capture economies of scale in production” (Love, 2005, p. 104). However, they recognised that industrialisation could not develop spontaneously in developing countries’ markets due to structural limitations such as information externalities, weak responsiveness to price signals or co-ordination problems within their economic systems (Lin, 2019). Moreover, these scholars argued that there are two fundamental sources of ‘structural deficiencies’ that would prevent industrialisation: those exogenous, such as the adverse conditions of trade and the limited capacity to import; and those domestic, such as accelerated population growth, premature urbanisation, underdevelopment of agricultural production, the reduced size of domestic markets and the presence of inefficient tributary systems (Missio et al., 2015). Therefore, to overcome these growth challenges, developing countries had to identify and address those limitations (Chenery, 1979, p. 310) and governments (the state) had to play an active role to correct those deficiencies as well as existing market failures (Lin, 2019).

Following structuralists, advise developing countries governments mobilised and allocated resources to build large-scale, modern industries. Similarly, countries built and improved hard infrastructure, such as power supply and road networks (Lin and Wang, 2020). However, despite the noble intention and the underlying theory, countries that adopted import substitution strategies (Latin American countries in the 1950s, for example) typically fell into universal failure, although the large-scale investments did bring rapid growth. The growth was followed by inflation and economic crises and the countries got stuck in long periods of stagnation (Sachs and Williamson, 1985; Lustig et al., 2016). According to some scholars, ISI strategies did not prosper because they ignored the endogeneity of countries’ economic structure and recommended the development on industries that were too capital intensive for their level of development and defied the comparative advantages determined

²ISI is a development strategy focusing on promoting domestic production of previously imported goods to foster industrialisation, mainly in developing countries (Bussell, 2018)

by their factor endowments (Lin, 2019). Furthermore, Lin explains that firms in the prioritised industries were non-viable in open competitive markets and required state subsidies and protection for their initial investment and continuous operation. Finally, it has been argued that the state's protection and subsidies needed for the implementation of ISI strategies led to interventions and distortions, causing misallocation of resources, rent seeking, corruption, and political capture (Lin, 2019).

The structuralist theory in Economics has been criticised. Several authors have argued that it does not support competition, ultimately affecting domestic innovation, and that the associated protectionism, to develop new industries, ended up failing (Williams, 2014). One of the main criticisms coming from "neoclassical" economists is that "trade restrictions and other kinds of state intervention stifled growth and development in many developing countries [the structuralist approach gives a more important role to the State, than traditional economists], and those countries that embraced relatively free market policies performed well" (Dutt, 2019). Other scholars have argued that, in practice, these ideas do not help the creation of national markets based on economies of scale and interfere with regional integration (Bulmer-Thomas, 2003; FitzGerald, 1998). Additionally, it has been stated that ISI did not increase productivity and that the process limited the expansion of traditional and new exports (Bruton, 1989). Finally, many authors argue that structuralism and ISI strategies failed because their implementation relied on bureaucrats that picked winner industries, which were usually not coherent with the endowment structure of a country Lin (2019); Lin and Wang (2020); Williams (2014); Dutt (2019). Moreover, Lin (2019) explains that "countries which followed these strategies placed a priority on capital-intensive heavy industries, that is, industries that made intensive use of a factor that they largely lacked, and that neglected to use many of the factors that they had in great abundance, such as unskilled labour and natural resources".

2.2.2 New structuralism and economic growth

New structuralist economists (NSE) have built on this structuralist tradition to argue that economic development is indeed a process of structural transformation, which involves technological innovation and industrial upgrading (Lin, 2017; Diao et al., 2017; McMillan et al., 2017; ECLAC, 1992). They emphasised that this new structural approach should focus mainly on three pillars: "technological upgrading; inequality reduction through productive employment creation, investment in human resources, social welfare, and transfer policies; and making the state more responsible to peoples' needs" (Dutt, 2019, p. 117). These thinkers also stresses that industrial structures need to be upgraded in order to achieve economic growth and development, particularly in developing countries (Lin, 2010). Likewise,

they argue that technological upgrading would open up new economic sectors that required and develop new skills (Dutt, 2019). New structuralists emphasised that structural change means moving into new productive sectors, but this is not a simple nor automatic process of capital accumulation, but requires policies to foster technological capabilities, as well as policies that develop specific infrastructure and institutions, human capital, specialised financial institutions, sophisticated local markets, and other components (Dutt, 2019; Lin, 2017, 2010; Milanovic, 2013). Also, Dutt (2019, p. 118) explains that "key sectors need to be identified, choices need to be made between different discrete paths, and a critical mass is required to take advantage of scale economies".

Both old and new structural approaches argue that the State needs to take a lead in development and economic growth. Both traditions stressed that a facilitating-state could achieve more growth goals than a pure *laissez-faire* approach (Lin and Chang, 2009). However, NSE propose the use of a neoclassical approach to study the determinants and impacts of the economic structure and its evolution, and argues for a market-friendly economic system (Lin, 2019). In contrast, old structural thinking relied heavily on a Keynesian tradition and promoted a stronger state intervention, mainly to achieve industrialisation (Missio et al., 2015). NSE stress that economic structures, including the structure of technology and industry, are endogenous to the endowment structure which is given at any specific time and recognise that this structure changes over time. Therefore, the new structuralism recognises that growth is a dynamic and evolutionary process of moving from one development stage to another (Lin, 2019), and that the role of the state is to encourage it, rather than to control it directly as in the previous approach.

Finally, Justin Lin (2019) explains that:

"The best way to achieve dynamic growth in an economy is to develop the economy's industries by following the comparative advantages determined by its endowment structure. The economy with such industries and appropriate hard and soft infrastructure will be most competitive, produce the largest surplus, have the highest possible returns to capital and thus savings, ensure the fastest upgrading of the endowment structure, and achieve the most rapid industrial upgrading and income growth. In this process, a developing country can have latecomer advantages and thus have a faster technological innovation and industrial upgrading than high-income countries, which leads to convergence with high-income countries".

2.2.3 Structuralism and economic complexity

In recent years, new tools have been developed to analyse economic systems and economic behaviours. These range from the dynamics of product adoption to financial contagions, where it is necessary to account for the patterns of interactions (Kirman, 2016; Barabasi, 2016; Jackson, 2010; Caldarelli et al., 2012; Beinhocker, 2006; Acemoglu et al., 2012; Haldane and May, 2011). Among these, the economic complexity approach has enabled us to quantify productive structures, which was previously dealt theoretically (Hirschman, 1958) or analysed based aggregated outputs (Input / Output Data) (Chenery, 1979). This approximation has gained much attention as its measurements are highly predictive for future growth, compared with other approaches developed to date (Abdon and Felipe, 2011; Bustos et al., 2012; Hidalgo, 2015; Cristelli et al., 2013, 2015; Felipe, 2012; Hausmann et al., 2014; Hidalgo and Hausmann, 2009; Hidalgo et al., 2007; Rodrik, 2006). Also, this new approach helps explain and predict international and regional variations in income, economic growth, income inequality, gender inequality and even greenhouse emissions (Hidalgo, 2021). Finally, the economic complexity debate has renewed the interest in the macroeconomic role of structural transformation (Hartmann et al., 2017a; Pinheiro et al., 2018; Alshamsi et al., 2018).

This economic complexity approach has fostered the development of new empirical research, which has confirmed, as previous structuralists argued, that productive structures of countries are not only determinants for economic growth (Djimeu and Omgba, 2019), but also a good approximation of the knowledge and know-how available in an economy (Milanovic, 2013; Hidalgo et al., 2007; Hidalgo and Hausmann, 2009; Hausmann et al., 2014; Hidalgo, 2015). This approach analyses the networks of relatedness between products, which allows us to understand the composition of a country's productive output. It shows that the mix of products (productive structure) that a country produces has strong connections with its level of economic growth (Hausmann and Hidalgo, 2011; Hidalgo et al., 2007; Hidalgo and Hausmann, 2009; Hausmann et al., 2014; Hartmann et al., 2017a). This approach has the potential to predict future patterns of diversification or specialisation, as well as the emergence of new industries in a country or region (Hartmann et al., 2017b,a; Hidalgo et al., 2018; Pinheiro et al., 2018; Asheim et al., 2011). In this sense, one could argue that the economic complexity approach offers new analytical / data driven tools to study the economic structure of a country and its evolution. These tools are enhancing our understanding on how productive structures change over time and offer guidance for structural productive reforms for countries to follow, based on their endowment factors (Hidalgo et al., 2007; Hidalgo and Hausmann, 2009; Alshamsi et al., 2018; Pinheiro et al., 2018). Therefore, this approach

has the potential to address some of the criticism that the original ideas and applications of structuralism thinking faced. For instance, Lin (2019); Lin and Wang (2020) argues that one of the main constraints that the application of these ideas and their ISI strategies was that the governments picked "winner" sectors based on industrialisation goals, ignoring countries' endowment factors. In this sense, economic complexity offers a robust analytical framework to understand these endowment factors. This could help us anticipate changes in specialisation patterns that would be key for industrial policy. Thus, new structuralism, supported by these new tools, has the chance to overcome "old" structuralist shortcomings. However, more empirical research is needed to assesses if this approach can succeed.

This structuralist literature is the theoretical foundation of the empirical chapters. This literature is crucial to help us understand that the economic configuration of an economy matters for its future growth. This literature will be expanded and the specific literature gaps, as well as the contributions will be further discussed in each individual chapter. In general terms, Chapter 3 draws from this literature to develop a methodology that assesses the importance of every economic sector of the Ecuadorian economy and identifies those industries that are key for its economic growth. The main contribution to this chapter's literature is the proposed methodology to assess the importance of economic sectors based on incomplete Input/Output data. Chapter 4 is based on the idea of *key sectors* proposed by the existing literature and analyses how a resource boom affect them as well as the productive structure of Ecuador. The main contribution of this chapter is that, by merging the structuralist literature with the resource curse debate, it empirically explores the structural effect of an oil boom. Finally, Chapter 5 is based on the new structuralist ideas, it uses the analytical tools developed by Hidalgo et al. (2007, 2018) to explore how the current productive structure of the Ecuadorian economy influences the emergence of pioneering industries, thereby helping to diversify the country. The main contribution here is the analysis built upon supply and consumer networks (as structures) for productive diversification of an oil exporting economy, which complements the current co-location approach.

To complement the structuralist literature, the next section will explore the debates regarding knowledge, technology and economic growth. This part of the literature review is key because the productive structure on an economy is strongly related to the existing knowledge that country has accumulated and developed over time (Utkovski et al., 2018). The next section will explain the link between economic growth, knowledge and technology as key factors to overcome the structural constrains for development.

2.3 Knowledge, Technology and Economic Growth

Since humans developed language, the transmission and accumulation of knowledge has been key for the evolution of our society. By exchanging ideas and practices, we have been able to develop better tools and find strategies that helped us to survive in complex environments. Human creativity has been essential to leave the stone age (millennia before) and the slow transition into our current digital era.

The explicit recognition of the importance of knowledge (first written references) dates back to Imam Ali, who was the ruler of the fourth caliph (which included the Rashidun Empire) in 599-661 CE. His teachings stressed that a connection between knowledge - power - wealth exists and he proclaimed that "knowledge is power" and "knowledge is a ruler and wealth is its subject" (Anand and Walsh, 2016). This notion was later popularised by Francis Bacon during the early seventeenth-century in England, who stated that "scientia potentia est" (García, 2001). These are but two examples that highlight the relevance of knowledge, also in connection to politics and economic growth. Throughout history, many scholars and philosophers have studied what knowledge is or means, how it is created, accumulated and transmitted, how it is used and so on.

Since medieval times, governments have understood that it was in their interest to subsidise the production and diffusion of knowledge "to support the useful arts, to extend education, to protect intellectual property, and to promote free trade" (Warsh, 2007). From China to France to the Americas, the state recognised the value of knowledge creation. In the Americas under Spanish control, schools and universities were created and it is well known that Napoleon did efforts to democratise it, from the *grandes écoles* to public schools (Chesnais, 1993). In the nineteenth century, German universities set the global standards for scientific excellence and it played a key role in its journey to become a powerful industrial nation (O'Boyle, 1983).

Nevertheless, it was only in the 1960s that knowledge and technology were formally included in economic debates, especially related with economic growth. During that time, the then existing theories emphasised that capital (physical), labour and other resources (factors of productions) with diminishing returns were the main drivers of growth (Kaur and Singh, 2016). Harrod (1939) and Domar (1946), for instance, argued that the economic growth of countries could be explained in terms of the level of national savings, the efficiency of their use and the productivity of capital investment. These elements enabled economists to study who should produce what goods, the existing work relationships, which responsibilities belonged to government and which were best left to markets, and so on (Warsh, 2007).

However, as Warsh (2007) argues, some circumstances in the human condition were simply taken for granted, for instance, the extent to which knowledge could influence productivity or the fact that human nature, expressed as tastes and preferences, influences market dynamics.

These ideas and models were further revised by Robert Solow (1957), who explained growth using three variables: capital and labour (as his predecessors), and "knowledge" or the "effectiveness of labour" (Romer, 2019). In this model, Solow assumed that ideas are a public good, although there is ambiguity regarding the source of technological progress, as Kaur and Singh (2016) point out. For Solow, ideas were considered an exogenous factor, meaning that they just accumulate over time. Many have argued that this is the main shortcoming of the model because the growth of new ideas and the creation of knowledge is endogenously produced by public or private investment in research and development Kaur and Singh (2016). Moreover, knowledge does not necessarily have the characteristics of the production factors in Solow's model and there was a sense that the model did not explain the sources of technological change and therefore left the underlying determinants of growth unaccounted for (Kugler, 2018).

By the time Robert Solow published his growth model, Trevor Swan (1956) independently developed a similar model. Many argued that Swan published a more complete analysis of technical progress, which Solow treated separately (Dimand and Spencer, 2009). However, the main difference between the two models is that Solow's diagrams focused on capital/labour ratio, whereas Swan's diagrams focused on the output/capital ratios and rates of growth (Dimand and Spencer, 2009). Solow (2007, p. 596) remarks that Swan's model "works exclusively in Cobb-Douglas³ terms. This allows an exposition entirely in terms of growth rates. Although this formulation fits in well with the literature at the time, it obscures the general-equilibrium character of the model. Between them, the limited generality and the preoccupation with growth rates may account for the fact that Swan's model of exposition did not catch on". Finally, many economists have argued that the Solow-Swan model was a very powerful technical tool to analyse long-term growth trends and perform accounting of how much each input contributes to output growth (Kugler, 2018). However, the productivity growth that resulted from technical change was ascribed to decisions and activities by scientists operating outside the scope of markets. This meant that there was no role for economic policy to directly impact long-term economic growth (Kugler, 2018).

³Cobb–Douglas production function is widely used to represent the technological relationship between the amounts of two or more inputs (particularly physical capital and labour) and the amount of output that can be produced by those inputs (Cobb and Douglas, 1928)

These shortcomings led some scholars to undertake the endeavour to bring "knowledge" into the growth model so that the role of technology, which had been treated as exogenous, was now endogenous and actually explained within the model. For example, George Akerlof and William D. Nordhaus (both Nobel Laureates, as Solow) proposed a model where they showed that "balanced growth in the presence of steady technological change may be, in a well-defined sense, a razor's edge case" (Akerlof and Nordhaus, 1967, p. 343). Later on, Nordhaus (1969) also worked on incorporating innovation and technological change into growth models. Kugler (2018) explains that part of the challenge that these scholars faced was "figuring out how to model, in a dynamic general equilibrium setting, the market power that firms must possess to cover R&D costs and thus be incentives to invest in innovation".

In the 1970s and 1980s, new efforts to address the shortcomings of Solow's models were developed and Endogenous Growth Theory was born. This theory consists of a set of models proposed mainly by Paul Romer (1990), based on previous work of Kenneth Arrow (1971), Rober Lucas Jr (1988) and others. These models internalised the role of knowledge and ideas as products of economic activity, arguing that technological change is indeed a result of intentional actions by people, through research and development (Peet and Hartwick, 2015), in contrast with Solow's model that assumed that knowledge creation and innovatio operated outside the scope of markets (Kugler, 2018). More specifically, Romer (1990, 1994) proposed a model of growth based on ideas, privately produced by for-profit firms in monopolistically competitive markets, which generate spillovers. Therefore, one of Romer's main accomplishments was to put "ideas" at the heart of growth theory, which open the growth models up to new venues of related research, such as the role of intellectual property, universities, capital markets (venture capitals), idea transmission, types of technology, agglomerations/density, among others.

In the 2000s, new research that complements Romer's ideas emerged. For instance, Oded Galor and Weil (2000) developed a unified growth model that captured the historical evolution of population, technology, and output. Charles Jones (1999) proposed a growth model that takes into account the production of ideas in the relation to population, the improvement of institutions that promote innovation and other factors. Gary Hansen and Prescott (2002), Robert Lucas (2002) or Galor et al. (2005) explored the causes of growth and technological change in the very long run.

Currently, a major topic within the study of economic growth is the role of technology. It is well recognised that technology differs across nations, whereby industries and firms are the main sources of productivity differences. Much of Daron Acemoglu's work on the direction of technological change can be interpreted as a second dimension of Romer's work (Jones,

2019). For instance, Acemoglu and Restrepo (2018) argue that ideas can augment either skilled or unskilled labour; Acemoglu (2003) asks if there is any reason why technical change would tend to be labour-augmenting instead of capital-augmenting. Hémous and Olsen (2014) and Acemoglu and Restrepo (2018), using an endogenous growth framework explore automation processes and its relation with labour. Also, connections between endogenous growth, technological change and international trade have been explored by Andrew Atkeson and Burstein (2010), Natalia Ramondo et al. (2016) and Costas Arkolakis et al. (2018). However, despite recent progress in modelling endogenous innovation and technology, it has been acknowledged that the forces shaping the diffusion of technology and innovation are still poorly understood.

2.4 Spatial characteristics and dynamics of knowledge and innovation

Much efforts have been made to understand how knowledge is created, why innovation takes place, how and where technology evolved. These questions have been studied in different fields beyond mainstream economics and with a wider variety of models beyond the traditional ones. Among these, a central concern of Economic Geography has been to understand why related economic activities or technologies agglomerate in a small number of locations (Delgado, 2018) and how these agglomerations are important for industrial and firm performance.

A pioneering scholar interested in this topic was Alfred Marshall (1920). In the late 1890's, he observed that industries specialise geographically and argued that this behaviour emerged because proximity favours the intra-industry transmission of knowledge. Moreover, he argued that industrial agglomeration happens because having closer supplier-buyer linkages, easy access to labour market as well as access to knowledge spillovers have a positive effect on the industries allocated in a given place (Marshall, 1920). Several contemporary scholars have shown and confirmed his findings and argues that the presence of related economic activity does matter for regional and industrial performance, including job creation, innovation and entrepreneurship (Feldman and Audretsch, 1999; Porter, 2003; Feser et al., 2008; Glaeser and Kerr, 2009; Delgado et al., 2014; Neffke et al., 2011; Balland et al., 2020b). Marshall's ideas were later complemented by Arrow and Nelson (1962) and Romer (1986), and formalised by Glaeser et al. (1992), giving rise to what is known as Marshall-Arrow-Romer (MAR) model. This model argues that the concentration of an industry in a region promotes knowledge spillovers between firms and facilitates innovation in that particular industry within that region (Glaeser et al., 1992; Ashraf et al., 2017). Furthermore, this model

explains that agglomeration not only promotes the transmission and exchange of knowledge, ideas, information, but also encourages the transmission and exchange of products and processes through imitation, business interactions and inter-firm circulation of skilled workers (Beaudry and Schiffauerova, 2009; Jara-Figueroa et al., 2018; Gomez-Lievano et al., 2018).

At the end of the 1990s, Feldman and Audretsch (1999) found compelling evidence that the specialisation of economic activities does not necessarily promote innovative output, despite previous claims. Their findings indicated that diversity across complementary economic activities sharing a common science base is more conducive to innovation than specialisation. In addition, their results suggested that location may provide a useful set of boundaries within which to organise innovation. In contrast with these findings, Glaeser and Kerr (2009); Diodato et al. (2018) argued that industrial diversity does not necessarily promote entry of new economic activities except through the "Marshallian channels"⁴ modelled. The clustering of related industries matter, but not diversity per se.

Following this line of research, Michael Porter (2003), Mercedes Delgado et al. (2010), Dario Diodato et al. (2018) found that industries located in regions with strong clusters (i.e. a large presence of other related industries) experience higher growth in new business formation and start-up employment. Similarly, in a study that uses a comprehensive set of industries, Delgado et al. (2014) found that industries participating in clusters with dual strength in innovation (patenting) and production (employment) grew faster in terms of innovation, as opposed to those that are not part of a cluster. Their research suggests that the "co-location of innovation and production matters for subsequent innovation and facilitates a broad set of linkages, including input-output linkages and knowledge spillovers" (Delgado et al., 2014, p. 332). Moreover, the authors argue that regional innovation is better understood in the context of co-located industries and firms that are connected by many types of linkages (Marshall, 1920; Porter, 2008; Delgado et al., 2015; Delgado, 2020; Balland et al., 2020a). Therefore, there is a strong relation between geographical concentrations of economic activities (firms or industries) and the production of new ideas and innovation (Balland et al., 2020a,b; Delgado, 2020).

In a similar line of work, Hidalgo et al. (2007); Hidalgo and Hausmann (2009) mapped the relations that exist between the products traded in the global economy and how regions and countries have specialised in the production of specific goods. This work showed (and confirming with previous research) that there is a strong tendency of the productive mix of a region to move from current products or industries towards related ones, rather than goods

⁴According to Alfred Marshall, there are three types of channels that foster agglomerations: large pool of skilled labor, easy access to local customers or suppliers and local knowledge spillovers (Diodato et al., 2018)

that are less related (Hausmann et al., 2007b; Boschma and Frenken, 2011; Boschma, 2012). In other words, there is a strong association between the number of **related activities that are already present in a given location** (relatedness by co-location) and the likelihood of that location to develop new industries. In later work, Boschma (2012, 2017); Hidalgo et al. (2018) explained that this co-location relatedness is relevant because the particular knowledge base of a region acts as an incentive for new industries to appear, because it not only offers local firms opportunities to further improve in familiar fields of knowledge, but also discourages knowledge creation that does not fit in the regional knowledge base.

This “Principle of Relatedness” has been explored and corroborated at multiple spatial scales, for different economic activities and diverse institutional regimes. It has been observed, for example, in analysis of entry and exit of products (Neffke et al., 2011; Zhu et al., 2017; Gao et al., 2017), industries (Hidalgo et al., 2007; Alshamsi et al., 2018), occupations (Muneepeerakul et al., 2013), and technologies (Kogler et al., 2013; Boschma et al., 2014), at the national, regional, and metropolitan scales (Hidalgo et al., 2018). In 2018, a new research that analysed 93 countries and their diversification paths, confirmed with previous findings, also showed that depending on the level of development, different countries diversified into more or less related economic activities than the ones that already existed in a given location (Pinheiro et al., 2018; Alshamsi et al., 2018).

In a different but related strand of literature, Comin et al. (2012) studied technology diffusion over time and found that technology diffuses more slowly to locations that are further away from technology leaders. In this study, the authors find that technological knowledge is likely to be more easily transmitted between agents in close places than between agents located far apart. Similarly, Keller and Yeaple (2013) found that the payoff to adopting a given technology may be affected by the adoption experience of neighbouring countries. By measuring cross country spatial barriers to the transmission of embodied or disembodied knowledge, the authors found that person-to-person communication costs increase with distance. The work of these authors suggests that the more knowledge intensive a production process is, the less likely its knowledge will spatially diffuse, which has important implications for cross-country income convergence.

As it can be observed, the aforementioned literature provides spatiality to the economic processes. Knowledge creation, combination and accumulation are bounded geographically, which explains why certain places develop and specialise in specific economic activities. In addition, these knowledge dynamics rely heavily on the pre-existing activities of a place, which make these processes heavily path- and place-dependent. These characteristics are

key to understand why countries have certain productive structures and why the technologies developed in certain countries foster their economic growth, as previously explained in this chapter. In this sense, when one talks about economic structures, as in old and new structuralism, we are talking about structures that are heavily linked with a space (country, city, region, and so on). Structuralism needs geography in order to understand these structures and their evolution over time. Therefore, this geographical approach is a key component to address the last objective of this thesis, which relates to economic diversification of an oil exporting economy. Evolutionary Economic Geography (EEG), which offers an empirical treatment of longitudinal patterns of regional economic evolution provides a theoretical framework to land structuralist thinking in a space. The relevant EEG literature as well as the contributions of the thesis will be further explained in detail in Chapter 5.

2.5 Literature Summary

As discussed in this literature review, economic growth has been a priority as a way to achieve the development. Economic growth is linked with the idea of material progress because it provides some of the means necessary to achieve other development objectives (if well managed) such as poverty alleviation, inequality reduction, betterment of health and education services, among others. However, achieving this is a complex task, which relies on an interplay of factors such as luck, geography, institutions, culture, among others.

Among the aforementioned components, this thesis focuses mainly on the geographical one, more specifically, the role that crude oil extraction has played in the development progress of Ecuador. To do so, this literature review began with exploring the structuralist literature in Economics. This tradition centres around the idea that countries have specific ‘structures’, which could foster or constrain economic growth. These structures are the result of historical (endogenous and exogenous) processes and are a crucial factor that drives development paths of countries. This literature pays specific attention to productive structures and argues that are sectors/industries that are key to promote growth. This section discussed the differences and similarities between the traditional and new structural thinking, as well as their advantages and shortcomings. In Chapter 3, this thesis proposed a new methodology to identify and explore the productive structure of an oil exporting economy, attempts to identify these key economic sectors in Ecuador and looks closely at the role that the oil sector plays in its economy. Chapter 4 links the structuralist literature with the resource curse debates. Its main objective is to explore the effect that a windfall of oil resources has on the productive structure of Ecuador.

Section 2.3 explores the connection that exists between economic growth and technology / knowledge. This literature argues that innovation and technological progress are key factors for growth - besides capital and labour - and is based on the fact that ideas, knowledge or technological change can augment either skilled or unskilled labour. This literature also recognises that technology differs across nations, which has an influence on their growth and argues that the main sources of their productivity differences are their industries and firms. Therefore, the technology that exists within countries is one of the main factors that configures their economic structures, as the structuralist literature argues, that could constrain or foster countries' growth and may signal certain development paths. In this sense, Chapter 4 explores the structural impact that a new technology and knowledge associated (the oil industry) had on the economic structure of Ecuador.

The last section of this review gives a spatiality to the previous literature and argues that all these structures and technologies are indeed associated with a physical space. Therefore, to study structural technological change (known as productive diversification), we need to first understand where these industries and supply chains are located and why there emerged in certain regions. The Economic Geography literature argues that these technological processes are heavily path- and place- dependent and the "Relatedness" literature, which is part of the Evolutionary Economic Geography literature, has shown that new industries or technologies emerge based on the density of related economic activities that exist in a given place. Chapter 5 draws from this literature and explores the emergence of new economic activities in all the counties of Ecuador, based on the co-location of firms and the location of the production chains.

The elements covered in this chapter provide the theoretical framework from which the rest of the thesis draws. In the following empirical chapters, we will find more detailed literature reviews in specific chapters. In sum, Chapter 3 explores in detail the structuralist literature and how the productive structure of a country could constrain its growth. Also, this chapter explores the role of key industries to foster growth. Chapter 4 builds on the previous literature and links it with the Resource Curse and Dutch Disease research. This sets the foundation for understanding the structural impact of a resource boom. Finally, Chapter 5 examines in depth the literature of knowledge creation and regional diversification in Economic Geography, and constitutes the foundation to understand how an oil exporting economy could diversify.

Chapter 3

Assessing the Productive Structure of an Oil Exporting Country

3.1 Introduction

The problems that oil exporting and other resource-dependent countries face are not new. Much research efforts have tried to understand: the vulnerabilities these countries have to macroeconomic shocks, usually caused by the volatility of commodity prices; their tendency to under-perform economically in comparison with non-resource rich countries; and the constraints that prevent them from productive diversification (Gelb, 2011; Lederman and Maloney, 2006). The above research is often grouped under the so-called "Resource Curse" hypothesis¹, which argues that natural resource dependency is correlated with low economic growth, weak institutions, lack of human capital, among others issues (Ahmadov, 2014; Auty, 1994, 2001, 2002; Auty and Pontara, 2008; Bornhorst et al., 2008; Gelb and Grasmann, 2010; Gelb, 2011; Humphreys et al., 2007; Lederman and Maloney, 2006; Murshed and Serino, 2011; Papyrakis, 2017; Sachs and Warner, 1995, 2001; Stevens et al., 2015; Sy et al., 2011).

Recent research has shown that the resource curse hypothesis is far from being conclusive (Auty and Furlonge, 2019; Ross, 2018; Papyrakis, 2017; Auzer, 2017; Stevens et al., 2015). The resource curse is in fact a much more complex phenomena and the previous attempts to generalise the relationship between resource dependency with the aforementioned factors have been largely unfruitful (Stevens et al., 2015; Papyrakis, 2017; Auzer, 2017). Most importantly, some countries experience the "curse" in different orders of magnitude whereas others do not seem to experience it at all (Auzer, 2017; Papyrakis and Pellegrini, 2019). Currently, the existing consensus among scholars is that studying the correlation between

¹This was briefly explained in the introduction of the thesis and will be covered in more detail in the following sections

economic performance and resource dependency must be dealt with on an individual basis, since its manifestation depends on factors that are unique to specific regions or countries (Papyrakis, 2017; Papyrakis and Pellegrini, 2019).

One way to study the unique features of a country’s economy is through the lens of structuralism, the ideas of which were originally developed in the late 1940s and states that countries have distinctive structures that impose constraints on growth for developing nations (Chenery, 1975; Prebisch, 1950; Singer, 1950a; Hirschman, 1980; Hausmann et al., 2014; Hidalgo et al., 2007; Ocampo et al., 2009). It argues that development occurs when there are changes in the productive structure of an economy and economies advance when such changes move towards ‘modern’ sectors (industrialised and technology-intensive (Misio et al., 2015; Lin, 2017; Diao et al., 2017; McMillan et al., 2017)). However, limited tools were available to explore the productive structure analytically (Hartmann et al., 2017a).

With ever-improving data and computational capacity, in the last two decades a new set of analytical tools have been developed for understanding these economic structures. These tools are based on graph theory (network science) and have shown that productive structures - measured in terms of products produced and exported, for instance - do indeed provide a good approximation of the economic growth or health of institutions in a given country and can be critical for understanding social welfare, inequality and poverty (Milanovic, 2013; Hausmann et al., 2014; Hartmann et al., 2017a,b). Further, this research has shown that countries’ productive structures are not only a determinant for economic growth, but also a good approximation of the knowledge and know-how available in an country, which could indicate future patterns of development and product diversification (Hausmann et al., 2014; Hausmann and Hidalgo, 2011; Hidalgo et al., 2007; Hidalgo, 2015).

Despite the recent scholarly advances, much of the research that focused on productive structures use international macroeconomic and trade data, for generalisability. In addition, previous studies have not looked at how resource-rich economies are configured internally nor investigated the interactions that exist between industries and products. Further, current research and their methodologies have focused on places where data are rich and easily accessible and excluded countries (mostly developing ones) where historical data is difficult to find or is incomplete. This chapter attempts to fill in these gaps.

3.1.1 Objectives

In order to overcome some of the shortcomings stated previously, the main objective of this chapter is to propose a network-based methodology that allows the study of the productive structure of a resource-rich economy, which aims to address the first research objective of

this thesis. More specifically, it aims to develop a methodology to assess the importance of different industries of a country based on their inter-industrial relationships, which also allows to examine what is the structural role of the crude oil industry.

Traditionally, Input-Output (I/O) analysis has been a good tool for identifying key sectors based on the linkages between industries. A key sector would have a high value in what is called an output multiplier. However, this kind of identification cannot be performed when data are incomplete, especially in cases where only supply or *use tables*² are available³. The proposed methodology, which will be further explained below, attempts to identify key sectors by using the tools developed by network science. Specifically it calculates the network centrality for different industries and products, based on national *use tables* to identify their structural position in a given economy. This methodology is tested using Ecuador as a case study. As explained previously, the country was selected given the relatively low complexity of its economy as well as the availability and accessibility of consistent Input/Output data for 30 consecutive years (1965-1995). The information to be used was compiled and validated by the United Nations Development Program and the Central Bank of Ecuador.

This chapter is the foundation of this thesis. Its first section will explore in detail the existing academic discussions about structuralism and economic growth, as well as the literature that links the role of key industries in achieving economic development. As discussed in the previous chapter, the productive structure of a country is one of the main factors that affects its economic growth. As such, understanding the productive configuration of an oil exporting economy could generate insights that may help explain their growth constraints. The second section of this chapter describes the data used for the empirical analysis of this investigation and it proposes a network-based methodology to extract the structural information of an economy. In the results section, the findings will be summarised with a special focus on the oil industry and its importance in the context of Ecuador. The findings will also be contrasted with traditional aggregate measurements and it will emphasise the differences between the traditional approach to understand the productivity of a country and the structural approach developed in this chapter. In addition, a discussion about how the proposed approach may or may not complement the ways in which economic systems are analysed. Finally, several conclusions are drawn, as well as the limitation of

²A use table shows where and how goods and services are used in the economy(UNSTAT, 2018).

³The methodologies used to study input-output relationships before the 1980s were not the I/O tables that are known today. According to the data examined for this thesis, there were only *use tables*, which were incomplete to build a current I/O table.

the proposed methodology. It is important to note that this methodology, the results and underlying ideas will be the foundation for the coming chapters.

3.2 Literature Review

3.2.1 Economic structures, economic growth and development

One way to study the unique features of an economy is through the lens of the structuralist tradition in Economics, which was originally developed in the late 1940s. Economists such as H. Singer, P. Prebisch and A. Hirschman argued that productive structures have specific characteristics that constrain the economic performance of countries or regions (Chenery, 1975; Prebisch, 1950; Singer, 1950a,b). They argued that those structures could explain phenomena such as balance of payments disequilibrium, the limited ability of economies to absorb workforce, the vulnerability to external shocks, or even unequal income distribution (Chenery, 1975). These structuralist thinkers promoted the idea that as a way to foster economic growth, nations should progressively change their imports and internal production structure (structural change) and promote import substitution industrialisation (ISI) (Prebisch, 1950; Singer, 1950a,b; Harvey et al., 2010a). Years later, Hollis Chenery argued that economic development itself can be seen as a set of interrelated structural changes (Chenery, 1975, 1979). However, some type of this ‘structural’ thinking could also be traced back to the work done by Joseph Schumpeter on creative destruction and how this resembled a process of ‘industrial mutation’ that changed economic structures (Schumpeter, 2013).

New structuralist economists, building on the structuralist tradition developed in the 1940s, have argued that economic development is a process of structural transformation, which involves technological innovation and industrial upgrading (Lin, 2017; Diao et al., 2017; McMillan et al., 2017). They stressed, as their predecessors, that industrial structures need to be upgraded in order to achieve economic growth and development, particularly in developing countries (Lin, 2010). Their ideas of structural transformations refer to a reallocation of economic activity across three broad sectors (agriculture, manufacturing, and services); and these transformations have the potential to create new jobs and increase labour productivity (Lin, 2010; Milanovic, 2013). Furthermore, these transformations are usually accompanied by improvements in infrastructure and institutions (Lin, 2017).

Alongside this argument, some development economists like Dani Rodrik, Ha-Joon Chang or Justin Lin pointed out that the success of countries such as China or South Korea has to do with the transformation of their productive structure rather than the role of free markets or ‘getting the price right’. They also argue strongly that the configuration of an economy

is crucial for the development of a country (Rodrik, 2006; Lin and Chang, 2009; Chang, 2002).

Recent research has confirmed that countries' productive structures are not only a determinant for economic growth, but also a good approximation of the knowledge and know-how available in an economy (Milanovic, 2013; Hidalgo et al., 2007; Hidalgo and Hausmann, 2009; Hausmann et al., 2014; Hidalgo, 2015). The ability of a country to produce sophisticated and diverse products depends on the productive knowledge that its population has and how people form social and professional networks to use this knowledge and combine it (Hartmann et al., 2017a). As such, the economic output and respective growth of countries whose knowledge is limited to the extraction of few natural resources would often be lower than in economies with a more diversified knowledge base and structure (Hartmann, 2014; Hausmann et al., 2014; Hidalgo, 2015).

Further, research has shown that these productive structures are also a proxy for the health of countries' institutions and critical for understanding social welfare, inequality and poverty (Hartmann et al., 2017a). Moreover, how production is structured in a country can constrain occupational choices, learning opportunities and even bargaining power of workers and unions (Hartmann et al., 2017a). Therefore modifying them through technological upgrade could provide new jobs and learning opportunities to workers in emerging economies (Milanovic, 2013). In addition, changing these structures could have a significant effect on the socio-economic progress of a given country. Finally, Lin (2010) and others have argued that industrial structures need to be upgraded in order to foster growth, which requires an upgrading of their endowment structures.

3.2.2 Understanding productive structures

Several efforts have been made to quantify productive structures. For instance, Wassily Leontief developed the input/output models (IO), which focuses on quantifying the fraction of the output of an economy used by different sectors or products and their corresponding inter-linkages and interactions (Leontief, 1986; Mirowski, 1991). This approach has been useful because it allows us to capture the inter-industrial relationships in an economy (Noguera-Méndez et al., 2016). However, doing this kind of analysis in some developing countries has technical challenges, because the data found in archives named as IO are usually *use tables*, which captured intermediate consumption and thus only contained part of the information required to undergo traditional IO analysis.

Meanwhile, John M. Keynes and other Keynesian economists developed the idea of national accounting to quantify the output, expenditure, and income activities of economic actors in

an economy, tracking their corresponding relationships with other countries' (Tily, 2009). Other economists such as Orris C. Herfindahl and Albert O. Hirschman proposed different aggregate measurements to quantify the diversity and concentration of exports by looking at trade between nations (Herfindahl, 1950; Hirschman, 1980; Imbs and Wacziarg, 2003).

With ever-improving data and computational capacity, a whole independent field of network emerged. Its tools, based on graph theory, have become key for analysing complex systems such as biological or social systems (Caldarelli, 2007; Barabasi, 2016; Jackson, 2010). In the last few years, network science has been applied to analyse economic systems and economic behaviours, ranging from the dynamics of product adoption to financial contagions, where it is necessary to account for the patterns of interactions (Kirman, 2016; Barabasi, 2016; Jackson, 2010; Haldane and May, 2011; Caldarelli et al., 2012; Beinhocker, 2006; Acemoglu et al., 2012).

This network approach has enabled us to quantify these productive structures and has renewed our interest in the macroeconomic role of structural transformation (Hartmann et al., 2017a). Further, it is relevant for understanding social welfare and inequality (Bourguignon, 2004; Ravallion, 2004; Hartmann et al., 2017a,b).

The economic complexity approach, based on network science, has gained much attention as its measurements are highly predictive of future growth, compared with other approaches developed up to date (Abdon and Felipe, 2011; Bustos et al., 2012; Caldarelli et al., 2012; Cristelli et al., 2013, 2015; Felipe, 2012; Hausmann et al., 2007a, 2014; Hidalgo and Hausmann, 2009; Hidalgo et al., 2007; Rodrik, 2006). This approach analyses the networks of relatedness between products, which allows to understand the composition of a country's productive output. It shows that the mix of products that a country produces has strong connections with its level of economic growth (Hausmann and Hidalgo, 2011; Hidalgo et al., 2007; Hidalgo and Hausmann, 2009; Hausmann et al., 2014; Hartmann et al., 2017a). This approach has the potential to predict future patterns of diversification or specialisation (Hartmann et al., 2017b,a).

Despite the above advancements, much of the research measuring productive structures mostly use international macroeconomic and trade data, for generalisability. In addition, previous studies have not looked at how resource-rich economies are configured internally nor investigated the interactions that exist between industries and products. Further, current research and their methodologies have focused on places where data are rich and easily accessible, excluding countries (mostly developing ones) where historical data is difficult to find or is incomplete.

3.2.3 Key economic sectors and growth

A way to explore the internal economic structure of a country is to understand the role that individual sectors and products play, as well as its corresponding relationships and dynamics over time. Noguera-Méndez et al. (2016) argues that changes on these inter-industrial relationships could affect the economic system as a whole and modifying their structural characteristics (density, centrality and network topology), as well as the position of the different actors (sectors or products), can have an important effect on the economic performance of a country (Noguera-Méndez et al., 2016).

Several scholars have argued that the way through which sectors and products are connected and their inter-dependency is a key characteristic of modern production; therefore how they interact are linked with the rest of the economy are crucial factors affecting economic growth (Meller and Marfan, 1981; Hausmann and Hidalgo, 2011; Hidalgo and Hausmann, 2009; Darwent, 1969; Albeaik et al., 2017). Meller and Marfan (1981) explain that these inter-dependencies indicate the capacity that sectors possess to stimulate others (Meller and Marfan, 1981). Moreover, Noguera-Méndez et al. (2016) explains that sectors with strong backwards and forwards linkages have the capacity to influence and expand others. Therefore, to identify key sectors, we cannot solely focus on whether a particular sector has a higher or lower weight in production, or whether its weight has been reduced or increased. Instead, we should find the structural positions of the sectors and systems that shape their relations, the complexity of the chains and production systems as well as the diversity of productive structures (Noguera-Méndez et al., 2016).

In accordance with the previous arguments, Kubo et al. (1986) conclude that there is a positive relationship between development and the interconnectivity of sectors. For instance, Dutt (2019) explains that Albert Hirschman favoured growth strategies that made use of the market linkage effect in key industries (especially backward linkage effects from input-using industries), to induce pressures that stimulate other industries. In their findings, the authors argue that a higher concentration of manufacture goods, which have the biggest consumers of intermediate inputs, are key sectors for an economy. This statement is coherent with the findings of Schultz (1977), where countries with high per capita incomes tend to have a high degree of division of labour, which is expressed by a dense network of intermediate supplies or deliveries. Finally, those sectors of the economy which, owing to their close technology-related ties, are in a position to promote or generate growth in other sectors through that of their own. This can be deemed strategic for achieving higher income levels (Schultz, 1977).

All these previous measurements are in line with the latest economic complexity research, where the product mix that an economy exports and how each product is linked and interacts with one another is highly correlated with economic performance and is one of the best existing predictors for economic growth (Albeaik et al., 2017; Hidalgo and Hausmann, 2009; Hausmann and Hidalgo, 2011).

3.2.4 Identifying key sectors and products

In the fields of development and regional economics, one way to identify key sectors - defined as sectors with a high potential of spreading growth impulses throughout the economy (Miller and Blair, 2009) - is by measuring the size of their linkages. This is based on the consideration that a sector with a stronger linkage can stimulate other sectors better, therefore promoting growth. Several methodological approaches have been developed to identify these sectors, examples being the replacement of direct backward linkages with the total backward linkages measured by the column sums of the Leontief-inverse (Rasmussen, 1956; Chenery and Watanabe, 1958), or the replacement of the row sums of the Leontief-inverse ⁴ with the row sums of the Ghosh-inverse in the case of total forward linkages ((Beyers, 1976; Jones, 1976; Rasmussen, 1956).

While traditional IO models have done a good job capturing inter-industrial relationships, they have not been able to capture the systemic nature of the complexity of economic structures (Noguera-Méndez et al., 2016), (Carvalho and Voigtländer, 2014) and (Kubo et al., 1986). For instance, IO analysis uses weighted indicators of the inverse matrix coefficients to measure structural change. However, Noguera-Méndez et al. (2016) explain that this is not enough to understand the complexity and the relationships between sectors. They argue that applying network analysis into the IO data offers relational information about the intermediate commercial relationships between the sectors that constitute an IO network, allowing us to perform structural analyses. In an IO table, the sector information could be used to generate a network that shows the changes on development processes, which can be analysed with network tools (Noguera-Méndez et al., 2016).

As previously discussed, despite the above advancements, much of the research measuring productive structures mostly use international macroeconomic and trade data, for generalisability. It was also stated that previous studies have not looked at how resource-rich economies are configured internally nor investigated the interactions that exist between industries and products. To address this gap in the literature, this chapter takes Ecuador as

⁴According to EUSTAT (2020), "the Leontief Inverse Matrix shows the coefficients (economic multipliers) that measure the successive effects on the economy as a result of the initial increase in production of an economic activity branch."

a study case and looks specifically at the configuration of its domestic productive structure, focusing mainly on its Input/Output and intermediate consumption data to better understand the relationships that exists between its industries. Furthermore, it was explained that existing research and their methodologies have focused on places where data are rich and easily accessible, excluding countries (mostly developing ones) where historical data is difficult to find or is incomplete. In this sense, to address these challenges, I spent three months on field work, looking for, obtaining, extracting and digitising the National Accounts of Ecuador (more specifically the I/O tables) from 1965 to 1995 from the historical archives of the Ecuadorian Central Bank. These data are key to explore how the economic structure of the Ecuadorian economy have evolved over time and to study the role of the oil sector in particular.

3.3 Methodology and Data

3.3.1 Data

To start the analysis, we first defined the ‘productive structure’ as the network of production and demand linkages that exist between products and sectors in an economy, which are captured by the relationships in the numbers that appear in the nationals, trade, and other accounts (Ocampo et al., 2009; Hirschman, 1958). In other words, it refers to the linkages and corresponding interdependency of the different components of a given economy, which are at different levels of aggregation, as summarised in table 3.1.

Table 3.1: Scale of Analysis

Level of aggregation	Scale
Lowest	Individual
Lower	Firm
Intermediate	Industry
Higher	Country

As it can be seen, the IO data are considered as an intermediate level of aggregation, which incorporates the flows of inputs across sectors in an economy. However, when exploring the IO information of Ecuador, which is recorded on its National Accounts (BCE, 1981, 1982, 1983, 1989), I realised that the data labelled as IO was in fact *use tables*. These tables contain product-industry intermediate consumption data of 31 sectors for the period 1965-1995, which includes the year of the oil boom (1972). The recorded intermediate

consumption is valued at purchaser prices (EuroStat, 2016) in Ecuadorian Sucres⁵ from 1972 (BCE, 1981). The table 3.2 indexes the national accounts used for this research.

Table 3.2: National Accounts - Index

National Accounts (# volume)	Publication year	Years covered
1	1981	1972-1978
2	1982	1965-1971
3	1982	1972-1980
4	1983	1974
5	1983	1973-1981
6	1984	1974-1983 (p)
7	1985	1980-1984(p)
8	1986	1981-1985
9	1987	1982-1986
10	1987	-
11	1988	1983-1987
12	1989	1984-1988
13	1990	1986-1989
14	1991	1987-1990
15	1992	1989-1991
16	1993	1989-1992
17	1994	1990-1993
18	1996	1992-1995

The information contained in the Ecuadorian *use tables* do not include fixed assets, whose consumption is recorded as consumption of fixed capital. Nevertheless, the information provides valuable insights of the goods and services consumed as inputs by the production processes.

Table 3.3: Use Table example

Industries Products	Sectors			Final uses			Uses
	Agriculture	Industry	Services	Final consumption	Gross capital formation	Exports	
Agricultural goods	34.00	59.00	1430.00	81.00	21.00	32.00	370.00
Industrial goods	106.00	119.00	77.00	123.00	103.00	62.00	590.00
Services	70.00	112.00	75.00	291.00	61.00	31.00	640.00
Total	210.00	290.00	295.00	495.00	185.00	125.00	1600.00
Value added	90.00	210.00	405.00				705.00
Output	300.00	500.00	700.00				1500.00

Data Cleaning

Before proceeding further, the data set must be balanced. A balanced data set contains observations of multiple phenomena obtained over multiple time periods for the same firms

⁵Currency used in Ecuador until 2000, when it was replaced by the US Dollar.

or individuals. This is important for the construction of networks, as well as the panel analysis to be done in the next chapter.

When we explored existing data, we found a few inconsistencies between the tables prior to 1979. According to the methodological documents that accompany the tables (BCE, 1981, 1982, 1983, 1989), the sectors corresponding to machinery and other manufactured products (before 1978) were counted as one sector. To overcome this issue, we split these sectors proportionally to the corresponding values for the years after 1979, according to the base format used for the period 1979 - 1995.

Additionally, in order to get consistency in the data, some sectors were re-ordered and their identification number (Id) was changed. In table 3.4, we can observe the changes highlighted in yellow. Oil refining, which had an ID 17 prior to 1978, was replaced with ID 7 and the rest of items were modified accordingly.

Table 3.4: Sectors and Products in the Ecuadorian Economy

1965-1978		1979-1995	
Id	Label	Id	Label
1	Banana coffee cocoa	1	Banana coffee cocoa
2	Other agricultural products	2	Other agricultural products
3	Animal products	3	Animal products
4	Forestry products	4	Forestry products
5	Hunting and fishing products	5	Hunting and fishing products
6	Oil and natural gas	6	Oil and natural gas
7	Other mining products	7	Oil refining products
8	Meat and processed fish	8	Other mining products
9	Cereals and bakery	9	Meat and processed fish
10	Sugar	10	Cereals and bakery
11	Various foodstuffs	11	Sugar
12	Drinks	12	Various foodstuffs
13	Tobacco	13	Drinks
14	Textiles clothing and leather products	14	Tobacco
15	Wood	15	Textiles clothing and leather products
16	Paper and printing	16	Wood
17	Oil refining products	17	Paper and printing
18	Chemicals plastics and rubber products	18	Chemicals plastics and rubber products
19	Basic mineral products	19	Basic mineral products
20	Machinery transport equipment	20	Machinery transport equipment
21	Electricity gas and water	21	Other manufactured products
22	Construction	22	Electricity gas and water
23	Commerce	23	Construction
24	Transport	24	Commerce
25	Communication	25	Transport
26	Financial services	26	Communication
27	Rental housing	27	Financial services
28	Business services	28	Rental housing
29	Hotels bars and restaurants	29	Business services
30	Household services	30	Hotels bars and restaurants
31	-	31	Household services

3.3.2 Methodology

As previously explained, the above Use Table data do not allow us to undertake the IO analysis. Here we propose a methodology based on network science, undertaking the IO analysis.

First, we make use of our *use tables*. A use table represents a relationship between products i and sectors j , which resembles a bipartite network or a two-node network (Figure 3.1). This kind of networks has two types of nodes, one representing the original node and the other representing the groups to which they belong. In the current case, if n is the number of products in the network and g is the number of sectors (industries), then the incidence matrix B is a $G \times n$ matrix having B_{ij} such at: $B_{ij} = 1$ if j belongs to group i and 0 otherwise. However, since the networks of analysis are *weighted*, instead of $B_{ij} = 1$, B_{ij} is monetary amount (measured in Ecuadorian Sucres form 1972) that a sector had used from a product in a given year. In other words, the links or edges of our networks could be interpreted as the monetary flows that had moved from a node i (products) to a node j (sector).

Table 3.5 shows how the network is structured.

Table 3.5: Use Table as a bipartite network

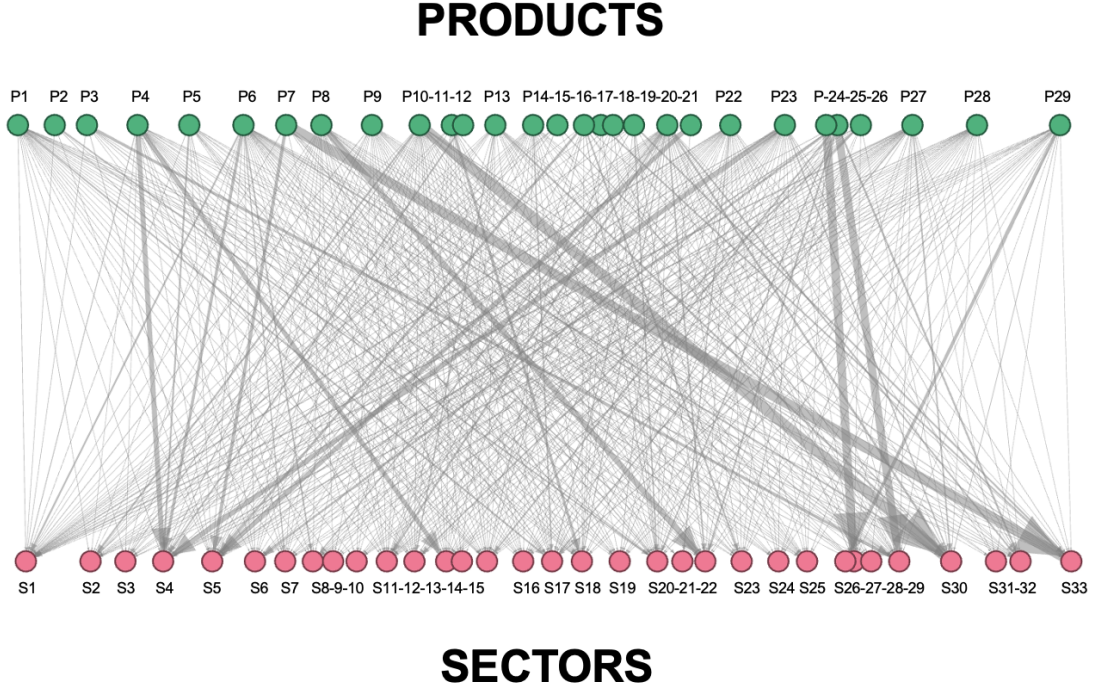
		Industries				
		1	2	3	...	30
Products	1 - Banana coffee cocoa	0	0	0	...	32
	2 - Other agricultural products	0	724	367	...	964
	3 - Animal products	0	0	135	...	1122
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	30 - Hotels, bars and restaurants	0	1	0	...	23

Florian Blöchl et al. (2011) explain that in complex network theory, identifying key sectors and ranking their roles in an economy entails applying an appropriate measure of node centrality. There are a number of measures of centrality in use, the most commonly used being degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality. The suitability of a centrality definition depends on the kind of network being considered. In this chapter, degree, weighted degree, and eigenvector centralities will be used and their interpretation in the context of our analysis will be further explained:

Centrality measurements

Degree Centrality A key property of a node in a network is its *degree*, which represent the number of links it has to other nodes. In other words, the degree centrality is the total

Figure 3.1: Use table as a bipartite network



number of links that a node i has. This is expressed mathematically as:

$$S_i = C_i^A = \sum_{j:j \neq i}^N A_{i,j}$$

where, i is the focal node and j represents all other nodes. N is the total number of nodes and A is the adjacency matrix, in which A_{ij} is equal to 1 if the node i is connected to node j (Barabasi, 2016). This quantity is a natural measure of the importance or centrality of a node i in the network. This measurement is the most intuitive topological measure of centrality: more connected nodes are more central.

In the case of our bipartite network, the degree centrality of a sector is the total number of products that are associated with it. In the case of a product, the degree centrality is the number of sectors associated with it. So, the highest degree centrality of a node means that the sector or product is the most connected entity within the network.

Weighted Degree Centrality Simply knowing how well the nodes are connected (degree centrality) is not enough to understand the complexity of a network. The weights of the

nodes are also important (Opsahl et al., 2010). The weighted centrality measures the strength of nodes in terms of the total weight of their connections (Opsahl et al., 2010; Barrat et al., 2004; Newman, 2004). In this case, the node strength S_i accounts for the number of connections of the node, weighed by the total cash in/outflows handled by each product or sector. In other words, this measure is the degree centrality weighted by the magnitude of their cash flows. This is expressed mathematically as:

$$S_i = C_i^w = \sum_{j:j \neq i}^N w_{i,j}$$

where, i is the focal node and j represents all other nodes. N is the total number of nodes and w is the weighted adjacency matrix, in which w_{ij} is greater than 0 if the node i is connected to node j , and the value represents the weight of the tie (Opsahl et al., 2010).

In our network, this means that the importance of a product (measured by its weighted centrality) represents the total number of sectors associated with it, weighted by the amount of product (measured in monetary terms) that flows into those sectors. On the other hand, the importance of a sector is the total number of products associated with it, weighted by the amount of products (measured in monetary terms) that the sector uses.

Eigenvector Centrality Calculating the Eigenvector Centrality is a natural extension of the degree centrality. In the previous case, the centrality of a node (sector or product) is measured based on the number of links that are connected with them and their respective weights. In the case of the eigenvector centrality, it evaluates the importance of a node if it is linked to other important nodes (Newman, 2010; Barabasi, 2016). In other words, Newman (2010) explains that instead of awarding vertices just one point for each neighbour [as in the case of degree centrality], eigenvector centrality gives each vertex a score proportional to the sum of the scores of its neighbours”.

$$C_i^e = \frac{1}{k_1} \sum_{j:j \neq i} w_{i,j} C_j^e$$

where, k_1 is the largest eigenvalue of w .

The aforementioned measures were calculated for our bipartite network, as well as, the average and the median of each centrality measure for the period of our analysis. This allows us to assess how relevant each sector and product is in the economy between 1965 and 1995, and gives us a clear view of how the economic structure is configured (Barrat et al.,

2004). Further, as a way to interpret the results obtained, we normalised and averaged the degrees. Lastly, the results above were compared with their absolute importance - measured in aggregate terms - in the Ecuadorian economy. This provides us with contrasting ways to understand economic growth for Ecuador. The next section presents and discusses the results derived from applying this methodology.

In addition, Newman (2010) explains that a bipartite network may give the most complete representation of a particular network but recommends working with direct connections between nodes of the same type. Thus, to extend the network analysis, I did a one-node projection on our *use table* data so that the nodes are of the same type. In *Appendix A*, I projected the bipartite network and explored the previous centrality measures. This exercise allowed us to double check the centralities of the nodes previously calculated.

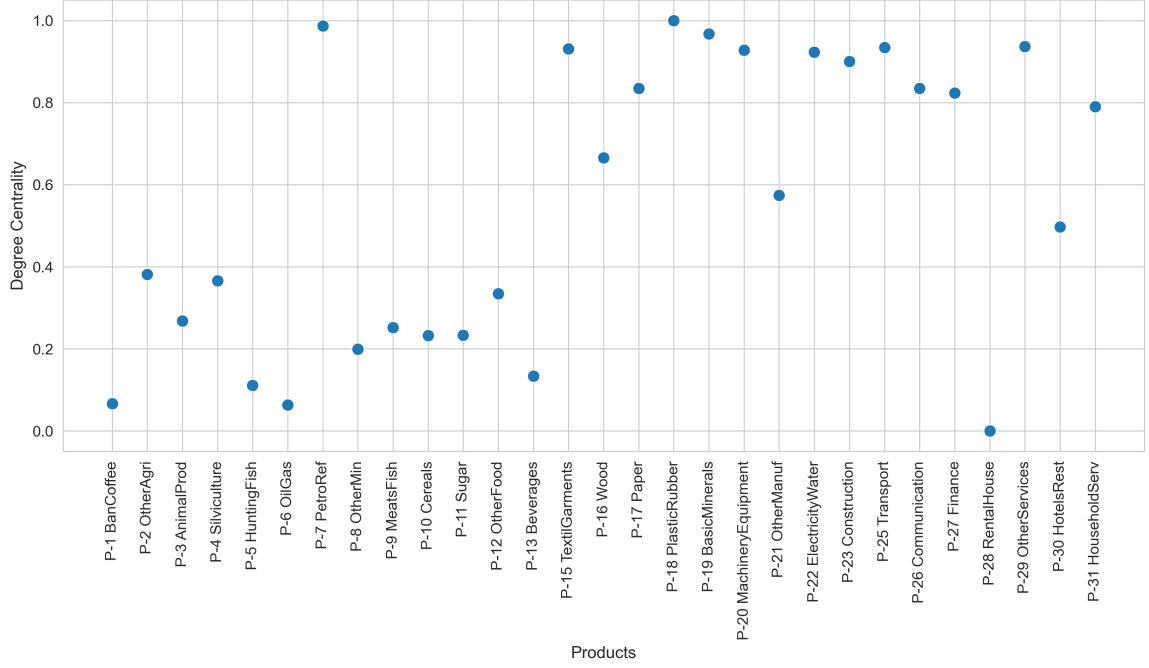
3.4 Results and Discussion

3.4.1 Key Products by Centrality

In the results summarised in the figure 3.2, there is a clear distinction between: primary products or raw materials (P1-P6), basic secondary products (P7-P13), more complex secondary products (P14-P21), and services (P22-P31). Primary products have the lowest degree centrality in relation to the other products, which means that the primary products have the lowest connectivity with the rest of the Ecuadorian economy. These results are coherent with the structuralist literature, in the sense that the raw materials products have the lowest number of linkages in comparison with other products. This means that increases in their productivity is not easily transmitted to other sectors due to their reduced connectivity (Ocampo, 2014; Britto et al., 2019). Similarly, the economic complexity literature states that these industries have the lowest Product Complexity Index (Hidalgo and Hausmann, 2009; Hausmann et al., 2014) and that these products are not good for promoting economic growth, as they are not well connected and thus cannot stimulate other products for growth. Indeed, the oil sector produces such products - it is one of the most disconnected sectors.

Figure 3.3 shows the results for the weighed degree centrality. As it was previously explained, this measures the number of links that a product is associated with and weights them by the magnitude of their respective transactions. In contrast with the previous results, some of the primary products increase in importance. The main change observed is related with crude oil and gas (P6), which has the highest centrality among other primary products. Also, some secondary products (as textiles - P15, wood products - P16 and paper

Figure 3.2: Products importance, based on degree centrality



- P17 and some services (electricity - P22, construction - P23, communications - P26, and others) reduced their importance under this metric. On one hand, this means that for the case of crude oil, the importance increases due to the size of their transactions (it is important to recall that this has been the main economic engine of Ecuador since 1972). On the other hand, for products and services that reduced their importance, albeit well connected, the magnitude of the transactions are low, in comparison to the rest of the Ecuadorian economy.

Figure 3.4 shows the results for eigenvector centrality. We observe that the results are similar to plot 3.2, which is caused by the topology of the network. As explained previously, the eigenvector centrality measures how central the nodes are based on the importance of the nodes that are connected with them. Since the network is bipartite, with only two type of nodes (*Products* and *Sectors*), the existing links are between them and not among products or sectors. Therefore, the eigenvector centrality coincides with the degree centrality.

3.4.2 Key Sectors by Centrality

Similar to the results obtained for products, the centralities of the different economic sectors vary depending on their kind (figure 3.5). For instance, we observe that primary ($S1 - S6$) sectors have a lower centrality than secondary ($S7 - S21$) sector and services ($S22 - S34$).

Figure 3.3: Products importance, based on weighted degree centrality

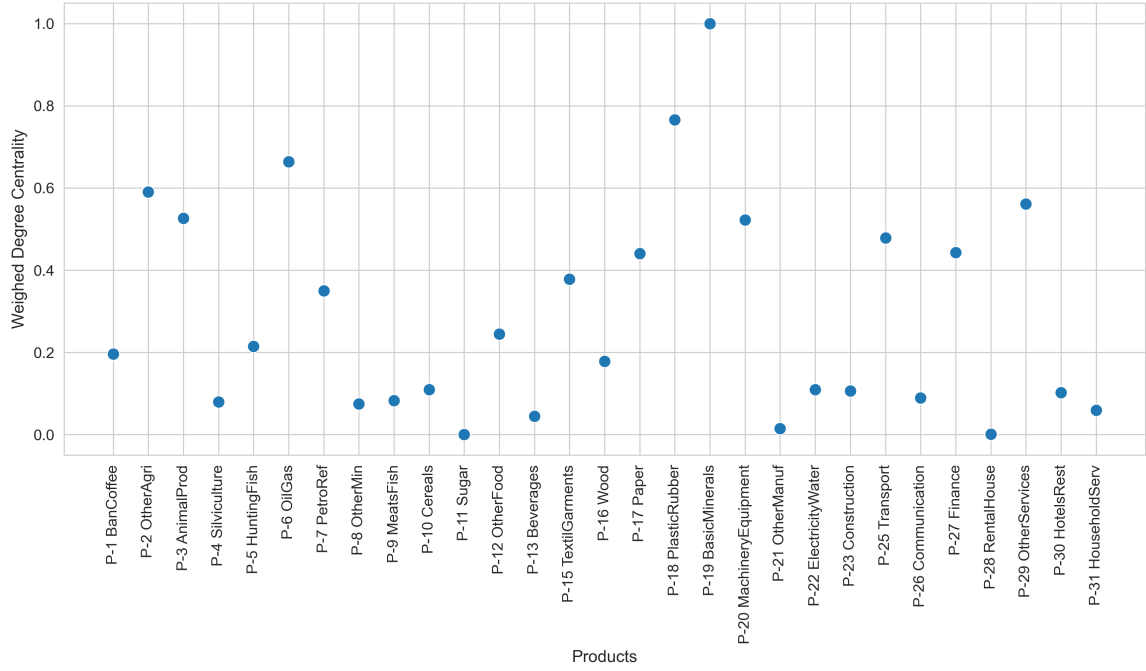
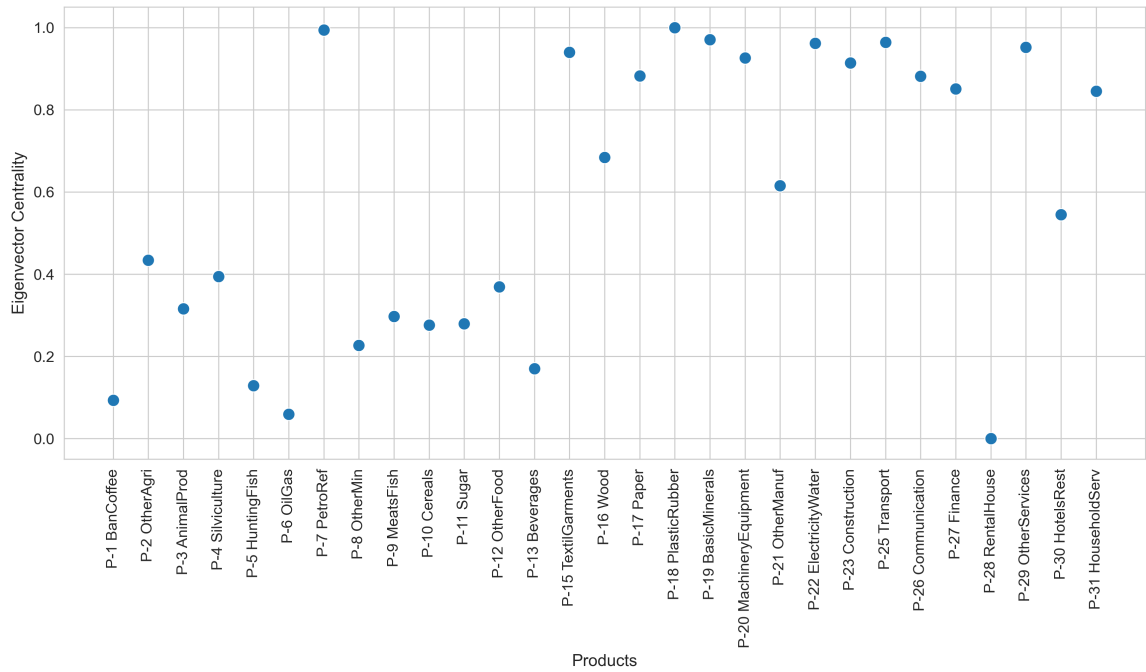
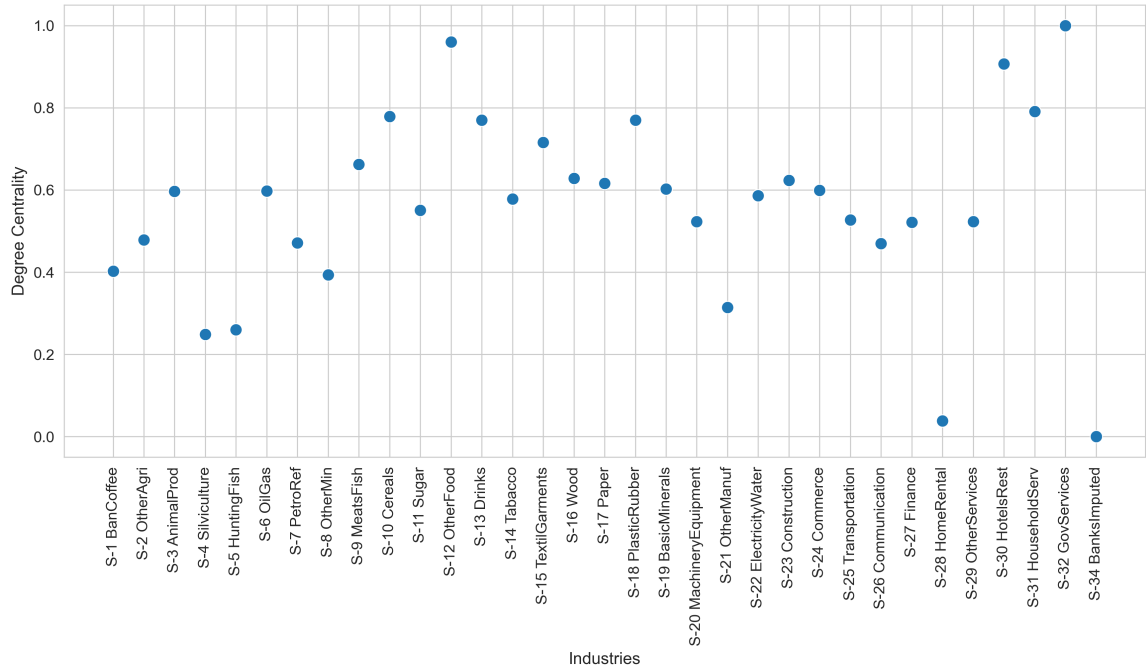


Figure 3.4: Products importance, based on eigenvector centrality



It is important to recall that the centrality in this case accounts for the number of products that "feed" the specific sector. Therefore, the oil and gas sector has a higher centrality among the primary sectors, which means that the sector has more inputs of products in comparison to the others.

Figure 3.5: Sector relevance, based on degree centrality



In figure 3.6, we observe a more distinguishable pattern between the sectors. It is clear that the primary sectors (S1-S6) have a lower average weighted degree centrality than the other sectors. This means that these sectors receive less inputs (not only measured by their inward connections, but taking into account the monetary inflows) than secondary sectors (P7-P21) and services (P22-P34). In other words, these primary sectors do not demand much products inputs in comparison with the rest of industries. Sectors such as commerce (S24), petroleum refining (S7), construction (P23) and commerce (S24) have the highest centrality, which means that they capture most of the inputs (in magnitude) from the Ecuadorian economy.

When we observe the eigenvector centrality in figure 3.7, there are patterns similar to the degree centrality plots. As it was explained previously, this similarity is due to the topology of the bipartite network.

Figure 3.6: Sector importance, based on weighted centrality

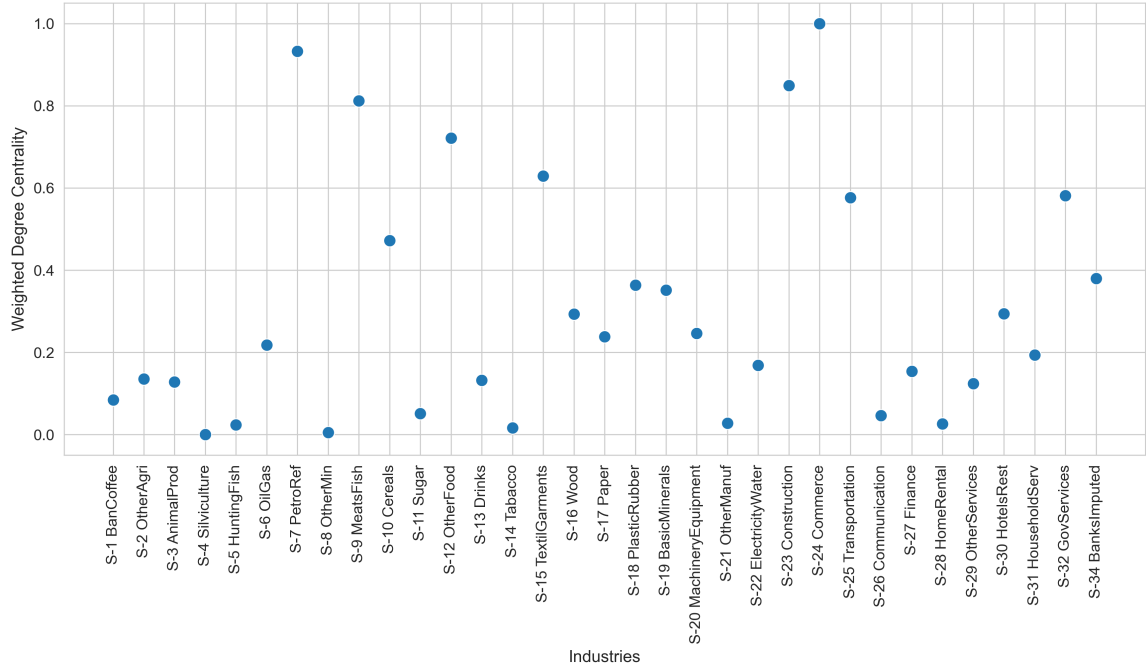
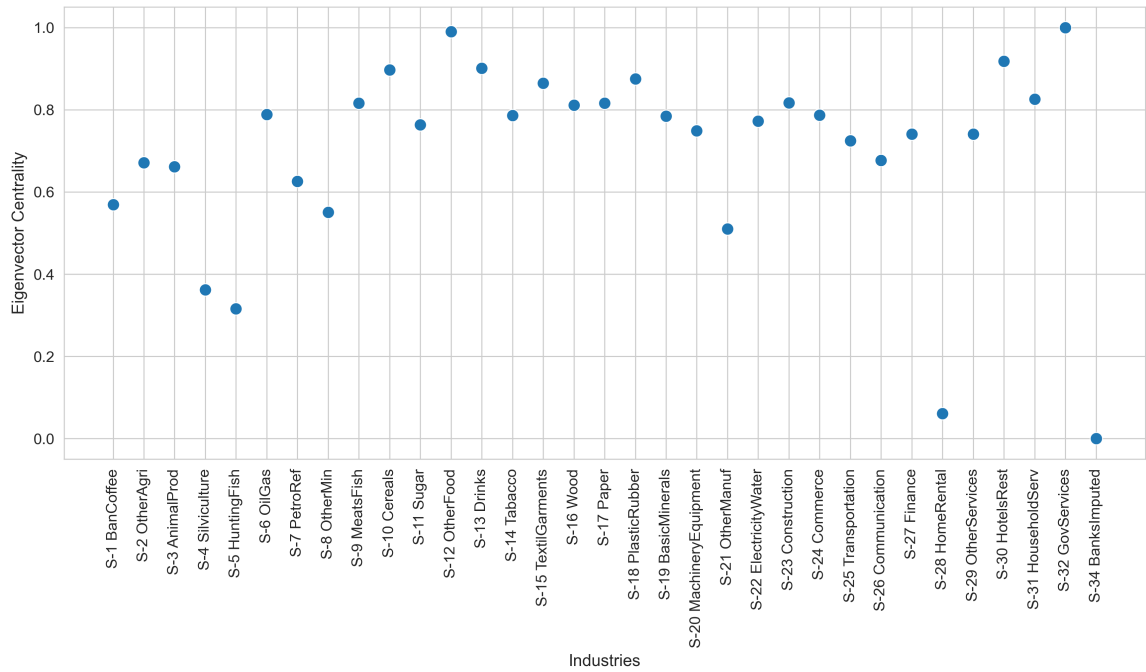


Figure 3.7: Sector importance, based on eigenvector centrality

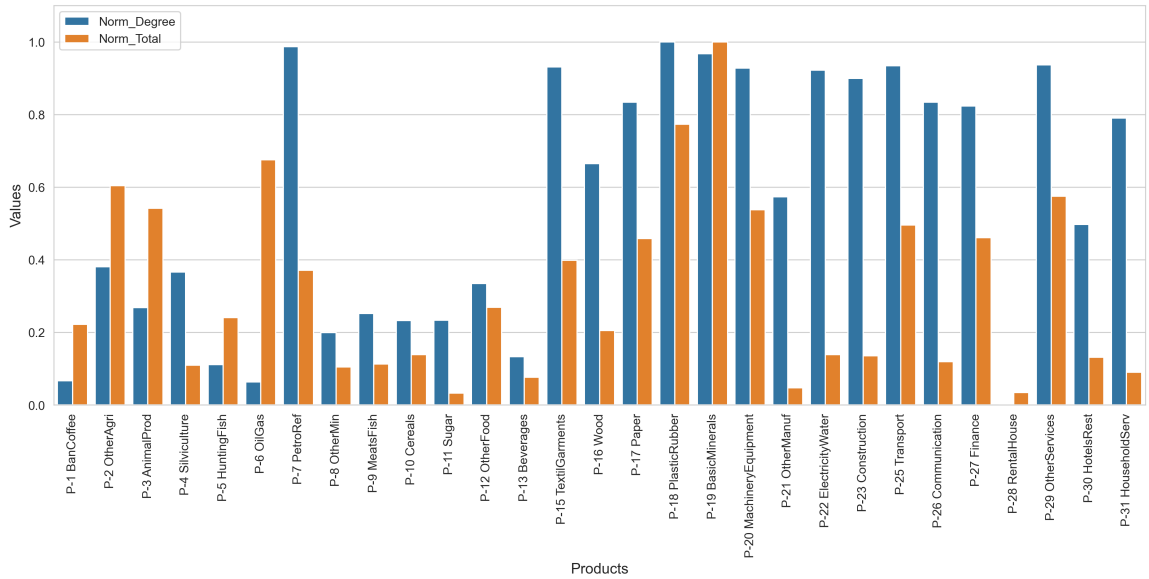


3.4.3 Comparisons between aggregated and network measurements

As it was explained before, a way to explore the internal economic structure of a country is to understand the role that individual sectors and products play and how they correspond to each other. This is important because the position (or centrality) of sectors in an economic system can have an important effect on the economic performance of a country (Noguera-Méndez et al., 2016).

Further, as Meller and Marfan (1981); Hausmann and Hidalgo (2011); Hidalgo and Hausmann (2009); Darwent (1969), and Albeaik et al. (2017) explained, how connected sectors or products are indicates the capacity that they possess to stimulate others. Therefore, investing in very interconnected sectors could foster higher intensity in the processes of economic development. In figure 3.8 and 3.9, the centrality and intermediate consumption measures were normalised ($Norm_{Total}$) to demonstrate how they compare with one another. We observe that in the case of crude oil, for instance, it is one of the most important for the Ecuadorian economy (i.e., with the highest weights). However, when we look at its centrality, it has one of the lowest values, meaning that it is one of the worst connected products in the Ecuadorian economy.

Figure 3.8: Comparison between standard (total demand) and network (degree centrality) measurements

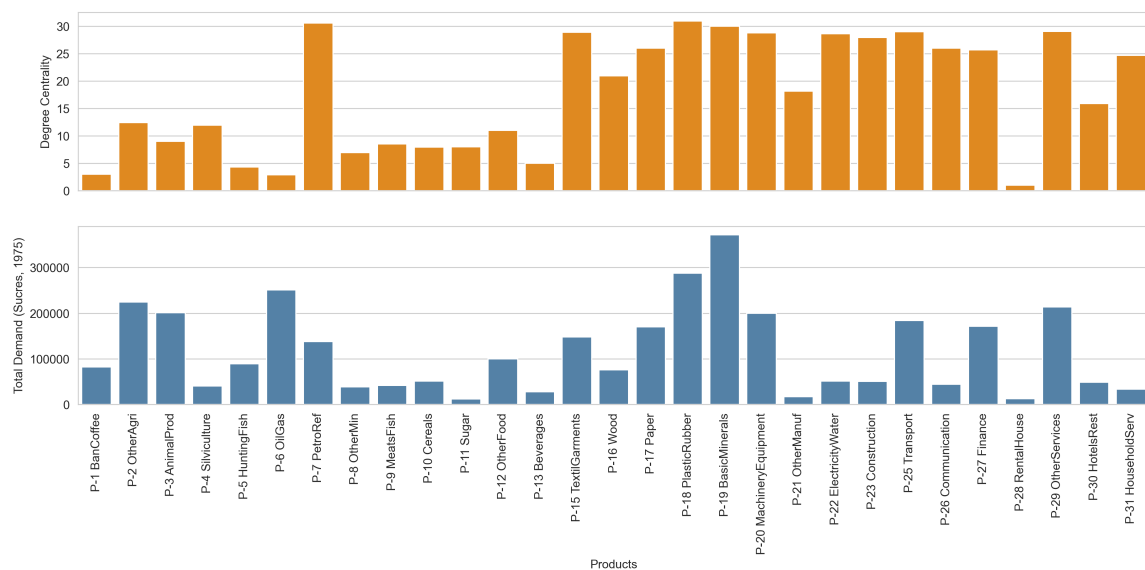


According to Noguera-Méndez et al. (2016), a relevant development question concerns the structural positions of the industries and systems that shape their relations. In this case, crude oil has the lowest position within a productive network. Therefore, relying on a sector that has a low degree of interconnectivity could mean that this one is not that relevant for

creating a stimulus for the system. Therefore, despite the oil sector generating huge amounts of revenues for the country, it only stimulates sectors that are connected with itself and not others. I will explore this further in the next chapter.

In figure 3.9, we can also observe the role that secondary and tertiary sectors play. Since manufacture goods are the biggest consumers of intermediate inputs, understanding their magnitudes of both intermediate consumption and centralities is also important.

Figure 3.9: Comparison between network (degree centrality) and standard (total demand) measurements



3.5 Conclusions

As discussed in the beginning of this chapter, a vast number of scholars has argued, for more than three decades, that natural resource dependency is correlated with low economic growth, weak institutions, lack of human capital, among others issues. In particular, these research have shown that oil exporting countries tend to have a lower economic performance in terms of growth, in comparison to non-oil exporting countries. However, it has been stressed that despite the efforts to study these countries there is not much work that have investigated how resource-rich economies are configured internally nor explored how different industries interact with each other. In addition, it has been stated that current research and their methodologies have focused on places where data are rich and easily accessible and excluded countries (mostly developing ones) where historical data is difficult to find or is incomplete.

To extend the existing knowledge about the structural characteristics of oil exporting

economies and its relationship with economic growth, the objective of this chapter was to assess how an oil-dependent economy is structured internally and to identify the industrial importance based on their centrality. To accomplish this, I proposed a network-based methodology that extracts structural information out of Input/Output data which allowed the identification of key economic sectors based on how connected they are (central) with the rest of the economy. Ecuador was chosen to test this methodology because of its long-standing dependency on crude oil exports and the availability of complete and consistent historical data since the mid-1960s, encompassing when its oil dependency started. In this section, I present two sub-conclusions. The first one relates exclusively to the methodology proposed in this chapter and the second one relates to the results.

3.5.1 Methodological conclusions

The main focus of this chapter, and the thesis as a whole, focuses on understanding economic structures and their relationship with economic growth. To do so, one needs to identify patterns of interactions between different economic actors (i.e., people, firms, industries) that traditional economic analysis usually circumvent (Jackson, 2014). The methodology proposed in this chapter is grounded in Network Science, whose main focus is to understand those interaction patterns and could provide insights into the economic behaviour of a system that is dynamic and constantly evolving (Barabasi, 2016; Jackson, 2014, 2010).

The core of this chapter uses the centrality metrics, which is a main component of network analysis. This metric allows identifying the role that every industry and product have, based on their connectivity, within an economic system. In the case of measuring the importance of a sector or a product based on their centrality, I conclude that using degree centrality and weighted centrality to examine these particular networks is more useful than the eigenvector centrality. As it was explained before, the eigencentrality of a node is assessed by measuring the importance of the other nodes that are connected with it, and their relative importance within the network. Therefore, due to the topology of our networks (*use tables*), the relative importance of the nodes is only based on the number of their connections (degree centrality) and as result their eigenvectors are similar to their degrees.

Finally, the centralities calculated allow us to distinguish between sectors that provide simple or more complex products or services, based on the inputs they required. For example, we observed that the sectors and products associated with the extraction or direct use of natural resources such as agriculture, mining or oil are the ones that have the lowest centralities, because they require less inputs than manufacturing or services. This finding is coherent with the work developed by economic complexity, since these products have the

lowest product complexity indexes and the sectors with higher centralities in our studies have the highest product complexity indexes, too. This will be further explained below.

3.5.2 Research conclusions

First, one of the main take always from this chapter is the differences that we can observe when comparing the aggregate output of the different sectors (traditional economic approach) to the centrality measurement (proposed network approach). On one hand, crude oil generated approximately 57.6% of the [exports] income of Ecuador and represented 13.6% of its GDP in our study period, which makes it the most important industry in the country in absolute terms (traditional approach)(BCE, 2019). On the other hand, when we measure its importance based on its links and the capacity to stimulate other sectors, as suggested by Meller and Marfan (1981), the oil industry has the lowest importance over other sectors. The results obtained show a dichotomy commonly experienced in developing countries that rely on extractive industries (oil, gas, mining): industries that generate big rents but are isolated (low connectivity) from the rest of the economy. Therefore, I argue that, if possible, every analysis of aggregated outputs should be complemented with network measurements so we could have a more complete picture of different economic dynamics.

Second, this finding is coherent with existing literature which states that crude oil has fewer linkages (i.e., require few inputs for production) and shares the fewest common characteristics with other products and sectors in an economy (Ahmadov, 2014) and (Lederman and Maloney, 2012). This means that these sectors generate income independently, without engaging with other domestic sectors (Auzer, 2017). Also, this industry is ranked between the lowest (20 out of 1219 products, according to the Product Complexity Index), which means that it has one the lowest complexities and is also one of the most isolated industries in the “product space”. This conclusion is also coherent with what many authors have argued for the extractive sector, because it is capital intensive and has limited links with other sectors of the domestic economy (Gelb, 2011). These insights will be explored further in Chapter 4 and Chapter 5, because this characteristic has a significant influence on the growth process of Ecuador as well as its opportunities for productive diversification.

Third, previous research has shown that there is a positive relationship between development and how industries are interconnected in an economy (Kubo et al., 1986). In their findings, Kubo et al. (1986) argued that a higher concentration of manufacture goods, which are the biggest consumers of intermediate inputs, is key for fostering sustainable economic growth. In our case, the income of the country relies mostly on an industry which is relatively isolated from the rest of the economy. Therefore, this could be one of the main constraints for growth

of Ecuador. Likewise, Ross (2019) also explained that it has the most difficult goods to diversify from. Therefore, this is a key factor to consider for economic diversification and the implementation of industrial policy.

Finally, the main contribution of this chapter has been to confirm our knowledge on key sectors and will serve as an analytical basis for the next chapter, which will examine how these centralities have evolved over time, and what was the role of the oil boom has had on its dependency. The limitations of this chapter will be stated in the conclusions of this thesis.

Chapter 4

Structural Productive Impact of a Resource Boom in Ecuador

4.1 Introduction

The Dutch Disease, which is the economic dimension of the Resource Curse, occurs when a country experiences a sudden windfall of income from an extractive industry, and subsequently its non-traded sector is crowded out by the growth in the natural resource sector (Corden and Neary, 1982; Humphreys et al., 2007; Auzer, 2017). Humphreys et al. (2007) and Auzer (2017) explain that the effects of the Dutch Disease take place because the exchange rates appreciate as a result of increased prices owing to increased aggregate demand; and the natural resource revenues tend to be consumed by public and non-tradable projects. So, these effects have the potential to deindustrialise the non-resource traded sector (Corden and Neary, 1982; Déés et al., 2008; Sachs and Warner, 2001; Gelb, 2011) - manifested by the manufacturing industry becoming less competitive as a result of an influx of cheaper imports (Humphreys et al., 2007; Auzer, 2017).

However, recent research has shown that the resource curse hypothesis, which includes the Dutch Disease, is far from conclusive (Stevens et al., 2015; Papyrakis, 2017; Auzer, 2017). The resource curse is in fact a much more complex phenomena and the previous attempts to generalise the relationship between resource dependency with the aforementioned factors have been largely unfruitful (Stevens et al., 2015; Papyrakis, 2017; Auzer, 2017). This means that studying the correlation between economic performance and resource dependency must be dealt with on an individual basis, since its manifestation depends on factors that are unique to specific regions or countries (Papyrakis, 2017).

Most of the existing literature on this matter has focused on analysing the aggregate outputs of different resource-dependent countries. However, it has examined neither how their

economies are configured internally, nor if a resource boom affected their productive structures. Drawing on the structuralist / network framework used in the previous chapter, this chapter aims to analyse how the Ecuadorian productive structure - measured by the centrality of their industries - has changed between 1965 and 1985, and identifying the effects caused by the 1972 oil boom with the help of a synthetic control method developed by Abadie et al. (2010, 2015). In other words, the main objective of this chapter is to assess if the Dutch Disease's effects can be observed on a structural level ¹ by identifying the crowding-out effect of the non-traded sectors.

The chapter is organised as follows: first, a literature review will be presented, which will summarise the existing debates around the Resource Curse and the Dutch Disease and its effects. Second, the evolution of the Ecuadorian economy will be described, with an emphasis on the role that the main commodities have played and the political conditions that surrounded the oil boom. Next, the methodology used will be described, including a discussion about its benefits and limitations. Finally, the results will be presented, followed by the corresponding discussion and relevant conclusions.

4.2 The Resource Curse and Dutch Disease

The "Resource Curse" hypothesis argues that countries, which are highly dependent on natural resource income tend to under-perform economically in comparison with non-resource rich countries (Ahmadov, 2014; Auty, 1994, 2001, 2002; Auty and Pontara, 2008; Bornhorst et al., 2008; Gelb and Grasmann, 2010; Gelb, 2011; Humphreys et al., 2007; Lederman and Maloney, 2006; Murshed and Serino, 2011; Papyrakis, 2017; Papyrakis and Pellegrini, 2019; Sachs and Warner, 1995, 2001; Stevens et al., 2015; Sy et al., 2011). It suggests that the under-performance could be explained by weak institutions, a lack of human capital, inequality, and in some cases high levels of corruption in rich-resource economies (Auty, 1994, 2001; Collier and Goderis, 2009; Papyrakis, 2017; Sachs and Warner, 1995, 2001).

The Dutch Disease (DD), which is the economic dimension of the resource curse, occurs in a country that experiences a sudden windfall of income from an extractive sector. This has an effect on its non-traded sector ², which is crowded out by growth in the natural resource sector and the non-resource traded sector ³ (Corden and Neary, 1982). In other words, Humphreys et al. (2007) and Auzer (2017) explain that this crowding-out effect

¹Changes in the industrial centralities calculated in the previous chapter.

²By non-traded sector I mean activities for which the domestic market determines prices, such as retail, health, education, construction.

³The non-resource traded sector corresponds to export industries (usually associated with manufacturing) with prices set by competition on world markets

causes an inappropriate allocation of resources between the traded and non-traded goods sectors, which at a macroeconomic level could result in lower economic growth.

Humphreys et al. (2007) and Auzer (2017) further explain that the effects of the Dutch disease take place because the exchange rates appreciate as a result of increased prices owing to increased aggregate demand, and the natural resource revenues tend to be consumed by public and non-tradable projects, such as roads, construction, power, telecommunications and other services. These effects have the potential to de-industrialise the non-resource traded goods sector (Corden and Neary, 1982; Déés et al., 2008; Sachs and Warner, 2001; Gelb, 2011) - manifested by the manufacturing industry becoming less competitive as a result of an influx of cheaper imports (Humphreys et al., 2007; Auzer, 2017).

However, recent research has shown that the resource curse hypothesis is far from conclusive (Stevens et al., 2015; Papyrakis, 2017; Auzer, 2017). The resource curse is in fact a much more complex phenomena and previous attempts to generalise the relationship between resource dependency with the aforementioned factors have been largely unfruitful (Stevens et al., 2015; Papyrakis, 2017; Auzer, 2017). Most importantly, some countries have experienced the "curse" in different orders of magnitude, whereas others have not experience it at all (Auzer, 2017; Tabata, 2012; Dobrynskaya and Turkisch, 2010; Lipscomb et al., 2017). For example, countries such as Nigeria, Ghana or Mexico have suffered the typical effects of the DD; other countries such as Malaysia, Botswana or Chile have managed to use their resource income to develop a more diverse economic productivity base (Auzer, 2017). In Russia, for instance, interesting work carried out by Tabata (2012) and Dobrynskaya and Turkisch (2010) shows that the country managed to increase the productivity of its manufacturing sector from 1999 to 2007, despite having observed symptoms of DD. Auzer (2017) explains that this can be considered to be just a specific case, due to particular monetary and rent-sharing policies implemented to maintain the country's international competitiveness. Likewise, Lipscomb et al. (2017) demonstrate than Indonesia considerably increased its manufacturing goods production between 2003 and 2008 after a resource boom. Yet there was not strong evidence of the country suffering from the DD and the weak performance in some sectors did not appear to be linked to the commodity boom.

This means that studying the correlation between economic performance and resource dependency must be dealt with on an individual basis, since its manifestation depends on factors that are unique to specific regions or countries. These factors include, for example, their type of natural resources and their linkages with the rest of the economy; how their relative importance is measured; the quality of the socio-political institutions in place; and so forth (Papyrakis, 2017; Papyrakis and Pellegrini, 2019).

In the previous chapter, we observed that the productive structure of a country is an important determinant for its economic growth. It also provided us with relevant insights into the structure of the Ecuadorian economy and allowed us to identify its key industries based on their centrality. This chapter is a continuation of the previous analysis and attempts to explore if the windfall of oil resources (oil boom) that Ecuador experienced in 1972-1973 had an effect on its productive structure. In other words, this chapter attempts to expand the existing literature on the Dutch Disease, by proposing a methodology that would allow exploring its effects at a structural level. As such, in this chapter, I hypothesise that oil boom had a negative effect on the centralities of the non-traded sectors.

To continue with this literature review, in the next section I will explore the historical primary exports dependency of Ecuador and give a brief overview of the current situation of the country.

4.3 Evolution of the productive structure of Ecuador

4.3.1 Colonial times to the 20th century

The dependency of the Ecuadorian economy on commodities export can be traced back to its colonial times. According to historical records, from 1534 to 1882, when the Ecuadorian territory was controlled by Spain, the main objective of its economy was to extract and provide natural resources (mainly metals) and manufacturing textiles for the Spanish Crown. By then, the institutions established in all the Spanish colonised territories had the sole objective of overseeing resource extraction (Lange et al., 2006; Benites, 2005). After Ecuador's independence in 1822, this economic logic lasted until the end of the nineteenth century (Bonilla et al., 1982). However, one of the few effects of the independence was a precarious insertion into world trade that also focused on the exportation of raw material, which affected significantly its emerging local manufacturing industry (Correa, 2012).

In the late 19th century, the decline of Spanish domain over its colonies and the rise of Britain's control over world trade had a direct impact on Ecuador (Marini, 2008). The shift of world powers deepened the "raw material and food supplier model" in Ecuador and helped to increase its exports to emerging markets (Correa, 2012). By the mid-20th century, Ecuador became an important provider of cocoa beans, bananas and other agricultural products. These revenues allowed the country to begin an incipient process of industrialisation (Correa, 2012).

4.3.2 Oil boom and Import Substitution Industrialisation (1960 - 1980)

In 1967, crude oil was discovered in the Ecuadorian Amazon. This discovery significantly boosted economic growth and in many ways changed the social structure of the country (Correa, 2012). Ecuador, once one of the poorest South American countries, became one of the middle income countries of the hemisphere. Its income per capita matched its neighbours (Colombia and Peru) and much social progress was made. Between 1960 and 1980, more than 10 years were added to Ecuadorians' life expectancy; death and infant mortality rates dropped by more than 40 per cent. School enrolment expanded rapidly. By 1980, the citizens were better educated, in better health, and better fed than in the country's history (Belli, 1984).

In the early 1970s, Ecuador's real GDP grew at an average rate of almost 15 per cent every year, of which more than 60 percent was attributable to crude oil exports. The value of exports (at constant prices) grew by 60 percent per year (Belli, 1984). The GDP per capita increased by approximately 72%; its exports rose from US\$ 400 million in 1970 to US\$ 1.1 billion in 1980 (1975 prices) and the real manufacturing output went up to 150% (Mateo and Garcia, 2014; Acosta, 2009b; Belli, 1984). Due to oil exports, Ecuador's terms of trade turned strongly in its favour (Belli, 1984). In 1973, Ecuador joined the Organisation of the Petroleum Exporting Countries (OPEC) and became a smaller member, which has not changed since. Ecuadorian production represented, on average, 1.2 percent of total output of the OPEC since the oil boom⁴ (figure 4.1) (Baquero, 2017). Additionally, state participation was relatively high, to the extent that 80% of its revenues were generated by its state-owned oil companies (Larrea, 2013).

The years after the 'boom' (1974-1979) were characterised by rapid and steady GDP growth (almost seven percent annually) (Belli, 1984; Larrea, 2013). During this period, the State pursued a development strategy based on import substitution industrialisation (ISI), investing in infrastructure, and using subsidised credit, tariff protection, tax incentives to promote growth and diversification (Larrea, 2013). The ISI strategy, accompanied by the drop of import prices, made manufacturing the leading growth sector, expanding at an average of 9.5 percent yearly (except for sugar, all sub-sectors grew faster than GDP) (Larrea, 2013; Belli, 1984). Textiles and mineral products alone accounted for over a third of total manufacturing growth (figure 4.2a) (Belli, 1984). However, the agricultural sector barely kept the pace with population growth (figure 4.2b) (Larrea, 2013).

⁴Ecuador left the OPEC in 1992, to joined later in 2007. In 2020, the country left the organisation again.

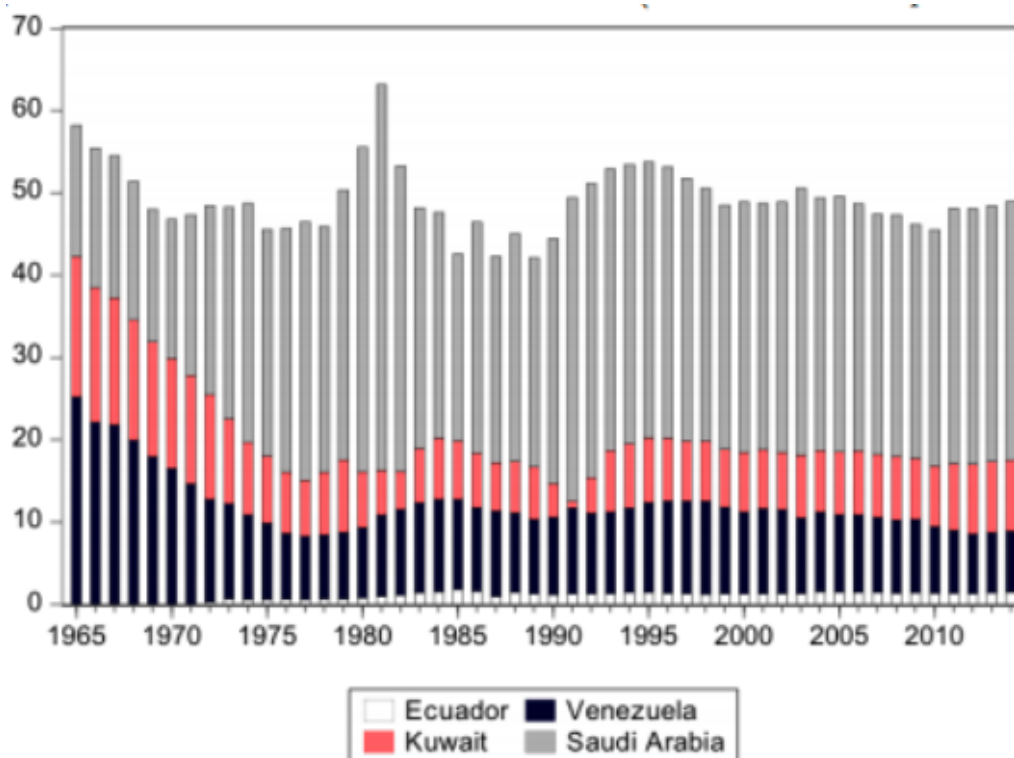


Figure 4.1: OPEC Oil Production in selected countries (% of OPEC total production)

In the 1970s, non-oil exports - mostly composed of bananas, coffee and cacao beans - enjoyed favourable international prices and received subsidised credit and other benefits. However, they were affected by the overvalued exchange rates during the oil boom period, and as a result, experienced slow growth (Larrea, 1992). With the important exception of shrimp, non-oil exports grew slowly in volume after 1974 and fell by a third in 1983 (Belli, 1984; Acosta, 2006). During this time, export diversification was also moderate (Larrea, 2013). Several authors argue that these effects were caused by the “Dutch Disease” (Acosta, 2009b; Larrea, 2013).

Although the growth of manufacturing was significant during this period, the sectors’ performance was weak (Belli, 1984). In fact, during the bonanza, structural diversification of manufacturing was minimal. The attempts for industrialisation were mostly short-lived, accompanied by a small demand expansion in protected sectors of the economy between 1972 and 1982 (Belli, 1984). Finished consumer goods remained dominant, with minimum development of intermediate, capital goods and of internal linkages (Larrea, 2013).

Larrea (2013) argues that a purely “horizontal” expansion prevailed, without the development of domestic internal linkages. Moreover, Vos (1987) and Belli (1984) analysed the changes in the internal structure in manufacturing and concluded that domestic linkages

Figure 4.2: Ecuador' Exports from 1972 to 1982

	Annual Growth Rates 1972-1982	Value Added (Composition)		Contribution to Growth 1972-1982	Average Annual Growth	
		1972	1982		1970-1973	1974-1979
Total Manufacturing	9.5	100.0	100.0	100.0		
Food, Beverages & Tobacco	8.5	48.0	46.9	35.5		
Fish & Meat	8.1	7.5	11.5	7.6		
Cereals	12.6	7.4	5.7	6.4		
Sugar	1.7	7.1	5.2	0.4		
Other Food	8.8	12.9	12.4	9.4		
Beverages	9.6	11.9	8.5	9.5		
Tobacco	14.0	1.2	3.5	2.2		
Textiles	9.4	20.7	21.5	22.0		
Wood Products	8.2	5.5	5.3	5.1		
Paper & Printing	8.6	5.8	5.4	5.1		
Chemicals	10.9	7.3	5.3	9.0		
Mineral Products	11.7	9.6	10.6	13.5		
Machinery & Equipment	8.1	3.2	2.8	4.8		
Other Manufacturing ^{a/}	---	---	2.4	4.9		
^{a/} Data not available for all years.					^{a/} Components do not add up to total because indirect tax	
Source: Statistical Appendix Table 8.1					(b)	
(a) Manufacture					Sector	
					growth	

in manufacturing were low and changed little during the oil bonanza. Larrea explains that the rapid growth in the 1970s was partially financed by foreign debt, which accumulated and became unsustainable under rising interest rates in the early 1980s (Larrea, 2013).

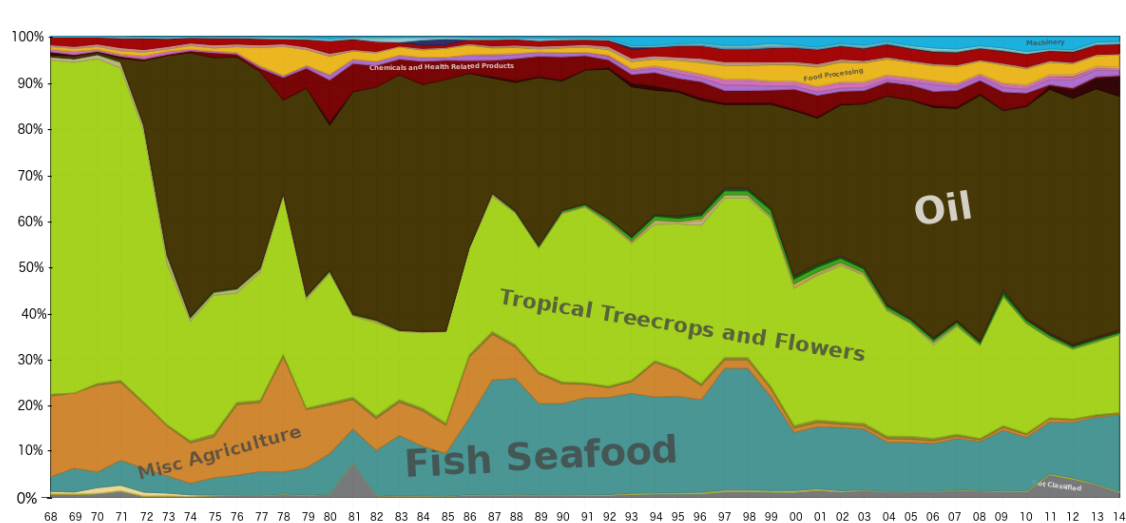
4.3.3 Structural Adjustment Period and Post-Neoliberal reforms (1980-2010)

After the Mexican debt moratorium⁵ in 1982, international financial institutions imposed rigid conditions on Latin American countries, the ISI strategy collapsed, and structural adjustment and liberalisation strategies were implemented (Larrea, 2013). As real oil prices dropped, mostly between 1986 and 1998, foreign debt became unmanageable. Moreover, some authors have stressed that since the 1990s, when the Ecuadorian economy opened up to foreign competition, manufacturing protection was drastically reduced (Larrea, 2013; Acosta, 2006). As a result, the manufacturing share in GDP declined, eliminating the gains of the 1970s (Acosta, 2006, 2009b). However, during the 1990s, banana export expansion became important, and a group of non-traditional exports maintained a steady growth rate from the early 1990s to the present, currently accounting for 9% of total exports. Nevertheless, export diversification has been low and insufficient (Figures 4.3 and 4.4) (Larrea, 2013).

⁵In August 1982, Mexico unilaterally announced a moratorium on its foreign currency debt, because the country had failed to raise a large enough loan to repay external debts. Other Latin American countries had similar debt problems as Mexico, so most commercial banks reduced significantly or halted new lending, which caused many problems for the region: incomes and imports dropped, economic growth stagnated, unemployment rose to high levels, and inflation reduced the buying power of the middle classes. (Stambuli, 1998; Cline, 1982)

In 2000, after Ecuador adopted the US dollars as its national currency, the oil sector increased its dominance over the economy and became the main source of incoming foreign exchange, severely affecting non-oil exports (Larrea, 2004). Later in 2004, oil extraction was boosted by a new pipeline (Oleoducto de Crudos Pesados - OCP), which doubled transportation capacity. Consequently, Ecuador experienced the largest expansion of oil exports in its history, as they grew strongly both in terms of quantity and prices (Larrea, 2013). Once again, Larrea argues that despite the public policies placed, economic diversification was poor.

Figure 4.3: Ecuadorian exports share 1968 - 2014



Since 2007, a set of policies were put in place to support “selective import substituting industrialisation”. However, Larrea (2013) contends that they failed to diversify the internal structure of manufacturing, with no significant changes observed between 2005 and 2009 (Table 4.1). Despite subsidised credit and other incentives, agricultural modernisation did not lead to growth in the sector. Per capita agricultural output for domestic consumption declined, and export agriculture expanded slowly (Larrea, 2013). By 2012, approximately 58% of the total Ecuadorian exports were crude oil products, which represented 38.7% of the state budget in that year (MRNNR, 2013). In addition, non-crude oil commodities and industrial products stood at approximately 32% and 10%, respectively (Mateo and Garcia, 2014).

4.3.4 Current situation

Currently, Ecuador still depends heavily on oil exports. According to Ecuador’s Central Bank statistics and information of capital flows and income of central government budget, oil

Table 4.1: Manufacturing structure by sub-sector (1965 - 2009)

SUB-SECTOR	1965	1972	1977	1982	1988	2005	2009
Food, beverages and tabaco	58.10	46.30	42.80	41.60	37.50	36.30	37.20
Textiles and clothing	16.20	22.10	23.60	22.10	19.80	8.80	8.10
Wood and wood products	6.20	6.00	5.90	5.30	5.70	5.70	5.70
Paper and printing	7.40	7.10	6.00	6.30	7.30	6.50	5.90
Chemical products	4.80	5.70	6.50	6.10	6.50	14.40	13.40
Non-metallic minerals and basic metals	5.40	9.10	10.00	12.40	12.70	12.70	13.20
Metal products and machinery	2.00	3.60	5.30	3.10	4.70	10.60	11.30
Other	0.00	0.00	0.00	2.90	5.80	5.00	5.20

Source: Larrea (2013, p. 12)

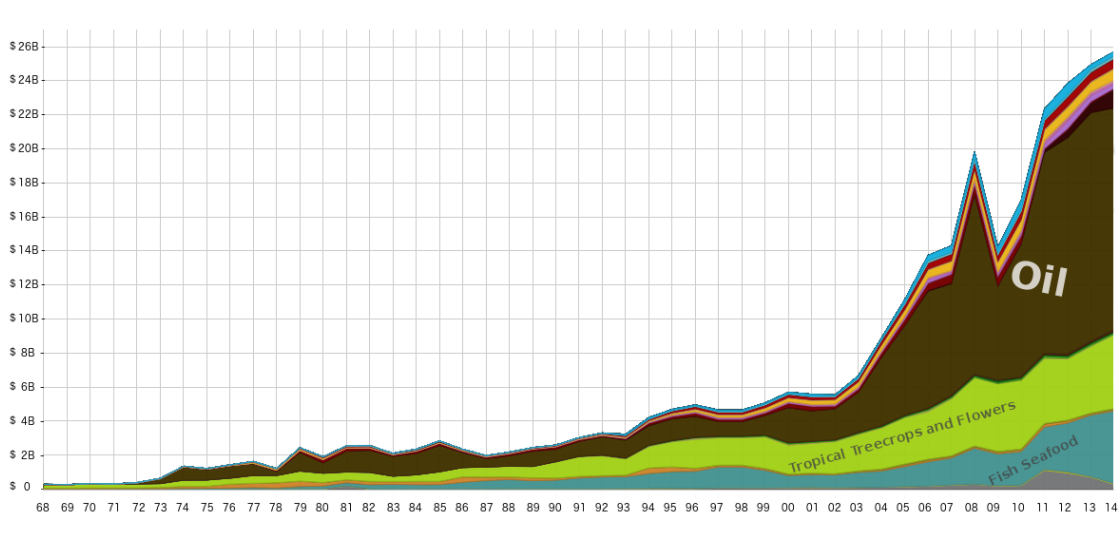
exports represent approximately fifty percent of its total exports and, on average, 30% of the total income of the state budget in the past six years (BCE, 2019). The country has proven reserves of 7,200 million barrels (MMbbl) of crude oil and 3,437 MMbbl of them are residual reserves. Its average production in 2012 was 504,000 barrels a day and generated a revenue of approximately 11,000 million US Dollars (MRNNR, 2013; Herrera and Jarrin, 2013). Furthermore, according to the most optimistic scenarios, there is a potential additional 1597 MMbbl of crude oil in the Ecuadorian Amazon, which was tendered internationally (MRNNR, 2013). However, it has been stated that these oil reserves have already reached the peak and will be exhausted within the next 30-40 years (Espinoza et al., 2019; Mateo and Garcia, 2014). In fact, in several scenarios, Ecuador could become a net oil importer between 2024 and 2035 and economic oil trade balance could be seriously affected, decreasing from a current positive value of around 2 billion USD to incur deficits of 0.6–16.7 billion USD in 2035 (Espinoza et al., 2019).

Meanwhile, in the last decade, the prices of oil rose and fell significantly. This volatility translates into variability in terms of trade of the country, affecting the purchasing power of its exports, and hence the real level of income (Zambrano, 2016). Acosta (2009a) explained that the importance for the Ecuadorian economy is such that a single reduction of one dollar in the price of the Ecuadorian oil barrel translates into a net decrease of fiscal revenues of approximately 57.8 millions of dollars [by 2010].

Finally, the overall economic growth in Ecuador after crude oil discovery was poor, with an average growth in per capita income (1.38% per annum), which was lower than the pre-oil period (2.42% per annum), with minimal productive diversification. In addition, only shrimps and flowers have been added to the export basket in the past four decades (Larrea, 2013). Manufacturing has not significantly increased its share in its GDP. The banana sector, which is Ecuador's biggest export after oil⁶, is facing serious challenges. It is possible

⁶Banana exports generate \$US 2.6 billion and 2.5 million jobs in the country could be seriously affected in the coming years (Perez, 2019)

Figure 4.4: Ecuadorian exports 1968 - 2014



that a fungus (*Fusarium R4T*), having been detected in Colombia, may affect Ecuador's production as well, with the potential to diminish the country's plantations (Perez, 2019). As such, it is not only important to understand its commodity dependence in more detail but also finding diversification pathways.

As it can be observed, Ecuador has a long-standing dependency on a few primary product exports and its economic performance has always been linked with the fluctuation of international commodity prices. Many studies have shown how the country experienced the Dutch Disease during its oil boom as well as discussions on its weak industrial capacity. In addition, studies have argued that all products exported by Ecuador face serious constraints in the near future, because its oil extraction capacity will inevitably decline given limited reserves and farmed shrimps and banana markets are almost saturated (Larrea, 2013). Therefore, Ecuador needs to examine its economic structure, how it has been affected in the past, so that it can take corrective measures to face a not-so-bright future. In this sense, this chapter explores how the Ecuadorian productive structures changed in the past, and would provide insights into how it may deal with the diversification challenges which will be analysed in the next chapter.

4.4 Methodology and Data

To address the objective of **Chapter 4**, which aims to analyse how the Ecuadorian productive structure has changed between and identifying the effects caused by the 1972 oil boom, I use the data generated in Chapter 3. This data contains information about the changes

of the industrial centralities of the Ecuadorian economy between 1965 to 1985. To explore the effects that the 1972 oil boom had in the productive structure of the country, we used the Synthetic Control method (SCM) proposed by Abadie et al. (2010, 2015), which will be described below.

In our particular case, ideally we would want to have data on countries which were similar to Ecuador and did not experience an oil boom. This would allow us to estimate effects of an oil boom because it would resemble an experiment. However, there are no data about the countries that could be comparable to the Ecuadorian economy. As such, we resorted to a solution offered by the SCM. This method facilitates comparative case studies in instances when no single untreated unit provides a good comparison for the unit affected by the treatment or event of interest. This is often the case when the treatment affects large aggregates like regions or countries, so that a limited number of untreated units are available (Abadie et al., 2015). In effect, it creates an "artificial" Ecuadorian economy without the oil boom. With this, it is possible to compare how the synthetically-generated economy is different to the "real"⁷ Ecuadorian economy after an oil boom.

4.4.1 Synthetic Control Method

The Synthetic Control Method (SCM), developed by Abadie et al. (2010), is a technique that uses a data-driven procedure to construct comparison groups to reduce discretion in the choice of the comparison control units. The authors explain that this method overcomes one of the main limitations of comparative studies, where researchers often select comparison groups on the basis of subjective measurements of affinity between affected and unaffected units. They argue that this approach forces researchers to demonstrate the affinities between the affected and unaffected units using observed quantifiable characteristics.

As the authors further explain, uncertainty about the values of aggregate variables chosen as sample units can be eliminated completely if aggregate data are available. In other words, a combination of units often provides a better comparison for the unit exposed to the intervention than any single unit alone. Similarly, Abadie et al. (2010) explain that this method, in contrast with traditional regression methods, has the advantage of making explicit:

- a. The relative contribution of each control unit to the counter-factual of interest.

⁷By "real" economy, I refer to the data recorded in the National Accounts. It is important to recall that this "real" economy leaves the informal economy, which it is difficult to quantify. According to some estimates approximately 44.5 per cent of the Ecuadorian economy in 2016 was informal, and this figure was larger in earlier decades (González and Cabrera, 2018)

- b. The similarities (or lack thereof) between the unit affected by the event or intervention of interest and the synthetic control, in terms of pre-intervention outcomes and other predictors of post-intervention outcomes. Because the weights can be restricted to be positive and sum to one, the synthetic control method provides a safeguard against extrapolation.

In addition, the authors argue that the synthetic control method allows researchers to decide on study design without knowing how those decisions will affect the conclusions of their studies, because the choice of a synthetic control does not require access to post-intervention outcomes (Abadie et al., 2010).

4.4.2 Application of the method to the Ecuadorian case

Let's suppose that we observe a sample of $J + 1$ units, which are indexed by j . Among this, $j = 1$ is the unit that is exposed to the event of interest during periods $T_0 + 1, \dots, T$, which in our case pertains to the effects of the 1972 oil boom; and the remaining units $j = 2$ to $j = J + 1$ are potential controls, in which the intervention has no effect during the pre-treatment period $1, \dots, T_0$. Furthermore, we assume that the unit to be analysed is uninterruptedly exposed to the intervention of interest after some initial intervention period, and that the sample includes a positive number of pre-intervention periods, T_0 , as well as a positive number of post-intervention periods, T_1 , with $T = T_0 + T_1$.

To use this method, a balanced panel is needed, which is a longitudinal data set where all units are observed at the same time periods, $t = 1, \dots, T$. In our particular case, the results obtained in the previous chapter are the sample to be analysed. They consist of annual national-level panel data containing the centrality measurements of 29 products from the Ecuadorian economy, for the period 1965-1995. These data have already been balanced, as required for the present analysis. The products that encompass the analysis are also uninterruptedly exposed to the effects of the 1972 oil boom. The sample period starts in 1965 because it is the first year of data available in the Ecuadorian National Accounts (NA). It ends in 1995, because in that year they stopped publishing the NA, to be resumed only in 2006. As the oil boom started in mid-1972, data includes 7 years of pre-intervention data, from 1965 to 1972, and a post-intervention period, from 1973 to 1985.

A crucial part of the analysis is the selection of the control units, which are meant to approximate the counter-factual of the case of interest. Abadie et al. (2010, 2015) explain that it is important to select the donor pool (potential controls) identifying the units that have not been affected by the intervention (oil boom) or whose levels of exposure to it differ notably. In our case, the units of interest consists of every product that is included

in the data set and the potential controls are a sample of the products that are not units of interest. However, this could represent a further challenge, because due to the nature of the Dutch Disease and the interconnectedness of a domestic economy, all the units might be affected by the intervention. In other word, we choose a product of interest to be studied and we build the synthetic product as a weighted average of potential control products. We repeat this procedure for every products. For example, we take the agricultural products as the unit to be analysed and we build the control based on the other products that are not being analysed. This approach could eventually lead to an underestimation of the effect of the oil boom, since some controls could have been affected indirectly by the boom.

Once we have selected the donors, we define a synthetic control as a weighted average of the units in the donor pool. That is, a synthetic control can be represented by a $(J \times 1)$ vector of weights $W = (w_2, \dots, w_J + 1)$, with $0 \leq w_j \leq 1$ for $j = 2, \dots, J$ and $w_2 + \dots + w_J + 1 = 1$. Choosing a particular value for W is equivalent to choosing a synthetic control. Following Mill's Method of Difference ⁸, we select the value of W such that the characteristics of the treated unit are best resembled by the characteristics of the synthetic control. Let X_1 be a $(k \times 1)$ vector containing the values of the pre-intervention characteristics of the treated unit that we aim to match as closely as possible, and let X_0 be the $k \times J$ matrix collecting the values of the same variables for the units in the donor pool. The pre-intervention characteristics in X_1 and X_0 may include pre-intervention values of the outcome variable (Abadie et al., 2015, 2010).

The difference between the pre-intervention characteristics of the treated unit and a synthetic control is given by the vector $X_1 - X_0W$. We select the synthetic control, W^* , that minimises the size of this difference. This can be operationalized in the following manner: for $m = 1, \dots, k$, let X_{1m} be the value of the m -th variable for the treated unit and let X_{0m} be a $1 \times J$ vector containing the values of the m -th variable for the units in the donor pool. Abadie et al. (2010); Abadie and Gardeazabal (2003) choose W^* as the value of W that minimises:

$$\sum_{m=1}^k v_m (X_{1m} - X_{0m}W)^2$$

where v_m is a weight that reflects the relative importance assigned to the m -th variable when the discrepancy between X_1 and X_0W is measured. For Abadie et al. (2015), it

⁸Mill's methods are attempts to isolate a cause from a complex event sequence. Its Method of Difference is when two or more instances of an event (effect) are compared to see what they all do not have in common (Baronett, 2015)

is crucial that synthetic controls closely reproduce the values that variables with a large predictive power on the outcome of interest take for the unit affected by the intervention.

Accordingly, those variables should be assigned large v_m weights. In the empirical application below, it was applied a cross-validation method to choose v_m , as proposed by Abadie et al. (2015). Let Y_{jt} be the outcome of unit j at time t . In addition, let Y_1 be a $(T_1 \times 1)$ vector collecting the post-intervention values of the outcome for the treated unit. That is, $Y_1 = (Y_{1T_0+1}, \dots, Y_{1T})'$. Similarly, let Y_0 be a $(T_1 \times J)$ matrix, where column j contains the post-intervention values of the outcome for unit $j + 1$. The synthetic control estimator of the effect of the treatment is given by the comparison of post-intervention outcomes between the treated unit, which is exposed to the intervention, and the synthetic control, which is not exposed to the intervention, $Y_1 - Y_0 W^*$. In this sense, for a post-intervention period t (with $t \geq T_0$), the synthetic control estimator of the effect of the treatment is given by the comparison between the outcome for the treated unit and the outcome for the synthetic control in that period:

$$Y_{1t} - \sum_{j=2}^{J+1} W_j^* Y_{jt}$$

The matching variables in X_0 and X_1 are meant to be predictors of post-intervention outcomes. These predictors are themselves not affected by the intervention. Critics of Mill's Method of Differences rightfully point out that the applicability of the method may be limited by the presence of unmeasured factors affecting the outcome variable, as well as by heterogeneity in the effects of observed and unobserved factors (Abadie et al., 2015). However, using a linear factor model, Abadie et al. (2010) argue that if the number of pre-intervention periods in the data is large, matching pre-intervention outcomes (i.e., on the pre-intervention counterparts of Y_0 and Y_1) helps control unobserved factors and the heterogeneity of the effect of the observed and unobserved factors on the outcome of interest. Abadie et al. (2015) explain that intuition of this result is straightforward: only units that are alike in both observed and unobserved determinants of the outcome variable as well as in the effect of those determinants on the outcome variable should produce similar trajectories of the outcome variable over extended periods of time. Once it has been established that the unit representing the case of interest and the synthetic control unit have similar behaviour over extended periods of time prior to the intervention, a discrepancy in the outcome variable following the intervention is interpreted as produced by the intervention itself (Abadie et al., 2015).

In our case, the predictors for centralities are product exports and their respective internal consumption. These variables are averages over the 1966-1972 period. Finally, using the techniques described by Abadie et al. (2010, 2015), we construct a synthetic product that mirrors the values of the predictors of weighted centrality before the oil boom. We estimate the effect of the oil boom on the weighted centrality of the respective products as the difference in weighted degrees between the product and its synthetic version in the years after the oil boom took place. Computation of synthetic controls was done using a freely available software scripts that is written for R (Abadie et al., 2011).

4.4.3 Method cautions

Abadie et al. (2015) explain that constructing a donor pool of comparison units requires some care. First, units affected by the event or intervention of interest or by events of a similar nature should be excluded from the donor pool. In addition, units that may have suffered large idiosyncratic shocks to the outcome of interest during the study period should also be excluded if such shocks would have not affected the treated unit in the absence of the treatment.

Likewise, to avoid possible interpolation biases⁹, it is important to restrict the donor pool to units with characteristics similar to the treated unit. Another reason to restrict the size of the donor pool and consider only units similar to the treated unit is to avoid over-fitting. Over-fitting arises when the characteristics of the unit affected by the intervention or event of interest are artificially matched by combining idiosyncratic variations in a large sample of unaffected units. In our case, due to the nature of the data, this risk exists and must be taken into account when the results are obtained.

The applicability of the method requires a sizeable number of pre-intervention periods. The reason is that the credibility of a synthetic control depends upon how well it tracks the treated unit's characteristics and outcomes over an extended period of time prior to the treatment. We do not recommend using this method when the pre-treatment fit is poor or the number of pre-treatment periods is small. A sizeable number of post-intervention periods may also be required in cases when the effect of the intervention emerges gradually after the intervention, or changes over time.

⁹Interpolation bias is the failure to adequately control covariates within the observed range of data, for example, when applying a linear model to a non-linear relationship (Jones and Rice, 2011).

4.5 Results and Discussion

4.5.1 Evolution of the Economic Structure of Ecuador

In this section, I present the results obtained by the application of the synthetic control method in our data set and explore the effects that the 1972 oil boom had on the centrality of individual industries of the Ecuadorian economy¹⁰.

I grouped the results in 5 different figures (figures 4.6-4.10), each containing two sub-figures: on the left-hand side, there is the data of intermediate consumption per product between 1965 and 1985; on the right-hand side, the centrality measurements per year. Finally, I looked at each product individually and examined the impact of the 1972 oil boom.

Figure 4.5a illustrates how the intermediate consumption (IC) of different products had grown constantly over time. There are products that grew more than others, such as meat products in comparison with agriculture ones. On the other hand, if we look at the figure 4.5b, we can observe clearly that most of the products' centralities did not change significantly over time.

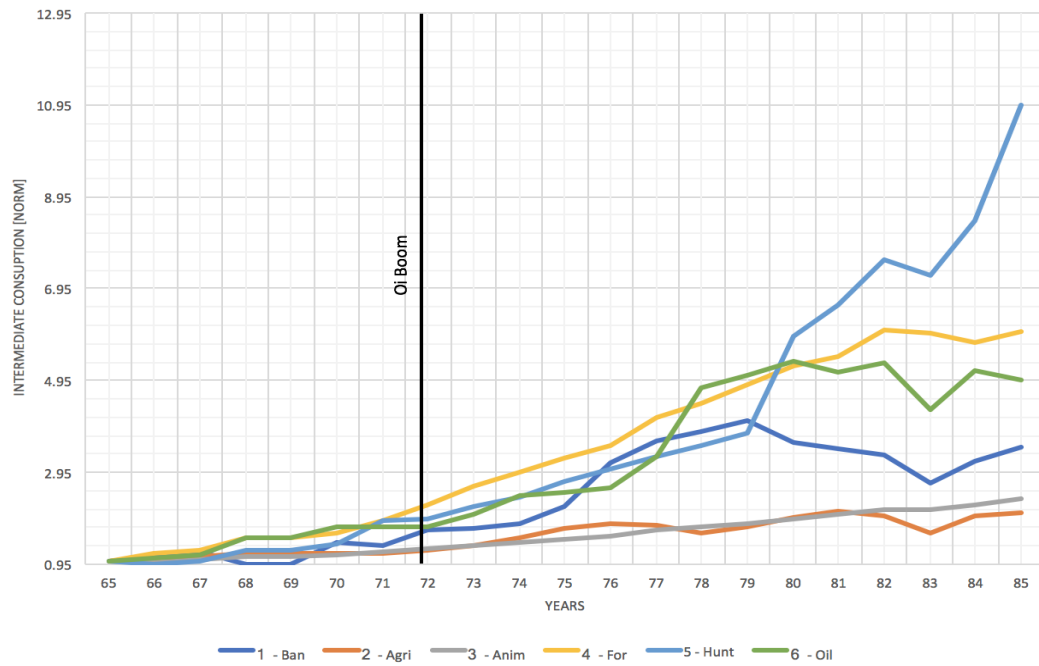
It can also be inferred that the centrality of crude oil grew substantially during the years before the oil boom (i.e., the observed rapid growth in figure 4.6b). This could be explained by the fixed gross capital investment into the sector from 1966 onwards, as preparation for oil exports by Texaco (nowadays, Chevron-Texaco) (Baquero, 2017). This meant that prior to the 1972 oil boom, the importance of the sector increased because of the investment that the sector had been absorbing.

Also, it is important to recall the differences in growth of the oil sector in figure 4.5a, in comparison with its centrality (figure 4.5b). In 1967 the centrality experienced an important growth, which was not reflected in the intermediate consumption. This shows us the differences that exist when we use aggregate measurements (intermediate consumption) rather than structural measurements (centralities). We will explore this further in the next section.

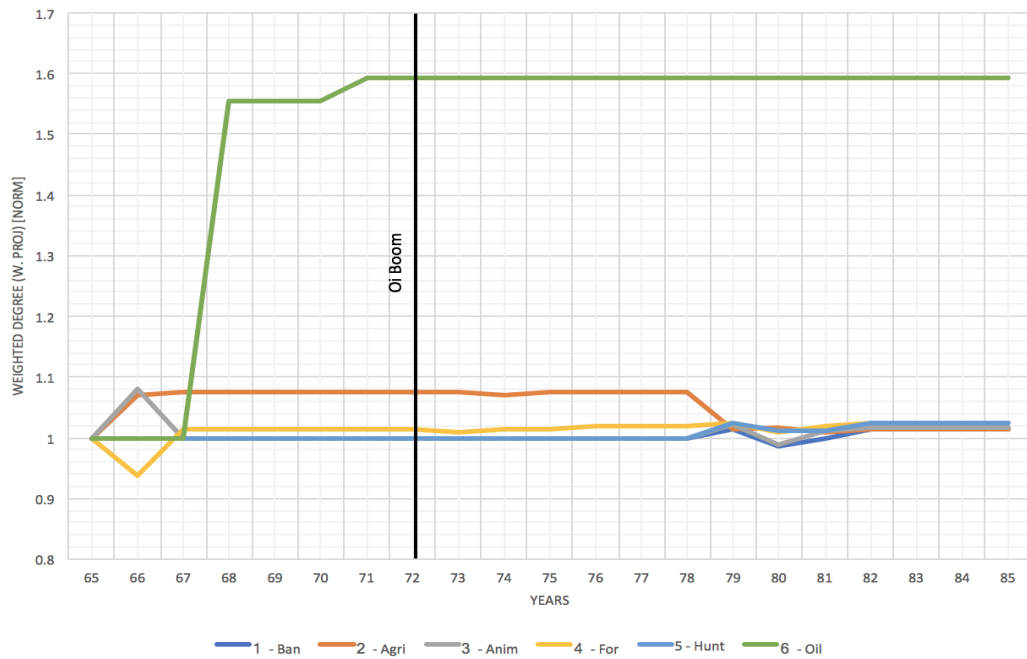
In figure 4.6a, we can observe that the IC follows a similar trend as the previous figures. In this case, refined petroleum and mining products are the ones that grew significantly compared with others that had initially remained constant (figure 4.6b).

¹⁰It is important to recall that in the analysis the informal economy is excluded. There is not data that quantified what was the share of informality by this time period. However, González and Cabrera (2018) estimated that 44.5% of the Ecuadorian economy was informal in 2016 and in previous decades the informality was higher

Figure 4.5: Evolution of the Ecuadorian Economy (1965-1985)



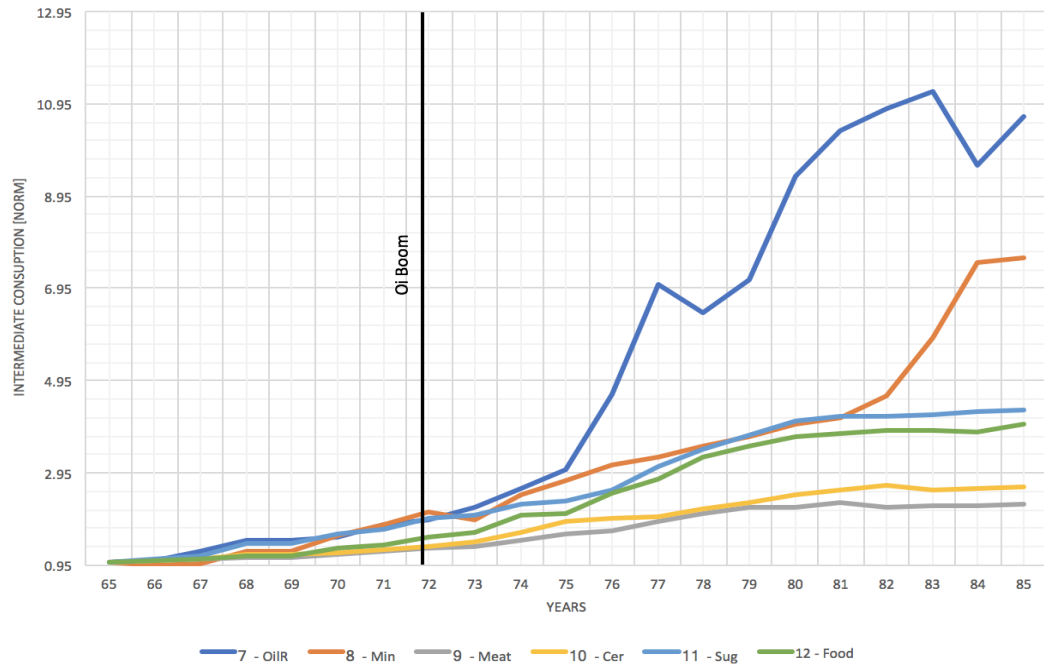
(a) Intermediate Consumption



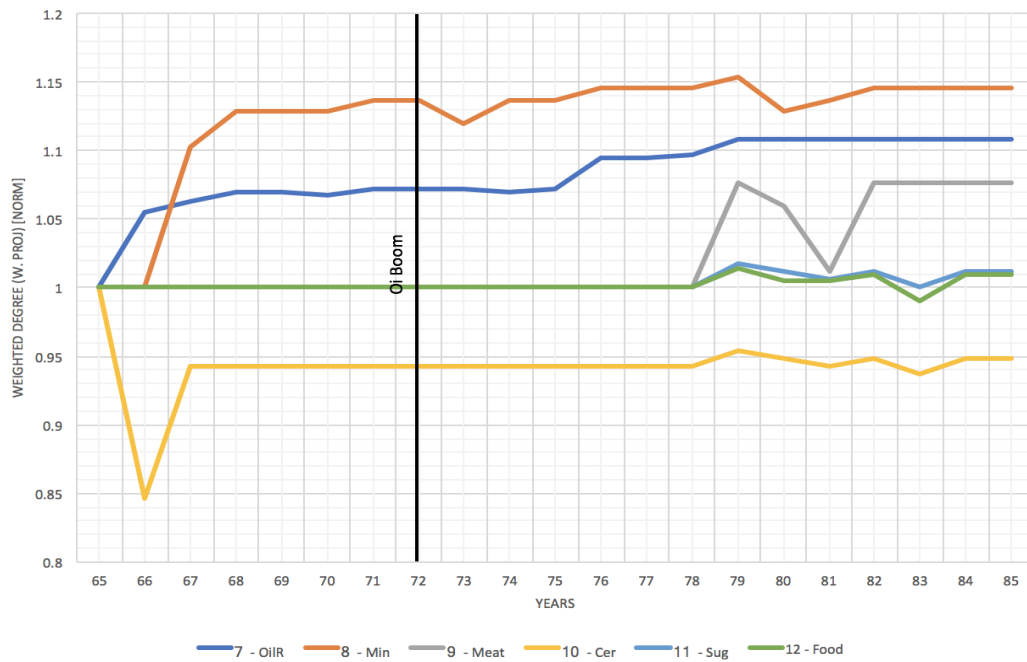
(b) Weighted Degree

The figures 4.7a and 4.7b, we can observe a similar trend for all sectors as in the previous figures, the exception being drinks (number 13) which did not increase its centrality.

Figure 4.6: Evolution of the Ecuadorian Economy (1965-1985)



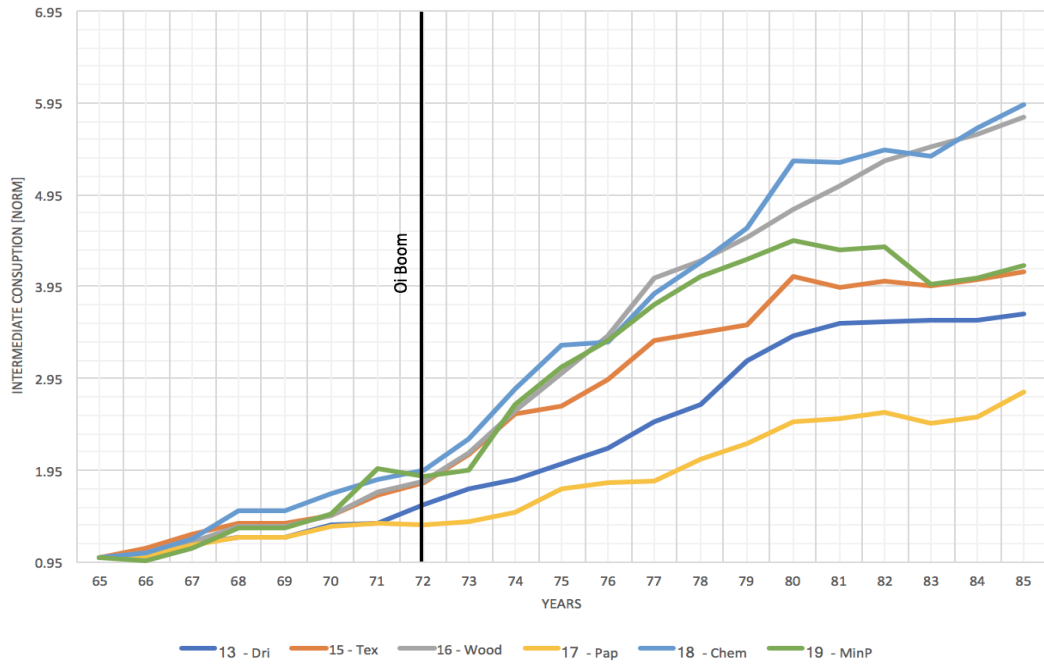
(a) Intermediate Consumption



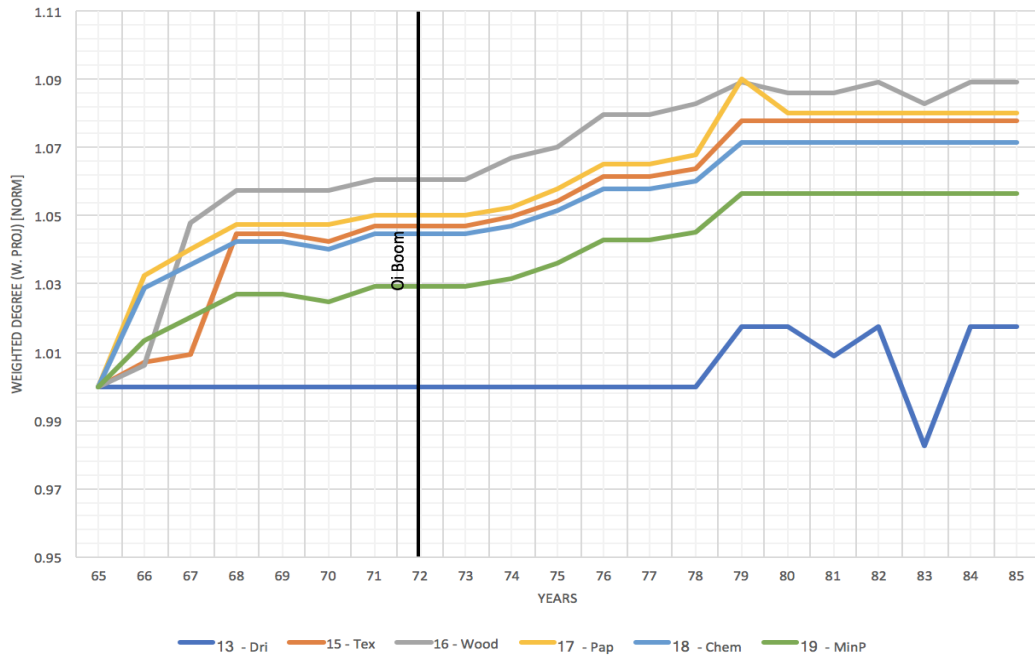
(b) Weighted Degree

The figure 4.8a, we observe a sustained growth of intermediate consumption in all the products. In the case of machinery, transport equipment and the construction sector, they experienced significant growth during the period of study. Similarly, manufacturing grew

Figure 4.7: Evolution of the Ecuadorian Economy (1965-1985)



(a) Intermediate Consumption

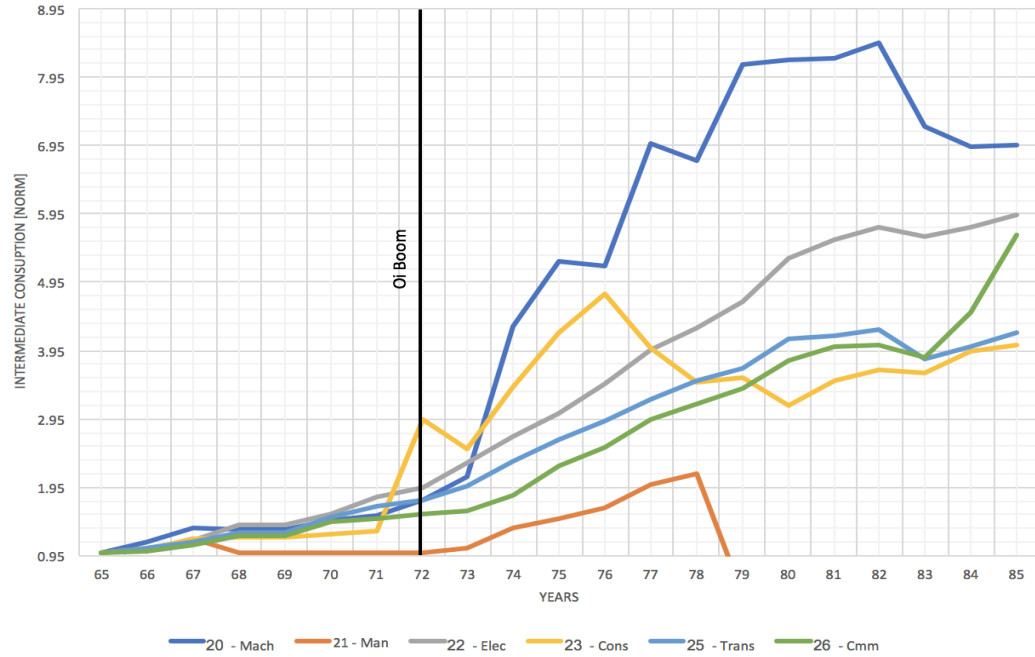


(b) Weighted Degree

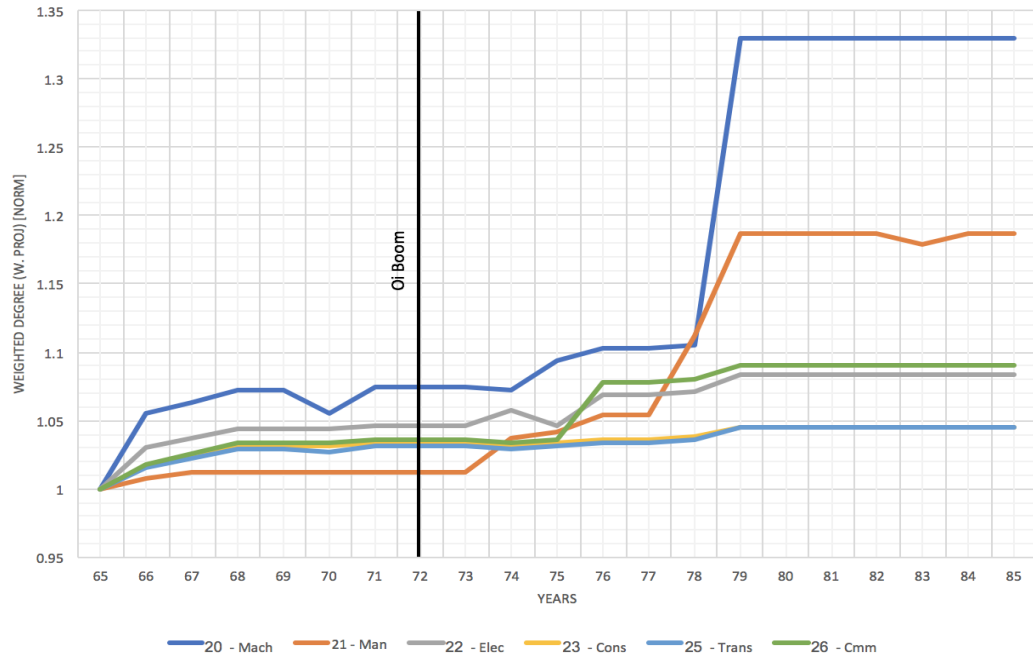
for a few years but collapsed later. All these observations are consistent with different economic literature in that most of the oil revenues were invested in the improvement of the infrastructure in the whole country (Belli, 1984; Acosta, 2006; Larrea, 2013). Similarly, in

figure 4.8b, we observe how machinery and transport equipment as well as manufacturing became more important for the Ecuadorian economy.

Figure 4.8: Evolution of the Ecuadorian Economy (1965-1985)



(a) Intermediate Consumption



(b) Weighted Degree

Finally, in figure 4.9a we can observe how the IC of different services have been growing

over time. This is because the economic expansion of an economy is linked with the growth in services such as water, electricity, business services and so on. However, when we look at the centralities in figure 4.9b, we cannot observe a clear common trend. We will analyse this in more detail in the following section.

4.5.2 Oil boom effects in the Ecuadorian productive structure

The analysis of the effects of the oil boom in the Ecuadorian economy is divided in four categories, all of which on a product basis: primary products or raw materials, basic secondary products, complex secondary products and services. For all figures in this section, the left-hand side contains the evolution of the centralities over time (solid line: observed centralities, dotted line: synthetic controls); the right-hand side figures represent the gap, as defined by the difference between the observed centralities and the synthetic controls. These figures allow us to capture the effect of the oil boom on different products in the Ecuadorian economy.

Primary Products

According to the Ecuadorian classification at their national accounts, the primary products can be analysed in six categories:

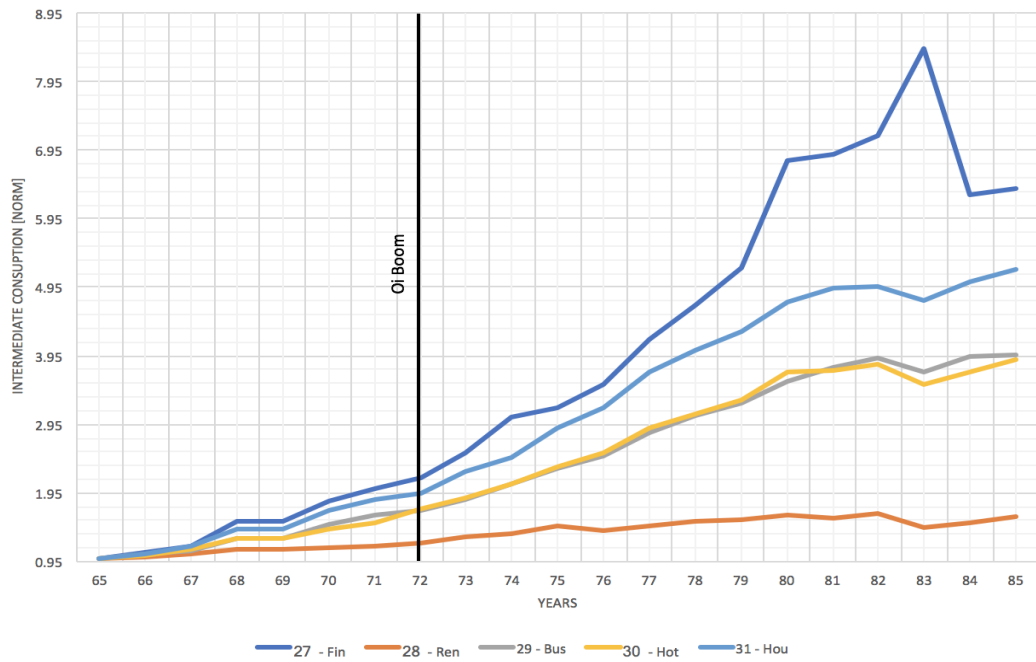
- a. Banana, coffee and cocoa beans
- b. Other agricultural products
- c. Animal products
- d. Forestry products
- e. Hunting and fishing products
- f. Oil and natural gas

In the case of ‘banana, coffee and cocoa beans’ (figure 4.10), we observe that the synthetically generated weighted degree (SWD) experienced a slight growth in comparison with the increase in the observed data, coinciding with the 1972 oil boom. The gap between the control and the observed data tells us that the oil boom could have limited the growth of this sector.

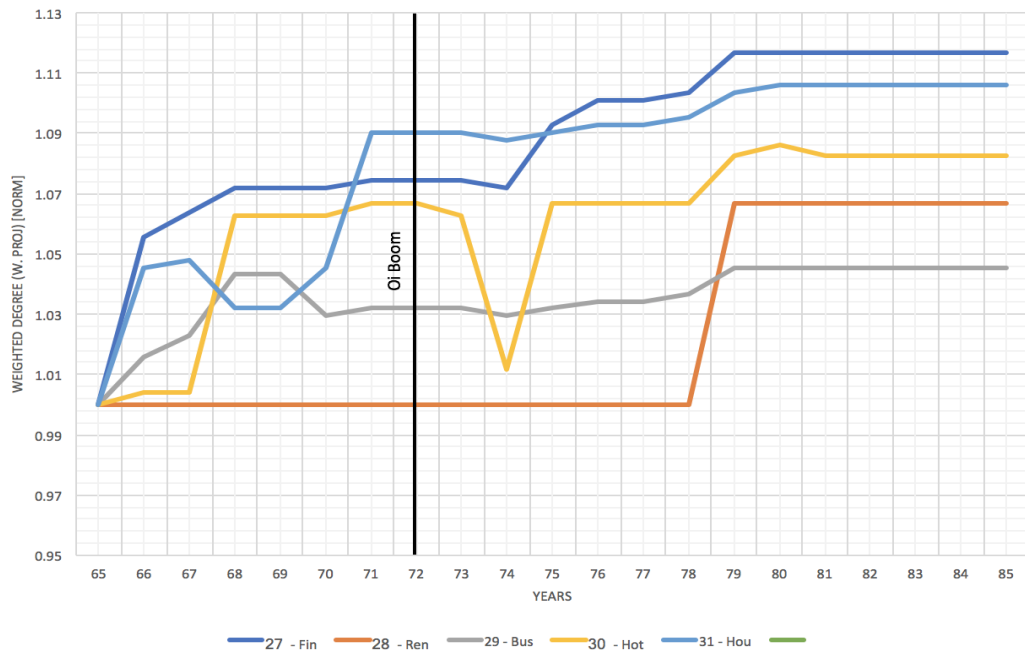
In the case of ‘other agricultural products’ (figure 4.11), the observed centrality did not show any change caused by the oil boom, although after five years the SWD experienced a small growth. We argue that the oil boom did not have an effect on the centrality of these products, since the observed centrality did not change.

In the case of ‘animal products’ (figure 4.12a), their centrality remains the same regardless of the oil boom and thus we argue that the oil boom did not have an effect on these products.

Figure 4.9: Evolution of the Ecuadorian Economy (1965-1985)



(a) Intermediate Consumption



(b) Weighted Degree

However, in figure 4.12b, according to the SWD, there was an expected growth only 3 years after the boom, which is not relevant for our study and will have to be examined in the future.

Figure 4.10: Banana, coffee and cocoa beans

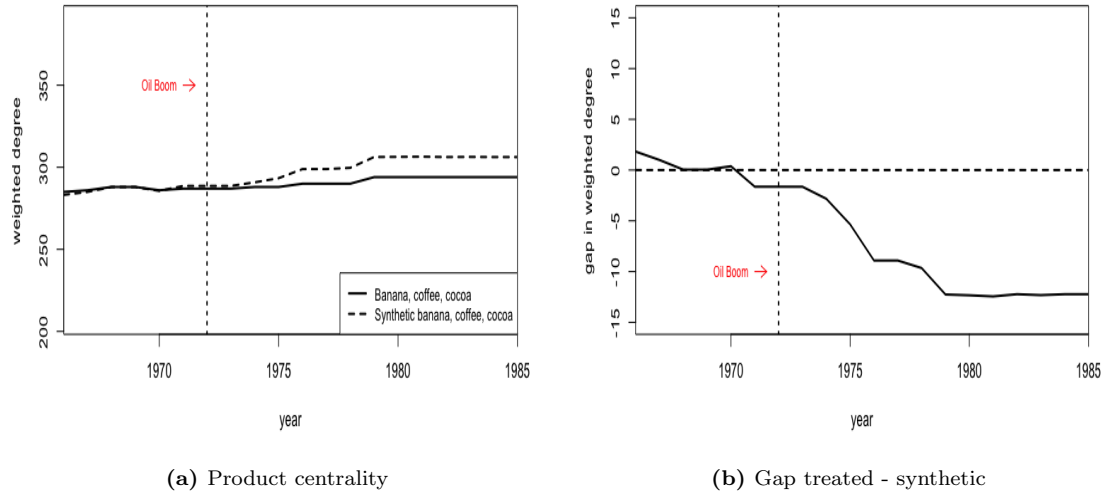


Figure 4.11: Other agricultural products

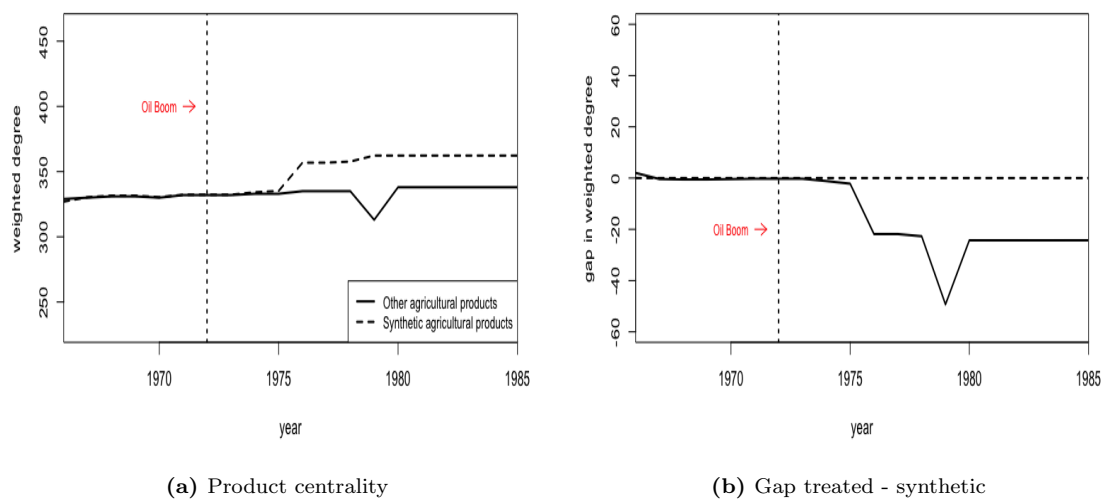
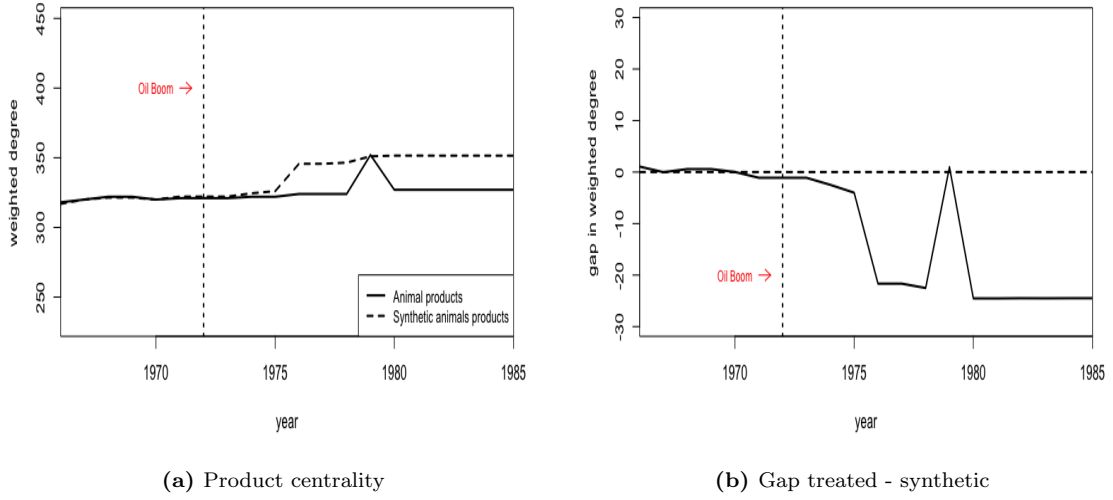


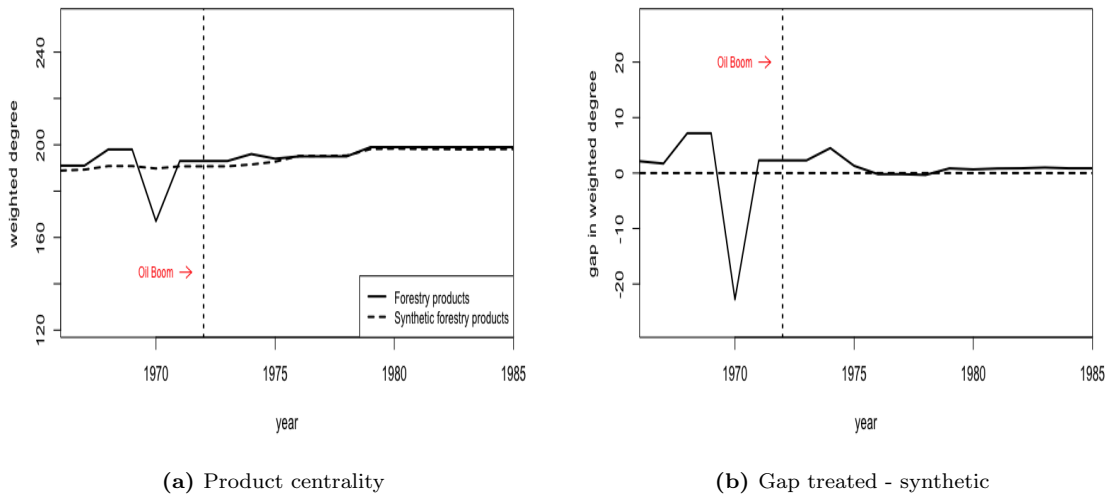
Figure 4.12: Animal products



In the case of forestry products (figure 4.13), we can observe that there are some changes on the centralities before the boom, but they lasted a short period. An explanation to this could not be found in the literature. Therefore, we cannot tell if this was caused by any event or whether there was an issue with the data. Nevertheless, for the matter of our analysis we can observe clearly that the centrality did not change due to the oil boom.

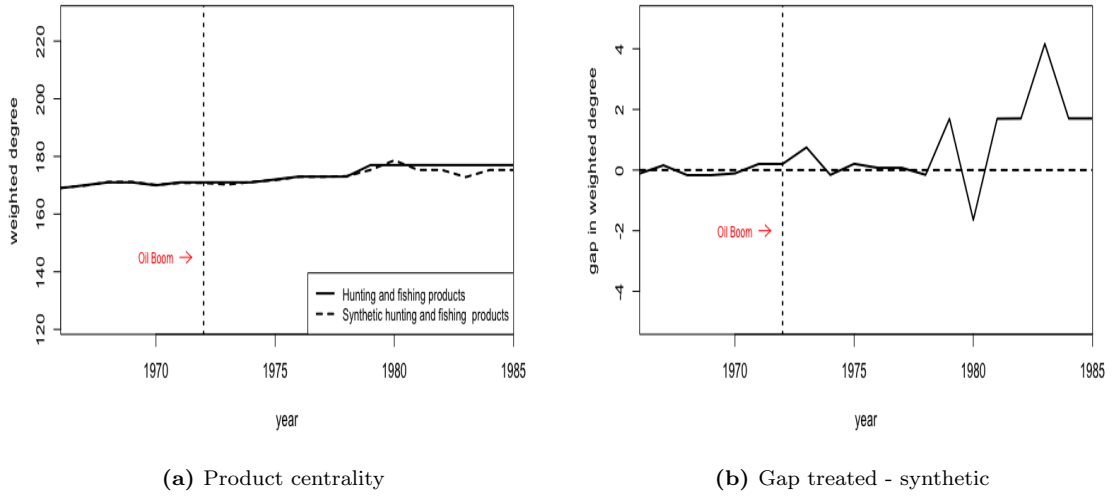
In the case of products related with hunting and fishing (figure 4.14), there is also not an apparent effect caused by the oil boom.

Figure 4.13: Forestry products



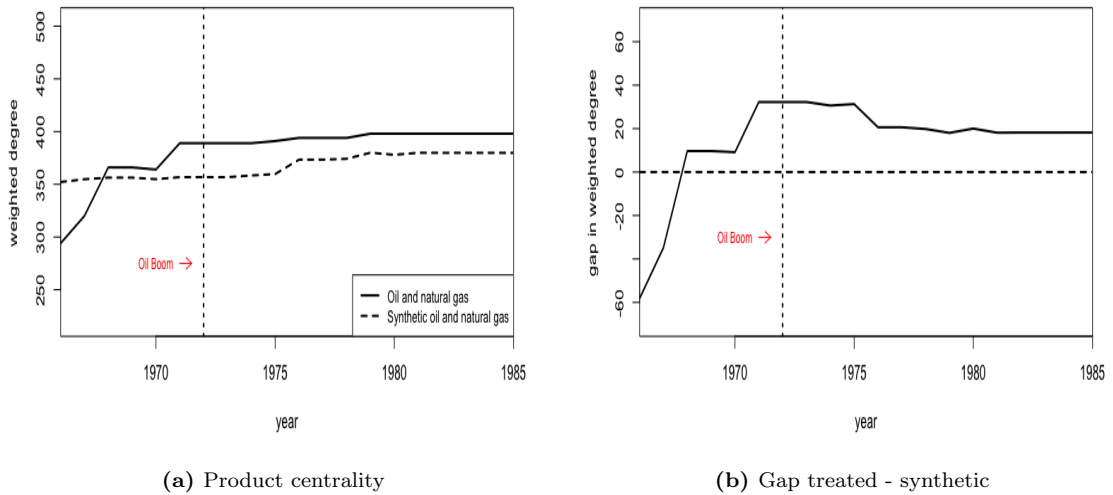
When we look at crude oil and natural gas (figure 4.15), we can clearly see an important change in their centrality. However, the changes occurred before the boom. As it was

Figure 4.14: Hunting and fishing products



explained in the previous section, this was caused by the investment that went into the area once oil was discovered. This investment triggered the subsequent oil boom.

Figure 4.15: Oil and natural gas



Basic Secondary Products

The 'basic secondary products' are those that have been transformed from raw materials into simple goods. In this case, we have the following products:

- Oil and refining products
- Other mining products
- Meat and processed fish
- Cereals and bakery

- e. Sugar
- f. Various foodstuffs
- g. Drinks

As it can be seen in figures 4.16 and 4.17, the products' centrality experienced small variations over time. However, the SWD projected a slight increase. In this case, we could argue that the effect of the oil boom constrained the increase of the products' centrality to some extent. However, there was not a significant impact on either product, overall.

Figure 4.16: Oil refining products

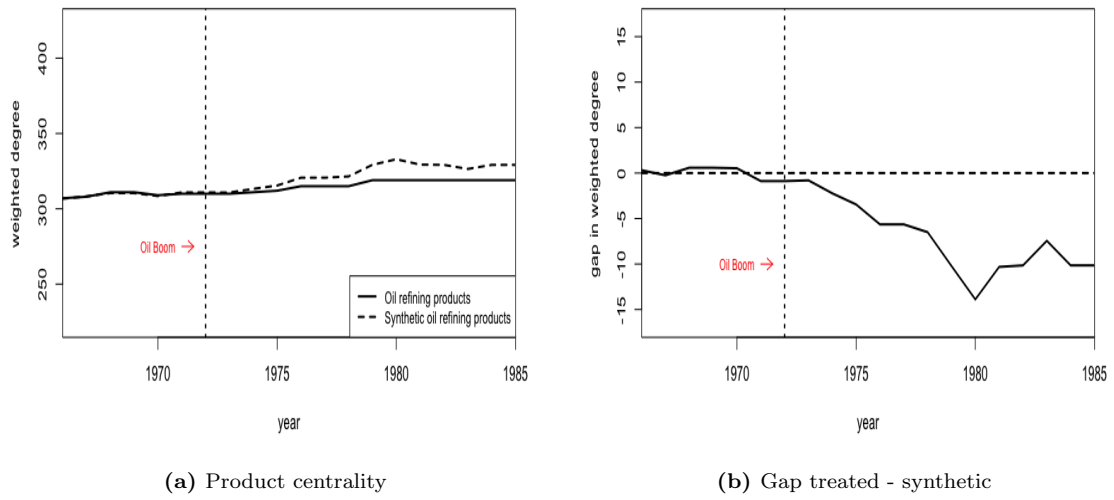
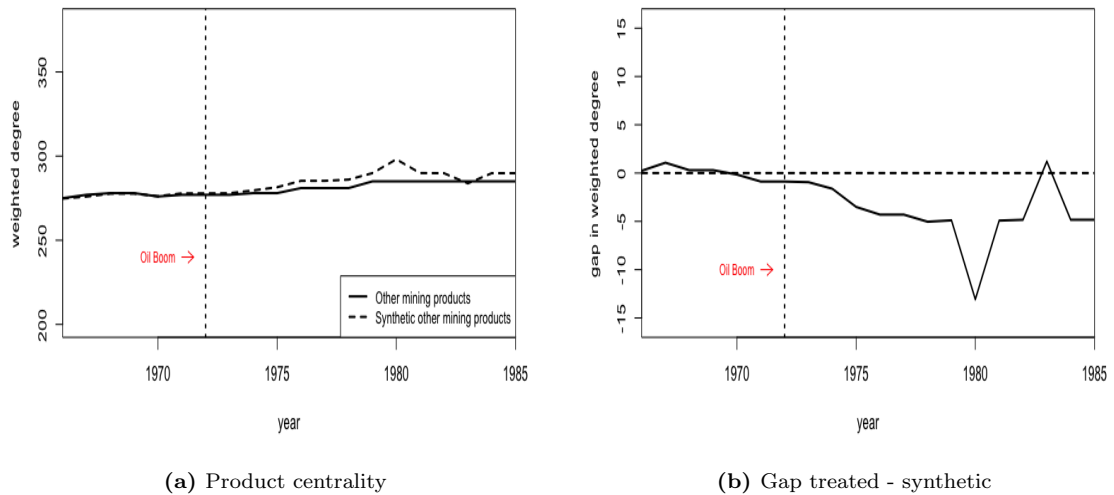


Figure 4.17: Other mining products



In the case of the products related with meat and processed fish (figure 4.18), cereals and bakery (figure 4.19) and sugar (figure 4.20), we can clearly see that the oil boom did not have an effect on these products. As the figures tell us, there are small temporal changes

on their centralities, but over time their centralities remained unchanged.

Figure 4.18: Meat and processed fish

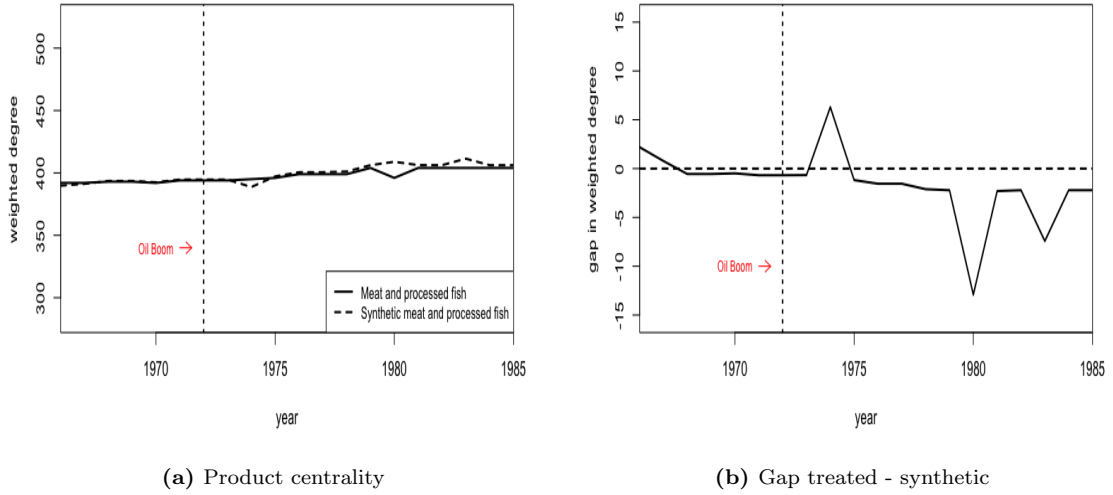
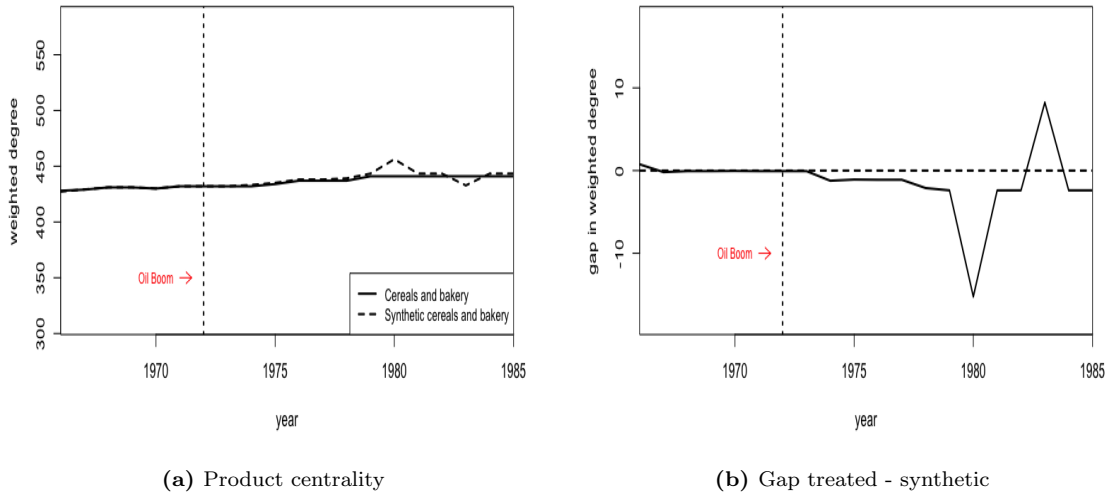


Figure 4.19: Cereals and bakery



Likewise, in the case of various food products (figure 4.21), we observe that the centrality did not change either. However, the synthetic control method used was not able to fit the data. Therefore, for this particular case, the methodology applied should be re-examined.

Finally, the oil boom did not have an effect on the centrality of drinks (figure 4.22). The same applies to cereals and sugar whereby the centrality remained the same.

Complex Secondary Products

The ‘complex secondary products’ are those that involve a more complex transformation process from raw or intermediate materials. In this case, we have the following products:

Figure 4.20: Sugar

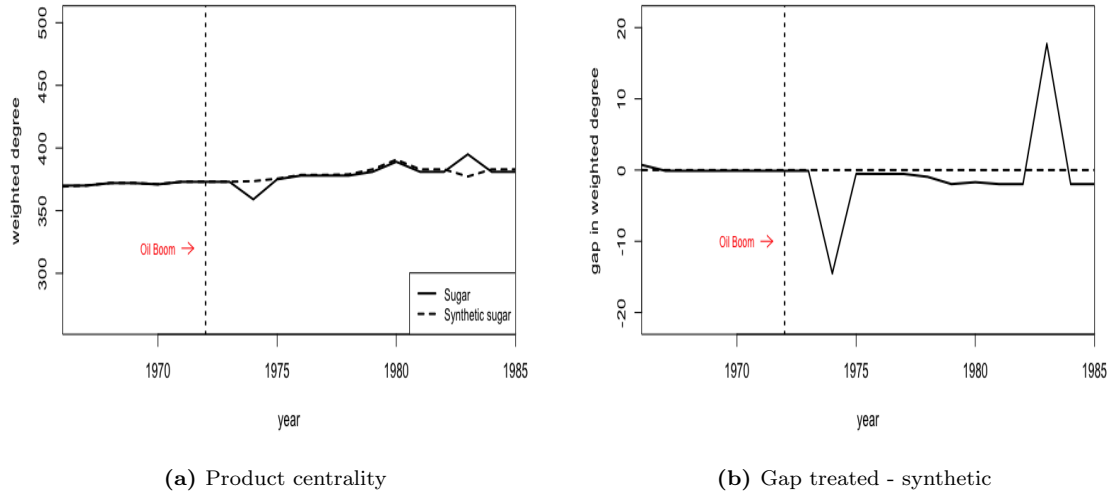


Figure 4.21: Various foodstuffs

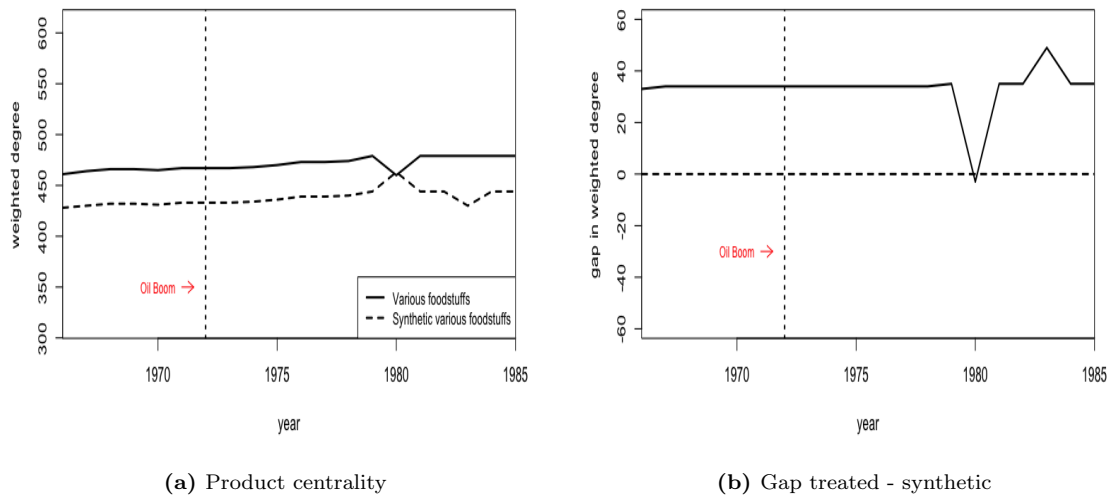
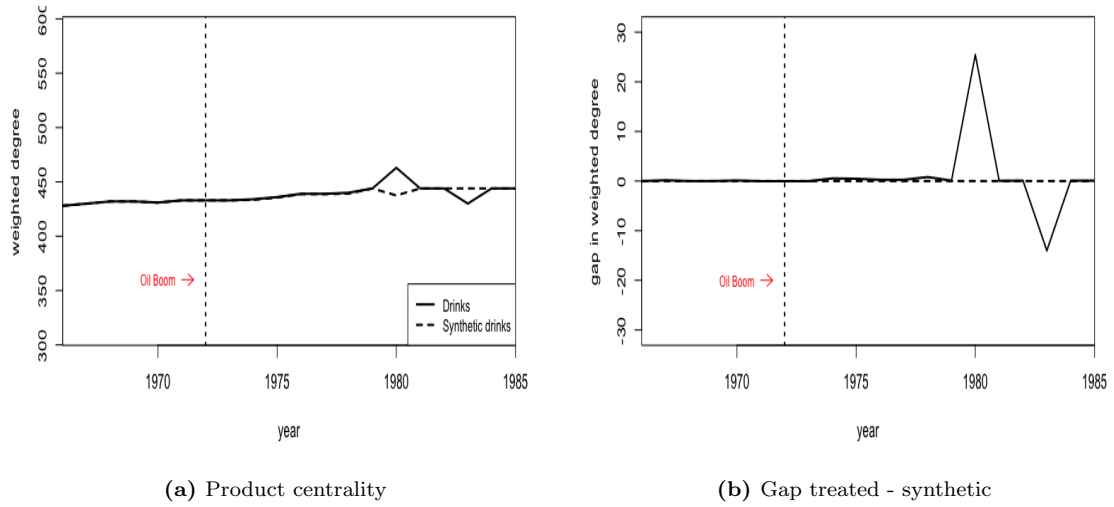


Figure 4.22: Drinks



- a. Textiles, clothing and leather products
- b. Wood
- c. Paper and printing
- d. Chemicals, plastics and rubber products
- e. Basic mineral products
- f. Machinery, transport equipment
- g. Other manufactured products

When we look at textiles, clothing and leather products (figure 4.23), as well as basic mineral (figure 4.27) and manufacturing products (figure 4.29), we can clearly observe an increase in their centralities although it did not happen at the same time as the oil boom, having occurred 3-5 years later. These findings are consistent with the literature in that the windfall of oil resources was used to promote industrialisation for import substitution. During this period, the oil revenues were invested in improving infrastructure as well as fostering the growth of the manufacturing industry.

In contrast with the sectors mentioned above, for wood production (figure 4.24) and paper / printing products (figure 4.25), we can observe a decrease in their centrality years after the oil boom. More research is needed to understand in detail what happened with these sectors during this period.

Finally, in the case of chemicals, plastic and rubber products (figure 4.26) and machinery and transport equipment (figure 4.28), we can observe an increase of their importance. This is associated with the inflow of oil revenues in the Ecuadorian economy, allocated in these

Figure 4.23: Textiles, clothing and leather products

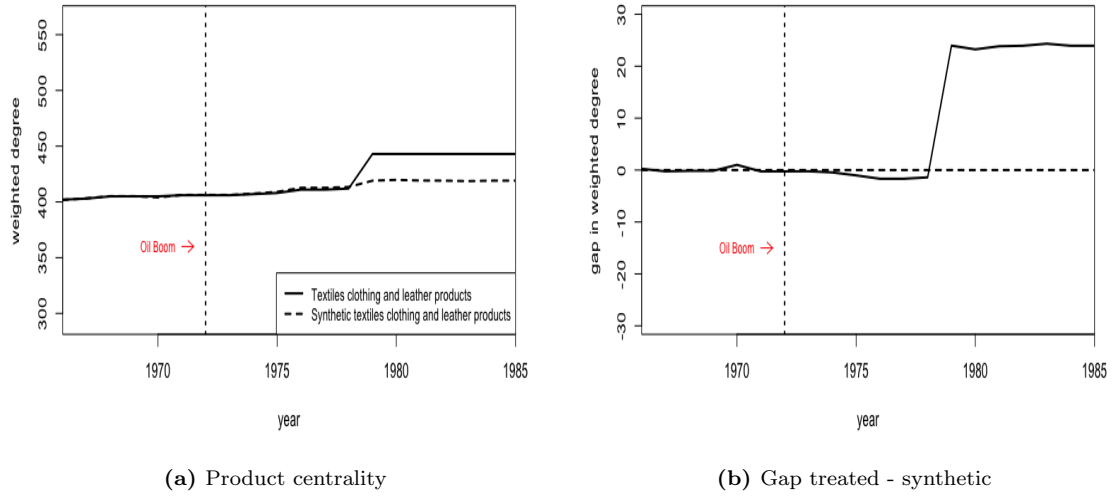


Figure 4.24: Wood

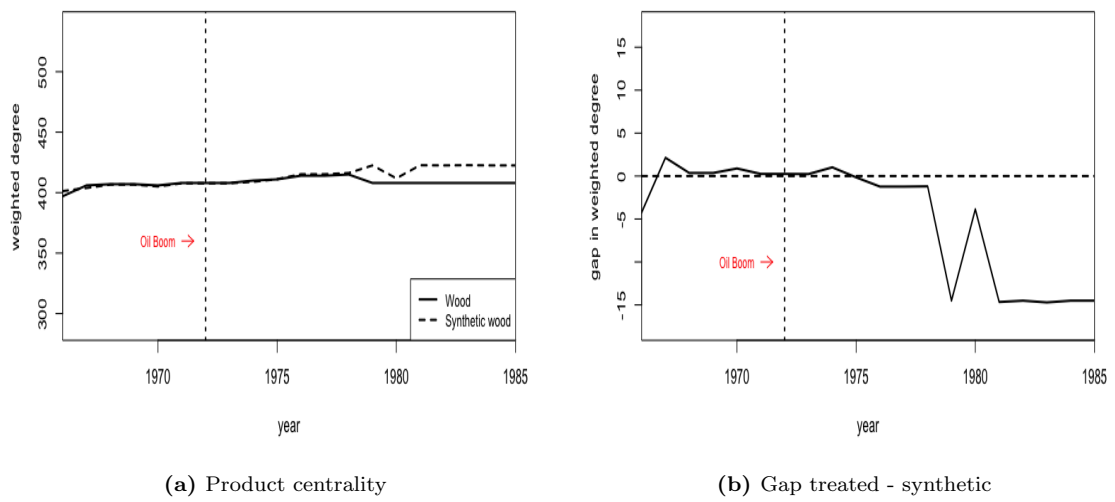
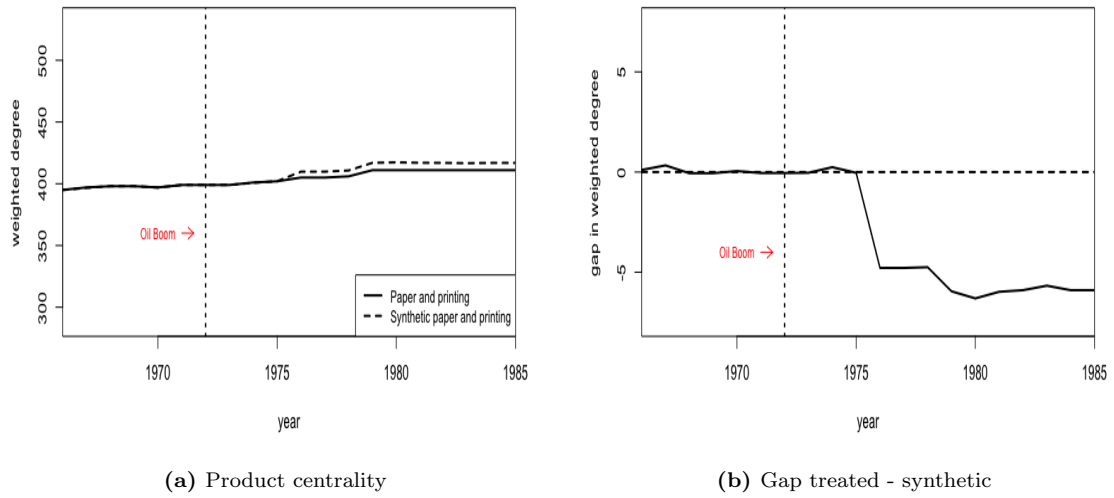
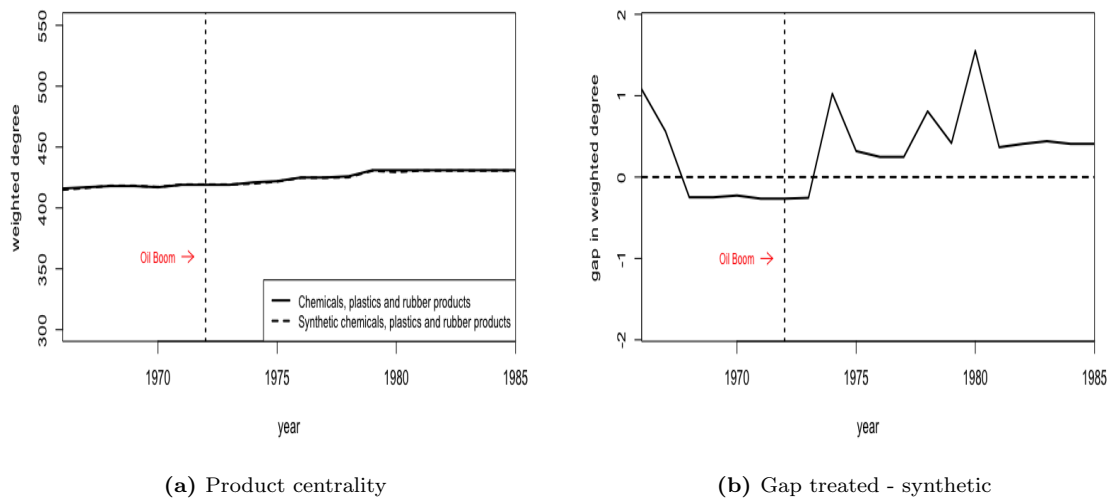


Figure 4.25: Paper and printing



sectors by the ISI policies in place, as the historical literature suggests. However, their importance did not grow as much as other manufactured goods.

Figure 4.26: Chemicals, plastics and rubber products



Services

These tertiary products involve the supplying of services to consumers and businesses, such as:

- Electricity, gas and water
- Construction
- Commerce
- Transport

Figure 4.27: Basic mineral products

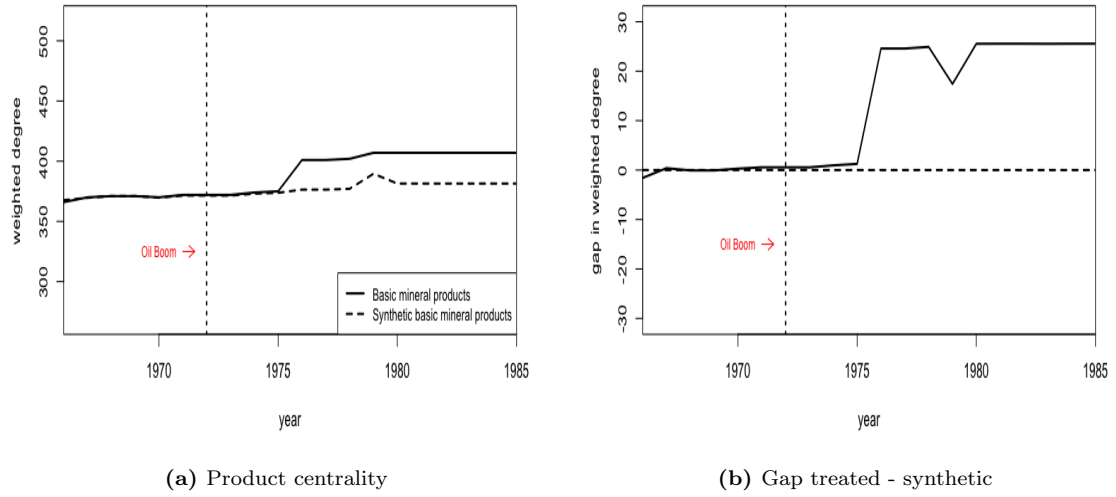


Figure 4.28: Machinery, transport equipment

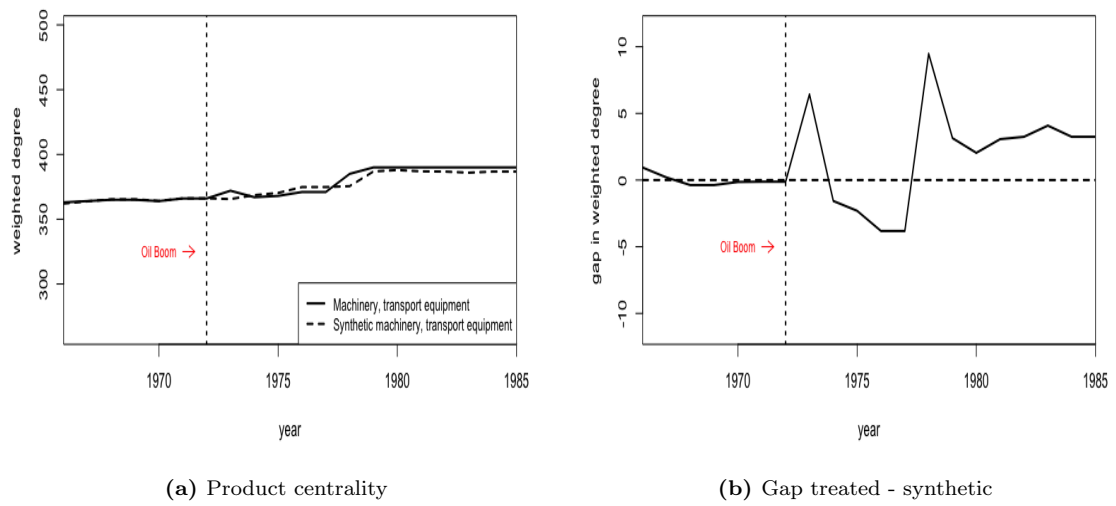
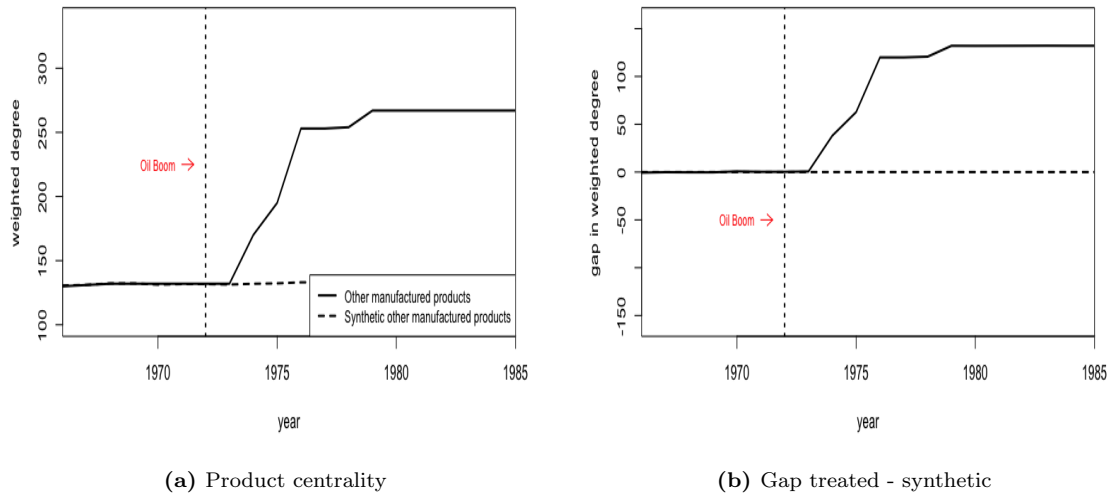


Figure 4.29: Other manufactured products



- e. Communication
- f. Financial services
- g. Rent housing
- h. Business services
- i. Hotels, bars and restaurants

When we look at services, we can observe an increase of the centrality of electricity, gas and water (figure 4.30) five years after the oil boom. This is expected, since the manufacturing centrality grew, as shown in figure 4.29, which requires the services of these industries. When we look at construction (figure 4.31), its observed centrality experienced a slight decrease in relation to the SWD. These results tell us that regardless of the increase in investment (shown in figure 4.8a) that went into this sector, the importance of the sector remained the same.

In the case of commerce (figure 4.32), we can observe a decrease in comparison with the SWD. However, this decrease is not directly associated with the oil boom.

In the case of transport (4.33), the oil boom did not have an effect and its centrality remains unchanged. However, there are three years where the centrality changes abruptly, which later on stabilised. In the case of communications (4.34), there is an increase of their centrality years after the oil boom. This could be related to the expansion of communication infrastructure that the country experienced in the 1970s.

In the case of financial services (figure 4.35), we cannot apply the synthetic control method because it is unable to replicate the observed data. The observed data tell us that the centrality had not changed.

Figure 4.30: Electricity, gas and water

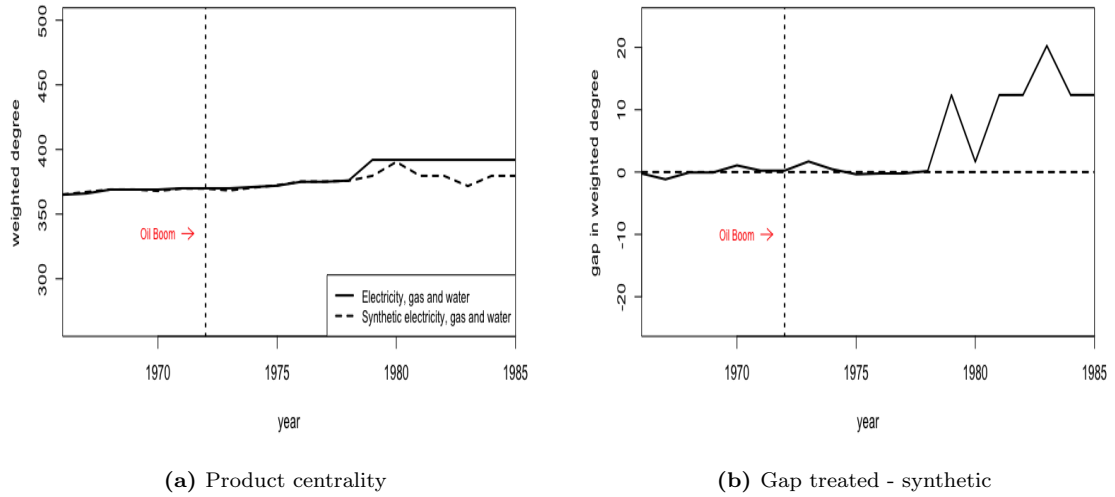


Figure 4.31: Construction

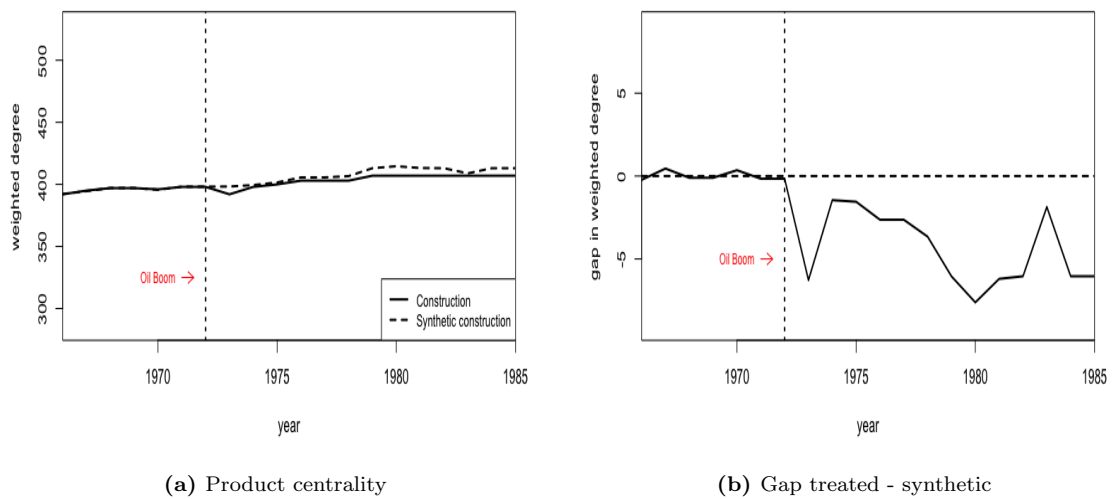


Figure 4.32: Commerce

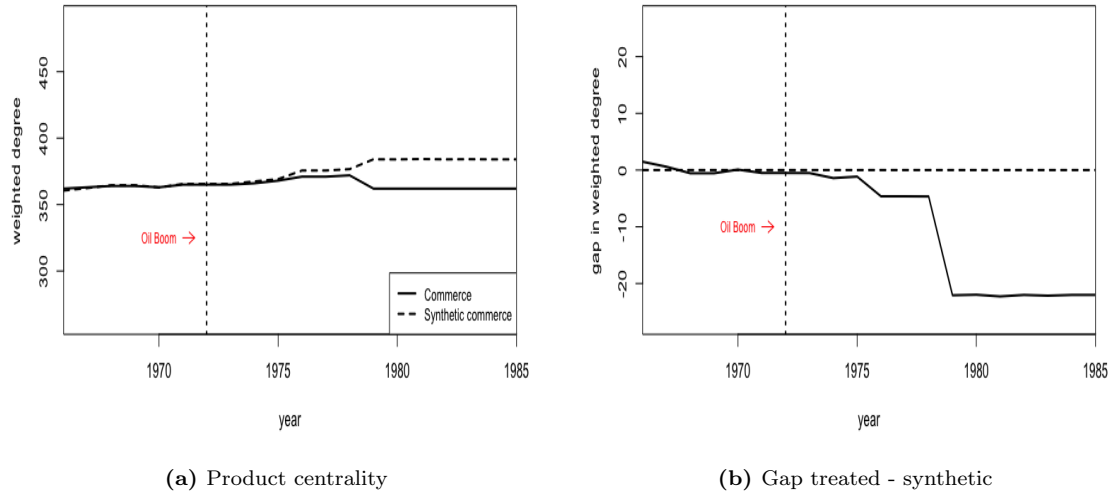


Figure 4.33: Transport

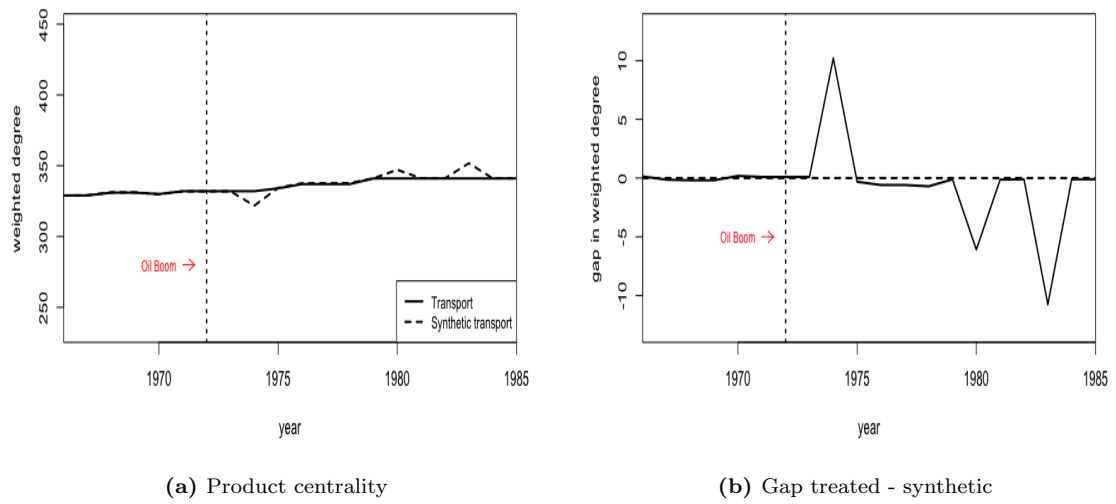


Figure 4.34: Communication

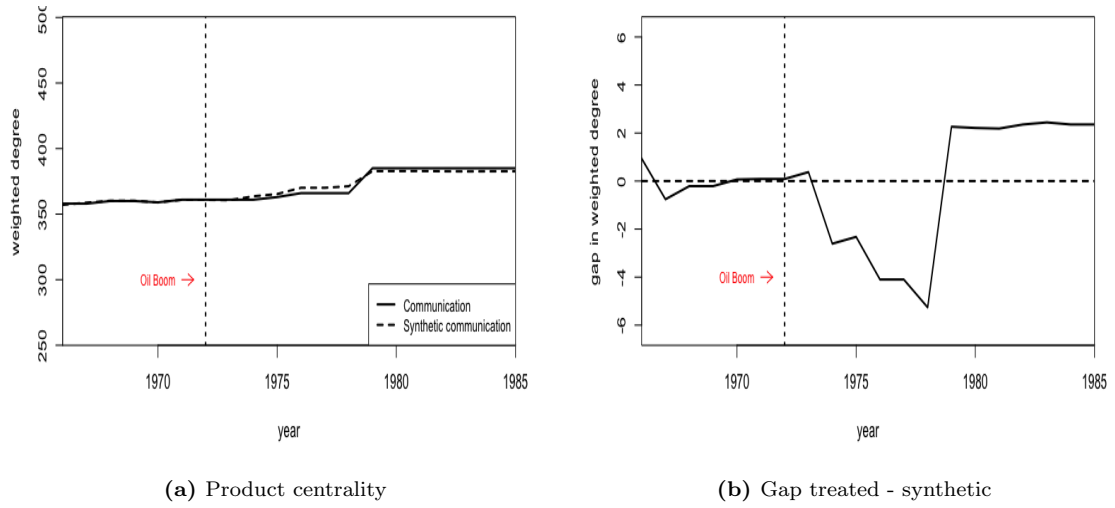
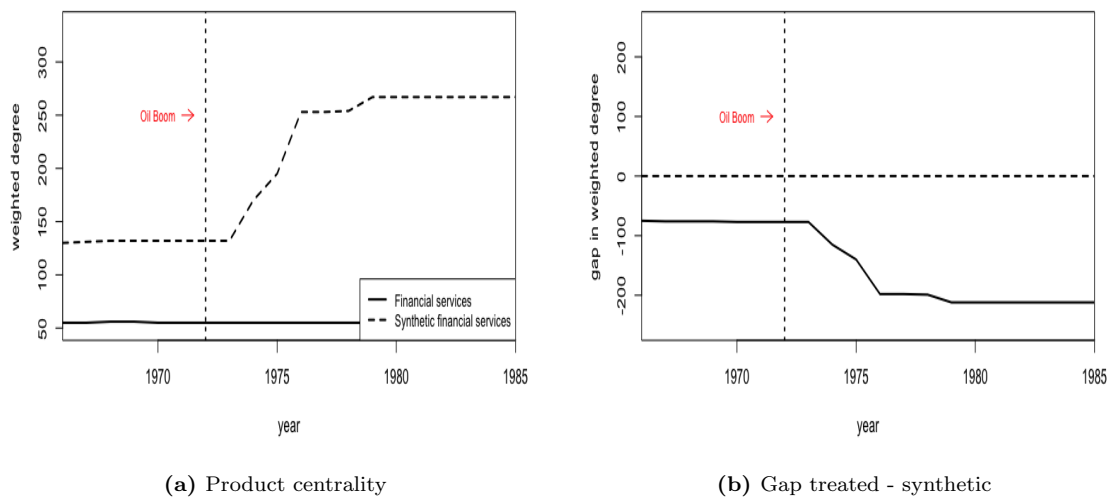


Figure 4.35: Financial services



Finally, in the cases of rental housing (figure 4.36), business services (figure 4.37) and hotels, bars and restaurants (figure 4.38), we can observe an increase in their centralities. This is expected because the new oil revenues encouraged consumption several years after the oil boom (Acosta, 2006, 2011; Correa, 2012).

Figure 4.36: Rental housing

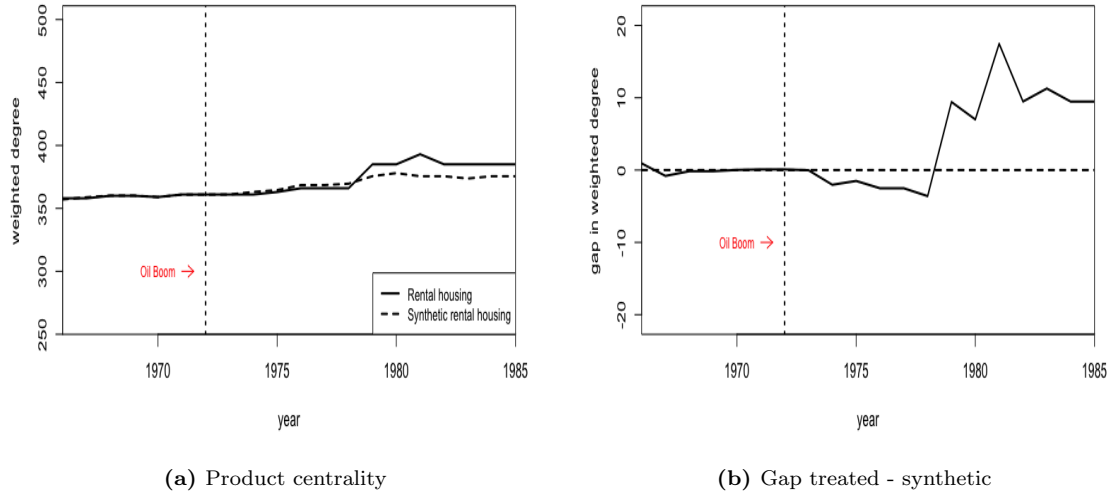
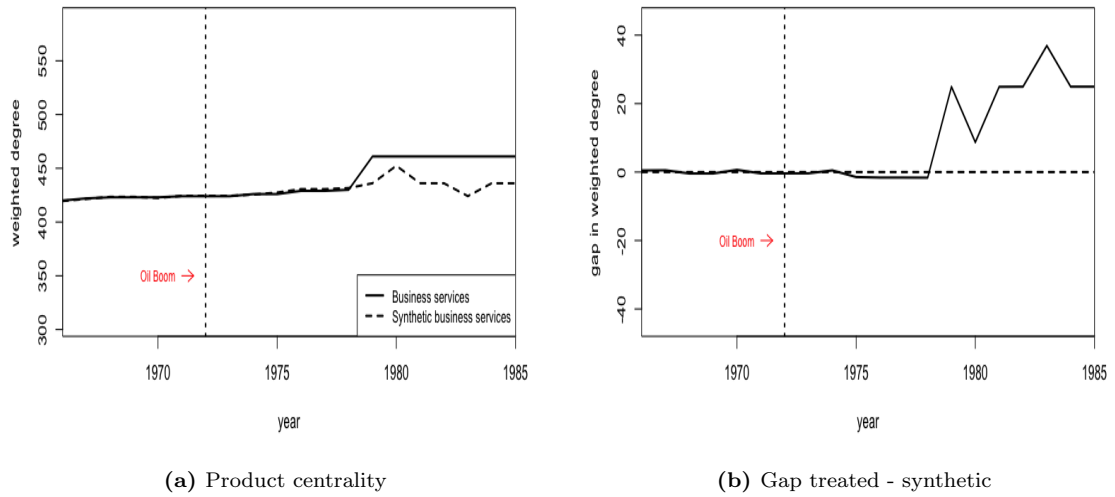


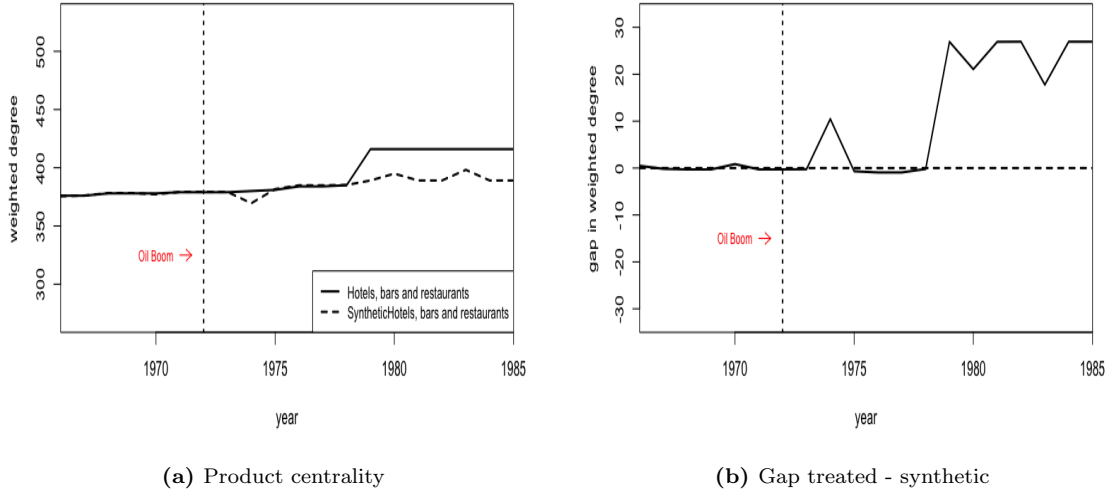
Figure 4.37: Business services



4.6 Conclusions

The aim of this chapter was to examine if the oil boom that Ecuador experienced in 1972-1973 had an impact on its productive structure. In other words, I attempted to identify if the importance of industries, measured by their centralities, changed due to the windfall of resources that came from this boom. To do so, a methodology was proposed to explore this

Figure 4.38: Hotels, bars and restaurants



effect and I hypothesised that oil boom would have a negative effect on the centralities of the non-traded sectors. Hereunder, I present two sub-conclusions: the former relates exclusively to the methodology proposed in this chapter and the latter relates to the specific results.

4.6.1 Methodological conclusions

The synthetic control method is a valuable tool to understand the effects of an intervention in comparative case studies. Its main objective is to estimate what would have happened to the treatment group if it had not received a given treatment. There are several elements that are key to produce reliable results using this method (Bouttell et al., 2018), among the most relevant for this study are:

- The method needs a large number of treated and control units and a considerable amount of observations pre-treatment for the controls generated to be more reliable.
- It is essential that the donor pool contains units that are similar to the treated unit but that are not affected by the intervention or whose levels of exposure to it differ notably.
- The treated unit or the potential control units in either the pre-intervention or post-intervention period can not be affected by other ‘shocks’ that might affect the outcome of interest.

These elements limit the applicability of SCM for this study case. First, as it was observed in the methodology section, the pre-intervention period was shorter than the post-intervention. The former was seven years and the latter twelve years. Even if the literature does not specify what is considered a good period of study, especially in relation to the pre-intervention

period, seven years could be a short time for the study. Second, in contrast to other interventions or treatments that are well-defined and precise in time (for example, the approval of a regulation), the oil boom is not well-defined and its effects are not precise. Third, the Dutch Disease and Resource Curse literature explain that the effects of an oil boom could take place after the resource discovery and prior to starting production (Cust and Mihalyi, 2017), mainly due to the large investments that are directed to develop the industry and the anticipation of the prosperity that the discovery causes. This phenomenon is observed in our data: the centrality of the oil sector increases a year before the actual boom. Therefore, the treated units are already affected by the intervention before it takes place. Fourth, as explained in the methodology, the donor pool entails the sectors that were minimally affected by the boom. However, due to the interconnections of the economy, the donor pool could have been affected by the intervention.

In this regard, I conclude that the Synthetic Control Method could be a valuable method for comparative studies in social sciences, including economics, although it proved to be unsuitable in the context of this analysis. First of all, the oil boom effects can not be easily isolated from the rest of the economy and it will be a source of endogeneity when trying to analyse its impact on other economic sectors. The SCM is used when the interventions are precise, whereas the oil boom effects take place even before the production of oil starts. As a way to overcome these issues, the donor pool could be constructed using data from countries that did not experience an oil boom. However, similar data (Use Tables) and similar metrics (as used in this thesis) must be calculated, and finding these data for several countries and for the same time period could be difficult. Based on the author's experience, this data is hard to access because it is stored physically in national archives and would require time to make it digital. Also, the methods to compile the Use Tables in other countries might not be consistent with the ones used in Ecuador.

4.6.2 Research conclusions

In spite of the unsuitability of the SCM to get insights about the effect of the oil boom in Ecuador, the analysis of the observed data (time series) reveals the centrality changes over time in the different economic sectors from 1965 to 1985. In this regard, the results show that the oil boom did not produce a visible change in the Ecuadorian productive structure, and the centralities of most of the industries were not affected by the boom, since most of them remained constant over time. However, there are a few industries that increased their centralities, which might not be directly related to the oil boom. The effect that they experienced was simply fuelled by oil income, but accompanied by a clear strategy of industrialisation by import substitution. The results obtained with the methodology

proposed are coherent with what the economic history of Ecuador tells us for the period under analysis. In more detail, it can be argued that:

First, one of the most important observations that we can make is that the centrality of primary products, which generate most of the income for the country, does not change over time. Many of them remained constant regardless of the oil boom. This finding is consistent with the Dutch Disease literature, which tells us that non-oil trade sectors are not affected in aggregate terms by a resource boom.

Second, it is important to recall that the centrality of crude oil grew significantly some years before the oil boom. Literature from this period explains that this change was caused by the investments triggered by Texaco's oil discoveries in the country (Belli, 1984; Vos, 1987; Larrea, 2013). Gelb (2011) explained that these effects are expected because the sector usually gains importance during the construction phases of the infrastructure needed for extracting the resource.

Third, data also shows that the centrality of manufacturing, machinery, construction and transportation increased in the years that followed the boom. The existing historical literature tells us that these changes were associated with industrial policy and ISI reforms implemented by the Ecuadorian Government after the oil boom. The new income that came from the boom fuelled public investment in these areas.

In the previous chapter, we observed that the amount of links that the oil industry has with other primary products is minimal and the sector itself has a small number of inter-linkages with the rest of the economy. For this reason, the changes in this industry due to the oil boom do not have strong effects on the other non-oil traded products. However, there are a few sectors that are more connected with the oil sector, such as the manufacturing and construction industries. Since these sectors depend directly on the oil revenues inflows, the oil boom and the variation in oil prices affect them directly. Thus, a network perspective is useful for understanding these interconnections (i.e., the industries inter-dependencies).

Finally, in Chapter 3 and Chapter 4, we observed clearly the productive structure of Ecuador in the past and the changes it experienced caused by the oil boom. Also, we observed how the country's oil dependency have prevailed since the 1970s. The limitations and recommendations of this specific chapters will be found at the concluding chapter of this thesis. In the next chapter, using a new and more comprehensive dataset, we will explore how this dependency could be overcome, which is the most relevant 'development' question for Ecuador and many other countries.

Chapter 5

Regional Economic Diversification and Industrial Relatedness

5.1 Introduction

As stated in the introduction of this thesis, oil and mineral export-dependency has been a long-standing issue for many developing countries. For many years, policymakers, academics and other experts have advised that diversifying their economies should be a priority, since relying on the income generated by few commodities can have detrimental effects on their economic growth and performance in the long run (Gelb, 2011; Ross, 2013). Moreover, empirical research has shown that economies with a more diverse productive base usually perform better over time, compared with their counterparts with limited products to export (Gelb, 2011; Lederman and Maloney, 2007; Hesse, 2009).

There are many benefits of having a more diverse economy. Many scholars explain that having a wider variety of products to export protects these countries from the effects of commodity price volatility. This is particularly so in the case of crude oil (Auty, 2001; Collier and Page, 2009; Lederman and Maloney, 2006; Ross, 2013; Papyrakis and Pellegrini, 2019). Productive diversity could help economies to create new jobs that are less dependent on the resource sector, and in this sense employment opportunities will not rely mainly on the rents created by reliance on natural resources. Moreover, economic diversification could help them ward off the resource curse (cf. Chapter 1, 2), and in the long run a more diverse productive base can prepare oil exporting economies to face future resource depletion or even to reduce the risks associated with carbon stranded assets (Caldecott et al., 2014).

Despite the importance of diversification, Ross (2019) stresses that little is known about how these economies can achieve it. In fact, existing literature on natural resource dependency has focused on the challenges that oil exporting economies face, having failed to shed light

on how to overcome their dependency on few commodities. There has been limited research on which aspects of diversification are important, the conditions under which diversification is desirable, and how best to promote it (Ross, 2019). He argues that scarce, missing or unreliable data regarding these countries is one of the reasons of the limited research on this topic (Ross, 2013). He further states that if data related with export diversification exists, it can be noisy or uninformative¹, which makes it difficult to know the true level of diversification in the domestic economy (Ross, 2019, 2013).

Understanding how countries or regions develop new growth paths and develop new economic activities (diversification) is one of the key research interests in Economic Geography (Balland et al., 2020a; Boschma et al., 2017; Hassink et al., 2019; Hidalgo et al., 2018; Isaksen, 2014). This means that the probability that a region will develop a new activity - exporting a new product, creating a new industry, or developing a new technology - increases with the number of related activities present in a location (i.e., co-location relatedness) or with the presence of that activity in neighbouring locations (Alshamsi et al., 2018; Hidalgo et al., 2018; Balland et al., 2019; Mealy and Teytelboym, 2020; Whittle and Kogler, 2019; Jun et al., 2019). Furthermore, research has shown that, regions at lower stages of development, as indicated by lower GDP, tend to diversify more frequently into related activities. However, when they are at intermediate stage of development, countries start diversifying towards more unrelated activities, which in turn can generate greater economic growth (Pinheiro et al., 2018; Alshamsi et al., 2018).

Drawing upon these ideas in Economic Geography, this chapter addresses the third research objective by exploring if the relatedness principle based on co-location holds true and whether it is a driver of regional diversification in the context of an oil-exporting-country². Furthermore, it has been explained that production and industrial activities are part of bigger networks that are fragmented and spatially dispersed (Alfaro et al., 2019; Coe and Yeung, 2015). In this sense, this chapter extends previous analyses and attempts to understand how production chains influence the emergence of pioneering industries in a given location. More specifically, it explores the idea of relatedness based on suppliers and clients networks at the industrial level within a city. In other words, it tests how the relations that exist between suppliers (supply chain) and consumers can explain the emergence of pioneering industrial activities in a particular place.

¹Ross (2019, p. 792) explains that "some oil-dependent countries report [for instance] impressive growth in their manufacturing exports because their governments classify refined oil products as 'manufactured goods'"

²Most of the existing literature on this topic focuses on developed countries. It is important to explore if these ideas hold true and to make these distinctions because studies have pointed out that agglomeration economies and human capital externalities might differ between developed and developing countries (Chauvin et al., 2017).

To accomplish these goals, this chapter uses a large and unique private data set, which was provided by the Internal Revenues Service (SRI) of Ecuador, under a Non-Disclosure Agreement. The data set comprises fine-grained information of all the legal (firms) and natural persons registered in the country from 2007 to 2015 (approximately 150,000 entities per year). It also includes records of every transaction between all these entities (approximately 7 million observations), their respective locations at the municipality level, and other characteristics. Using this data, I calculate several well-known relatedness indicators based on existing literature (relatedness density for instance) and propose two new, which incorporate the networks of suppliers and clients. Using logistic regressions and machine learning techniques, I assessed the likelihood of a pioneering industry, which serves as an indicator of diversification, to emerge in a given location.

The chapter starts with a brief overview of the existing literature on economic diversification in oil exporting economies (Section 5.1) and the role that relatedness plays in regional diversification (Section 5.2). Subsequently, a description of the data used will be made, accompanied by descriptive statistics, and an explanation the econometric and machine learning models proposed for the analysis. In the last section, the main findings will be reported, focusing in particular on the oil/mining sector (extractive industries) and their comparison with manufacturing. Finally, key takeaways will be drawn.

5.2 Literature Review

5.2.1 Understanding Economic Diversification in Oil Exporting Economies

For over five decades, scholars have argued that primary commodity production and exports offer limited development opportunities (Djimeu and Omgba, 2019). Raul Prebisch (1950), Hans Singer (1950a) and more recent academics such as Harvey et al. (2010b) noted that there is a long-run downward trend of the terms of trade between commodities and manufactured goods, which would have a strong negative impact on commodity-dependent economies. These authors also argue that technological improvement, combined with the emergence of new suppliers and the development of several policies in developed countries, has contributed to a rapid expansion of world commodity supplies, affecting others that have historically been reliant on a few commodities. An example of the former was the fast rise of the United States to become one of the world's top oil exporter and its impact on global oil prices (Blas, 2018).

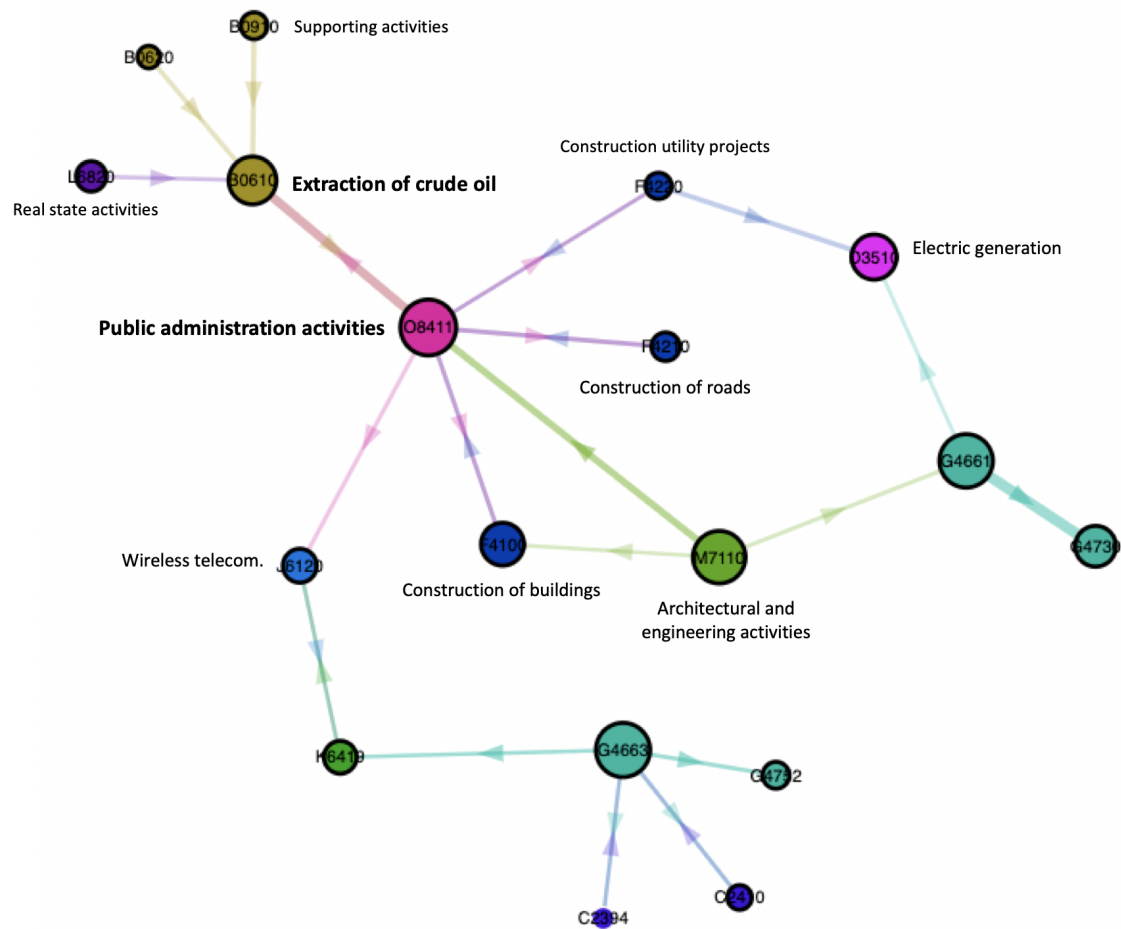
On the other hand, many studies have shown the benefits of having a more diverse economy. Firstly, scholars explain that having a variety of products (mainly non-commodities) to export protects countries against macroeconomic shocks caused by commodity price volatility

(Auty, 2001; Collier and Page, 2009; Lederman and Maloney, 2006; Ross, 2013). This is specially the case of crude oil since its rents tend to be highly concentrated and largely realised in the form of fiscal revenues, resulting in an over-reliance on state investment. In these cases, state investments feed other industries but when oil revenues stop flowing the rest of the economy slows down. In the case of Ecuador, for example, Figure 5.1 shows how the oil sector is linked with other industries. Based on the analysis of the monetary flows between all the economic sectors in the country, one can observe how the wealth generated by the extraction of crude oil is directly linked with the state, which in turn invests this wealth mainly in construction activities and telecommunication. Furthermore, as this investment keeps trickling down into other important sectors, when the crude oil wealth reduces, it also trickles down the whole economy. Since international oil prices are volatile, the economy of a oil-exporting country (including Ecuador) that relies on its revenues becomes highly vulnerable to price changes. This volatility also increases financial uncertainty in these countries, which might discourage further investments (Djimeu and Omgba, 2019; Budina et al., 2006; Hausmann and Rigobon, 2003).

Secondly, a more diverse production makes an economy less vulnerable to the physical exhaustion of natural resources, while minimising the risk associated of their assets, such as crude oil, to be "stranded" due to a global decreasing demand for fossil fuels (Gelb, 2011; Caldecott et al., 2014). Indeed, it has been estimated there would be a sustained decline in global production of conventional crude oil before 2030 (Miller and Sorrell, 2014). A study done by the IEA (2012) estimated that existing fields will reduce their production from 68.5 million barrels (mb) per day in 2011 to only 26 mb per day by 2035. In the case of Ecuador, a recent study estimates that the country will become a net oil importer between 2024 and 2035 (Espinoza et al., 2019). If this happens, its oil trade balance could be seriously affected, decreasing from a current positive value of around 2 billion USD to incur deficits of 0.6 - 16.7 billion USD in 2035 (Espinoza et al., 2019). Likewise, Other oil exporting economies will face similar problems if they do not develop more productive alternatives to compensate for these losses.

Thirdly, a vast amount of research has shown that moving towards manufacturing generates higher benefits for the society, compared with primary production and some services (Haraguchi et al., 2017). The manufacturing sector has the ability to achieve higher levels of capital accumulation, generate economies of scale and foster technological progress, to name but three benefits (Szirmai, 2013). One of the most interesting aspects of manufacturing is that, as Haraguchi et al. (2017, p. 295) state: "it tends to accelerate the rate

Figure 5.1: Crude oil wealth flow in Ecuador (2015)



of technological progress of the economy as a whole, partly due to the absorption of surplus labour and the creation and dissemination of innovation in certain industries through linkage effects”. In addition, the development of manufacturing could help create new jobs outside the resource sector, providing employment opportunities that do not rely mainly on natural resources’ rents (Gelb, 2011).

Despite the benefits of having a diversified economy and the urge (previously explained) that oil exporting economies face, achieving it is a complex task and few countries have managed to break free from their resource dependency. There are two key reasons that prevent oil-exporting-countries from diversification. First, as it was explained earlier in this thesis (cf. Chapter 4), the Dutch Disease can lead to currency appreciation, which increases the profitability of the oil industry but also makes other tradable sectors less competitive (Hausmann and Rigobon, 2003). This effect draws human capital and other resources away from any non-tradable activity, thereby harming export diversification (Hausmann et al., 2010) and reduces the incentives to create new economic activities. Second, Ross (2019) states that the oil industry generates limited knowledge spillovers (learning-by-doing) that prevent countries from acquiring capabilities that can be applied in other industries, which in turn limits their potential to develop products that can support higher wages (Gelb, 2011).

Researchers have stressed that there are many elements to take into account when driving diversification. First, oil exporting countries have different local contexts, including population, workforce and skills, location and connectivity with international markets, oil reserves and potential for other resource-based activities (Gelb, 2011). These conditions should shape diversification priorities and policies. Second, there are different diversification paths that a country could take, such as, diversifying within the resource sector (moving from crude oil to gas), moving downstream in the production chains to increase value-added activities, or moving towards labour intensive manufacturing, and many more (Gelb, 2011). Third, according to Alshamsi et al. (2018) and Boschma (2017), the diversification routes that countries or regions can follow are place- and path-dependent, which means that their diversification is usually constrained by the existing capacities and related economic activities available in a location.

5.2.2 Regional Productive Diversification and Industrial Relatedness

Existing literature on productive diversification also stresses how complex it is to achieve it. Garud and Karnøe (2003) and Garud et al. (2010) stated that the creation of new productive paths is an iterative construction process wherein networks of distributed actors

jointly develop knowledge, create new market segments and user profiles, adapt regulation, lobby for subsidies, and/or define new technical standards and thereby ultimately create the environment that helps a new industry to develop and prosper in a given region. Ron Boschma et al. (2017) add that creating productive novelty is likely to be a messy, step-wise and experimental process, requiring entrepreneurship not only to be proficient in the relevant knowledge fields, but also to be capable of embedding new ideas in a wider social environments.

Research has found that the creation of these productive paths (i.e., diversification) is conditioned by four systemic resource formation processes: knowledge creation, investment mobilisation, market formation and technology legitimisation, which can be mobilised from inside or anchored from outside the region (Binz et al., 2016). Further, productive and technological relatedness across regional industries, combinatorial knowledge dynamics, and branching processes are key explanatory factors for where and how new industries develop (Asheim et al., 2011; Balland et al., 2019; Neffke et al., 2011; Boschma and Frenken, 2011; Mealy and Teytelboym, 2020). In that sense, regions are assumed to branch into productive related fields in a path-dependent related diversification processes (Martin, 2012, 2010; Martin and Sunley, 2006; Neffke et al., 2011; Boschma and Frenken, 2011).

In 2007, Hidalgo et al. (2007) mapped the relations that exist between the products traded in the global economy. This research, termed "The Product Space", formalised the previous ideas and allows us to explore analytically how countries have changed their productive paths and with what implications for future economic growth (Hidalgo and Hausmann, 2009). This work confirmed that there is a strong tendency of the productive mix of a region to move from current products or industries towards related ones, rather than goods that are less related (Hausmann et al., 2007b; Boschma, 2012). In other words, it confirmed that there is a strong relation between the number of **related activities that are already present in a given location** (relatedness by co-location) and the likelihood of that location to develop new industries. In later work, Boschma (2012) explained that this co-location relatedness is relevant because the particular knowledge base of a region acts as an incentive for new industries to appear, because it not only offers local firms opportunities to further improve in familiar fields of knowledge, but also discourages knowledge creation that does not fit in the regional knowledge base.

This "Principle of Relatedness" has been explored and corroborated at multiple spatial scales, for different economic activities and diverse institutional regimes. It has been observed, for example, in analysis of entry and exit of products (Neffke et al., 2011; Zhu et al.,

2017; Gao et al., 2017), industries (Hidalgo et al., 2007; Alshamsi et al., 2018), occupations (Muneepeerakul et al., 2013), and technologies (Kogler et al., 2013; Boschma et al., 2014), at the national, regional, and metropolitan scales (Hidalgo et al., 2018). In 2018, a new research that analysed 93 countries' diversification paths confirmed previous findings and also showed that, depending on the level of development, different countries diversified into more or less related economic activities than the ones that already existed in a given location (Pinheiro et al., 2018; Alshamsi et al., 2018).

5.2.3 Literature Gaps

A thorough search for relevant literature did not yield any study that combines productive diversification of oil exporting economies and regional diversification. On one hand, as it was stated previously, existing literature on natural resource dependency has focused on the challenges that oil exporting economies face, having failed to shed light on how to overcome their dependency on few commodities and diversify their productive bases (Ross, 2019). On the other hand, the literature on regional diversification has not focused on countries that heavily rely on primary exports. In this chapter, I argue that understanding the domestic productivity (productive networks) and regional dynamics is a key component to move forward the diversification debate in oil exporting economies.

As already explained, the debates on regional diversification has focused mainly on the idea that the number of related products, industries or technology in a given location (known in the Economic Geography literature as "Related Density") is a strong predictor of the future activities that could emerge in a place. However, production or industries are not necessarily bounded to specific locations. Usually, industries and firms are part of larger production networks that are fragmented and spatially dispersed (Alfaro et al., 2019; Coe and Yeung, 2015). Therefore, relatedness purely based on co-location is limited³. Incorporating how networks of clients and suppliers influences the emergence of pioneering industries, products or technologies in a place is important for a better understanding how productive diversification happens.

Furthermore, as far as the author is aware, existing research on relatedness has not yet made a distinction between which firms are producers or which are consumers, nor explored their specific roles on diversification. This is relevant because both producers and clients generate different incentives in an economy and their location could explain better some diversification dynamics. This chapter explores each case individually to assess its effects

³Studies also suggest that non-local capabilities is a key component for production (Trippel et al., 2018), so that the knowledge required to produce a good or to generate a service is not necessarily concentrated in a location.

on the entry of pioneering firms in a location. Moreover, there are limited studies that had looked and compared the influence that relatedness has on different industries. For instance, the agricultural sector relies heavily on the resources available in a given location, while manufacturing might rely more on the proximity with physical infrastructure to produce and distribute, and the service industry might rely more on the location of the customers.

In order to fill in the previous gaps, one of the objectives of this chapter is to merge the two separate debates on economic diversification of oil exporting economies and regional diversification (in Economic Geography), which as far as the author understands, has not been done before. Secondly, by testing the previous ideas of relatedness by co-location, I will incorporate the idea of relatedness based on supply chains as well as networks of consumers. My main argument is that the concentration of suppliers and clients is key for diversification, exerting a differentiated influence for the emergence of pioneering industries. Knowing these differences provides better insights on how pioneering economic activities may emerge. Thus, I will desegregate the previous analysis to explore the role of suppliers and consumers for the emergence of pioneering firms in a place. An analysis is also developed to compare the previous measurements in two economic sectors separately: extractive industry (oil and mining extraction) and manufacturing.

5.3 Data and Methodology

5.3.1 Data

To address the aforementioned literature gaps and objectives, a large and unique data set, which was provided by the Internal Revenues Service (SRI) of Ecuador, was used. It comprises of fine-grained information about all the legal (firms) and natural persons registered in the country between 2007 and 2015: 328,640 firms and 3,942,095 natural persons. It includes information of every transaction between all the entities, aggregated per year, and contains information of individual entities on:

- Industrial classification (ISIC⁴ Code - 4 digits, Rev. 4)
- Taxpayer classification (public sector, private sector, IGO, NGO)
- Fiscal address at the municipal level
- Total tax base (US Dollars)
- Total sales and purchases

⁴ISIC means: International Standard Industrial Classification of All Economic Activities is a standard classification of economic activities arranged so that entities can be classified according to the activity they carry out

This dataset also includes transactions between registered entities and some foreign companies, which are not registered in the country. Since the focus of the analysis is on the domestic product chain, and because their information is incomplete, all these transactions were excluded.

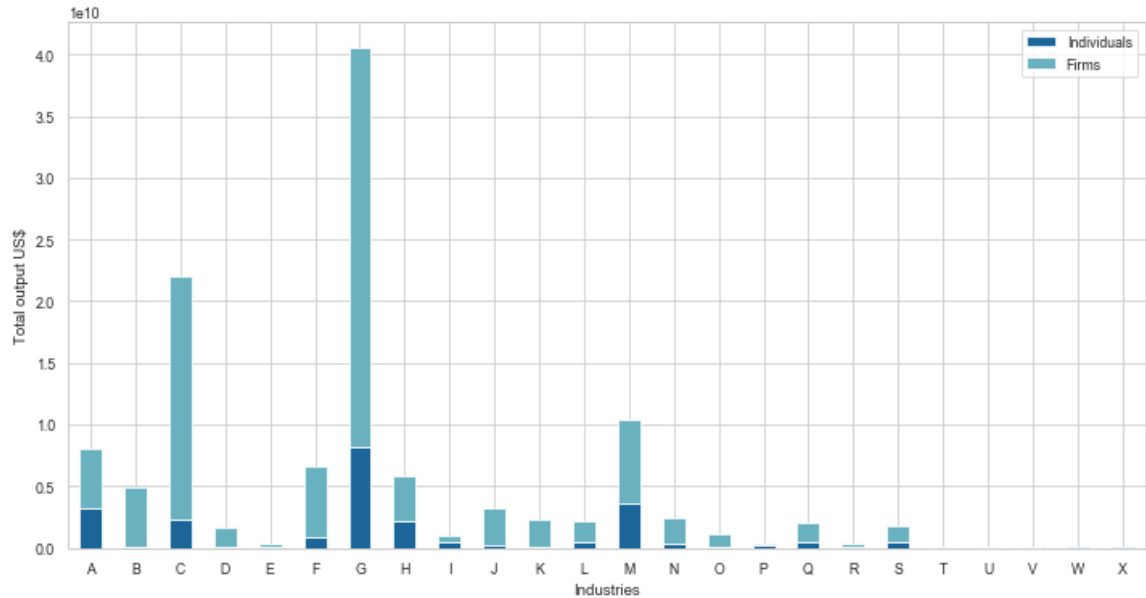
Before proceeding to the analysis, it is worth mentioning some of the limitations of this dataset. First, it only contains information of the firms as a whole, so there is no information about the establishments that firms might have in different locations. This is particularly problematic in the case of the wholesale and retail sectors, which are the main economic agents in the country.

Second, other firms may be registered in the main cities (such as Quito or Guayaquil), although their activities take place in a different location. This could generate bias in the analysis because some industries that operate outside the main urban areas will be included as part of the urban cohort. I correct for this in the analysis, explained in the methodology later. Through the analysis, it was found that the most problematic industries were those related with crude oil extraction and some tourist activities. In the case of crude oil, the national oil company and other oil foreign companies are registered in Quito, but their activities take place in two Amazonian provinces (Sucumbios and Orellana). Similarly, the main touristic activities in the country take place in the Galapagos Islands, but their headquarters are registered in Quito or Guayaquil.

Data Description

Figure 5.2 shows the two biggest economic sectors (measured by total output) in the Ecuadorian economy are *wholesale* and *retail trade* (ISIC: G), which generated more than 40,000 million (MM) US Dollars in 2015. These are followed by *manufacturing* (ISIC: C) that generated around 23,000 MM US Dollars, and *professional, scientific and technical activities* (ISIC: M) that generated 10,000 US Dollars. Moreover, as it could be expected, most of the output is generated by firms rather than individuals (natural persons). When the sectors are desegregated to a lower level (ISIC 2 digits), it can be observed that wholesale (ISIC: G46, 47) is the dominant sub-sector, followed by manufacturing of food products (C10), crop and animal production (A01), architectural and engineering activities (M71). All this data can be explored in detail in: www.pastudillo.com/data-1.

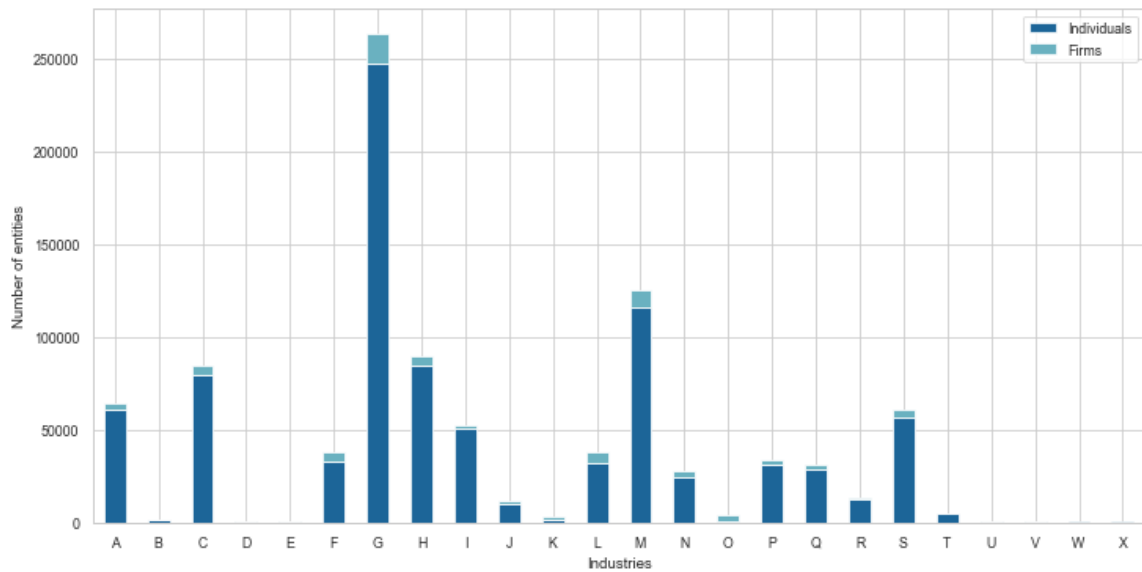
Figure 5.2: Total Output by Industry in Ecuador, by ISIC (2015)



A	Agriculture, forestry and fishing	I	Accommodation and food service activities	Q	Human health and social work activities
B	Mining and quarrying	J	Information and communication	R	Arts, entertainment and recreation
C	Manufacturing	K	Financial and insurance activities	S	Other service activities
D	Electricity, gas, steam and air conditioning supply	L	Real estate activities	T	Activities of households as employers
E	Water supply; sewerage, waste management	M	Professional, scientific and technical activities	U	Activities of extraterritorial organizations
F	Construction	N	Administrative and support service activities	V	No economic activity - ISIC
G	Wholesale and retail trade; repair of motor vehicles	O	Public administration and defence; social security	W	As a dependent private sector
H	Transportation and storage	P	Education	X	As a dependent public sector

In what concerns the sector that has the highest number of entities (firms and individuals), the wholesale and retail sector (G) stands out, as figure 5.3 shows. It encompasses more than 250,000 entities, mostly natural persons. This economic sector is followed by professional, scientific and technical activities (M) with around 125,000 entities, transportation and storage (H) with almost 90,000 entities and manufacturing (C) with similar values. When these values are desegregated to two digits, it can be observed that wholesale and transportation have the highest number of entities registered, followed by crops and animal production.

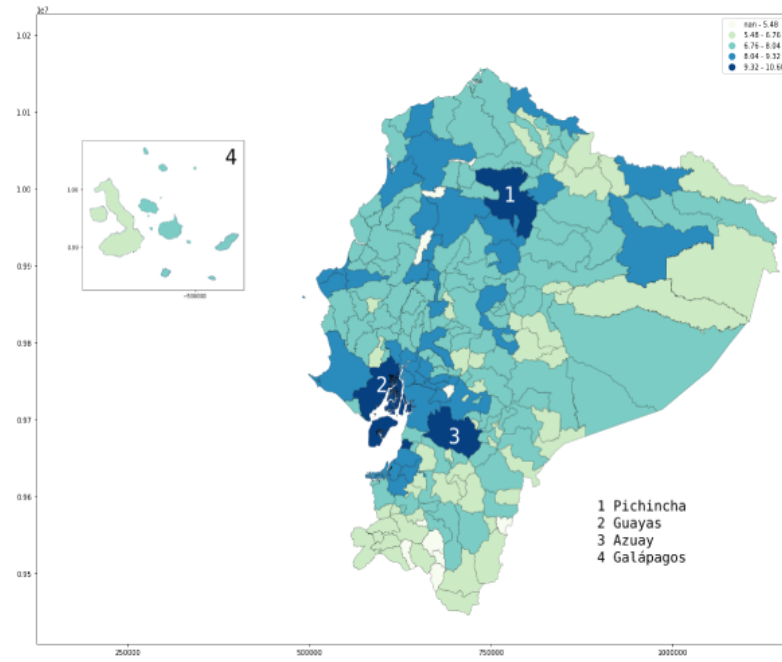
Figure 5.3: Total Number of Entities by Industry (2015)



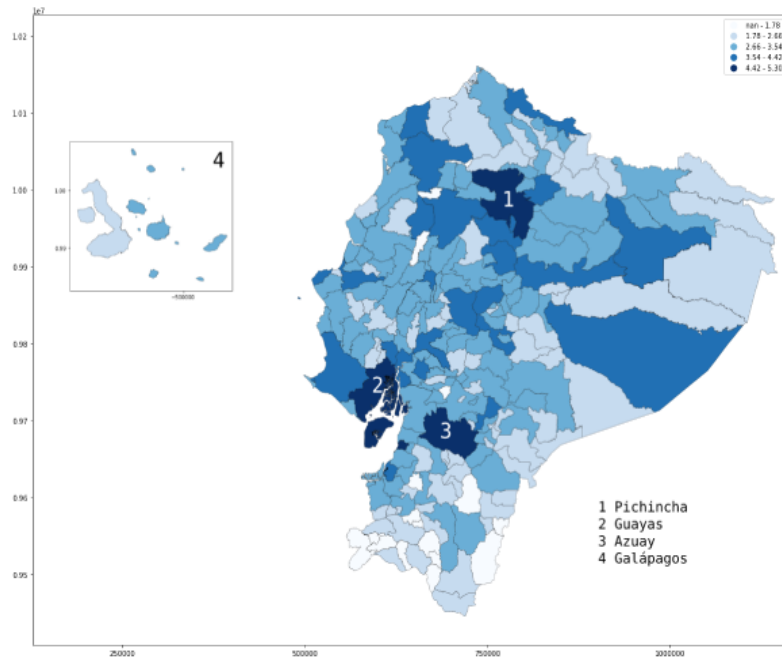
When comparing the two graphs, it can be observed that most of the sector's output is generated by firms (legal persons). Even when the number of individuals (natural persons) that perform economic activities is higher, their contribution to the total output of the sector is marginal.

Figure 5.4 shows how industries are distributed throughout the country. First of all, we can see that the sector's output is heavily concentrated in two provinces: Pichincha, where the capital (Quito) is located, and Guayas, where the second biggest city and port is located (Guayaquil). In contrast, natural persons are more evenly distributed, although their main concentration remains in three provinces: Pichincha, Guayas and Azuay.

Figure 5.4: Economic Activities by Province (2015)



(a) Total Output [Scale: Log10]

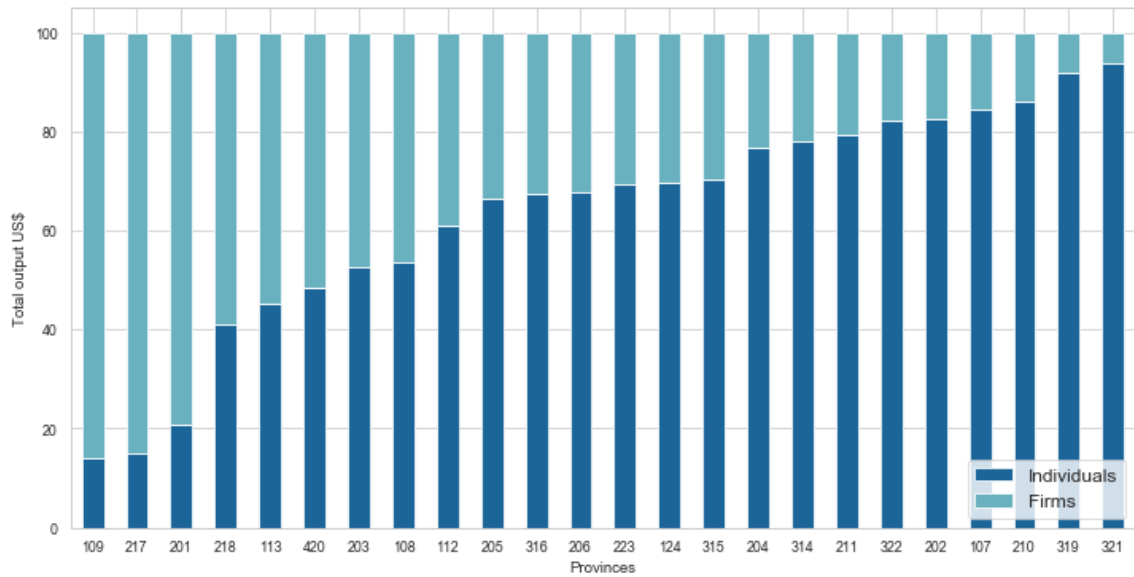


(b) Number of Entities [Scale: Log10]

An interesting fact that can be observed in figure 5.5 is that the share of the output generated by firms or persons varies according to the distance to the main urban centres. We can see that in more urban and populated provinces (Pichincha-217 and Guayas-109), most of the

output is generated by firms. This distribution changes (decreases) as we move into more inner provinces, such as Sucumbios-321 or Zamora Chinchipe-319 that are located in the Amazon region. This observation is coherent with the literature and can be observed in other countries (Balland et al., 2020b; Glaeser, 2011a,b).

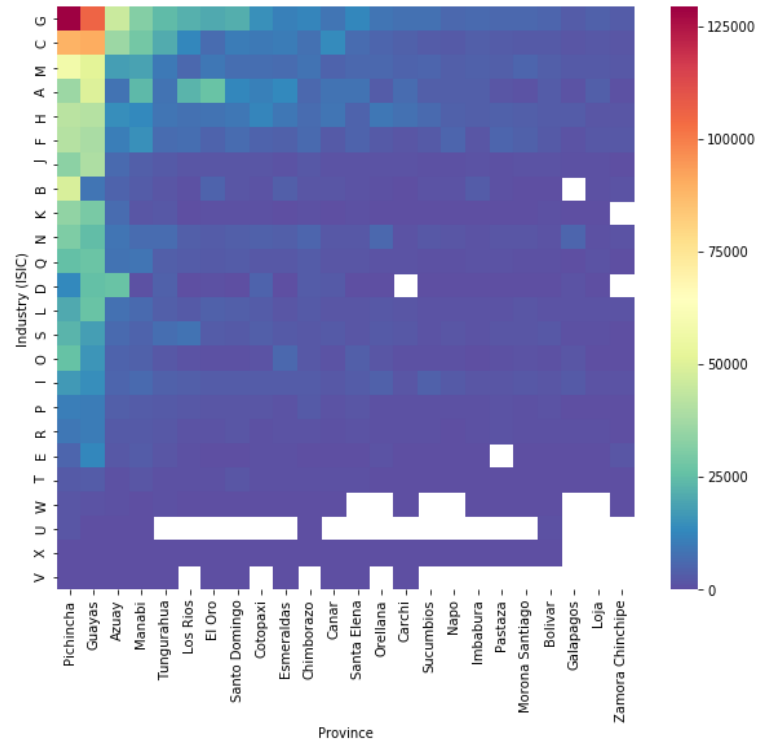
Figure 5.5: Output Shared Generated by Province (2015)



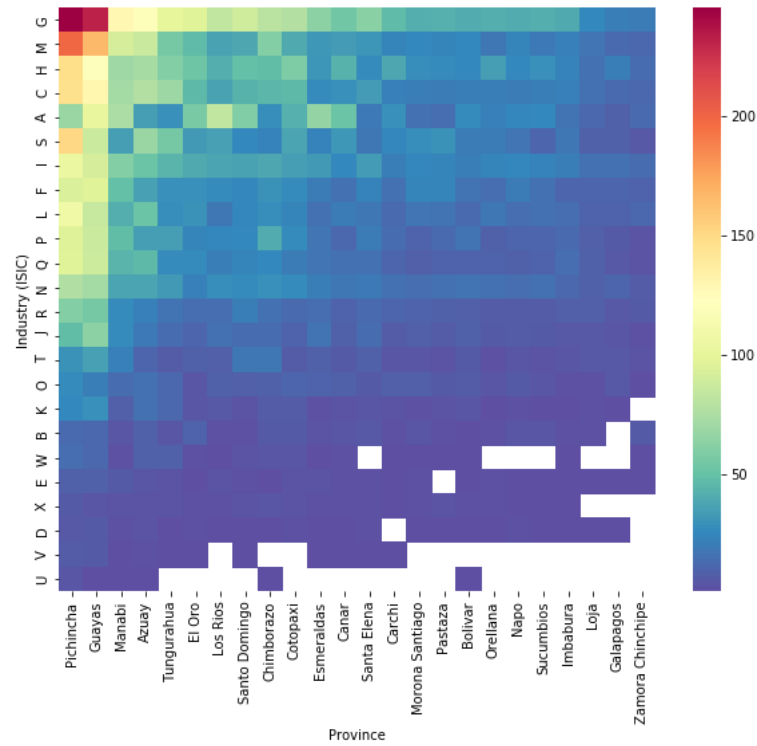
107	El Oro	201	Azuay	210	Imbabura	315	Napo
108	Esmeraldas	202	Bolivar	211	Loja	316	Pastaza
109	Guayas	203	Canar	217	Pichincha	319	Zamora Chinchipe
112	Los Rios	204	Carchi	218	Tungurahua	321	Sucumbios
113	Manabi	205	Cotopaxi	223	Santo Domingo De Los Tsachilas	322	Orellana
124	Santa Elena	206	Chimborazo	314	Morona Santiago	420	Galapagos

Figure 5.6 shows the distribution of total output and number of entities by industry in each province. In line with what has been described thus far, most of the output and entities are located in a couple of provinces. However, certain activities have a stronger presence in some provinces. For instance, agriculture (ISIC: A) has a stronger presence in Los Rios and El Oro than in Pichincha. Similarly, most of the industrial fish production (ISIC: A) is located in Manabi. Although the industrial output is more concentrated in certain regions, the number of entities is more widely distributed within the country. For this reason, the latter could be a more reliable indicator of the economic specialisation for regions than the industrial output. This will be a key consideration for the upcoming analysis.

Figure 5.6: Economic Activities by Industry per Province (2015)



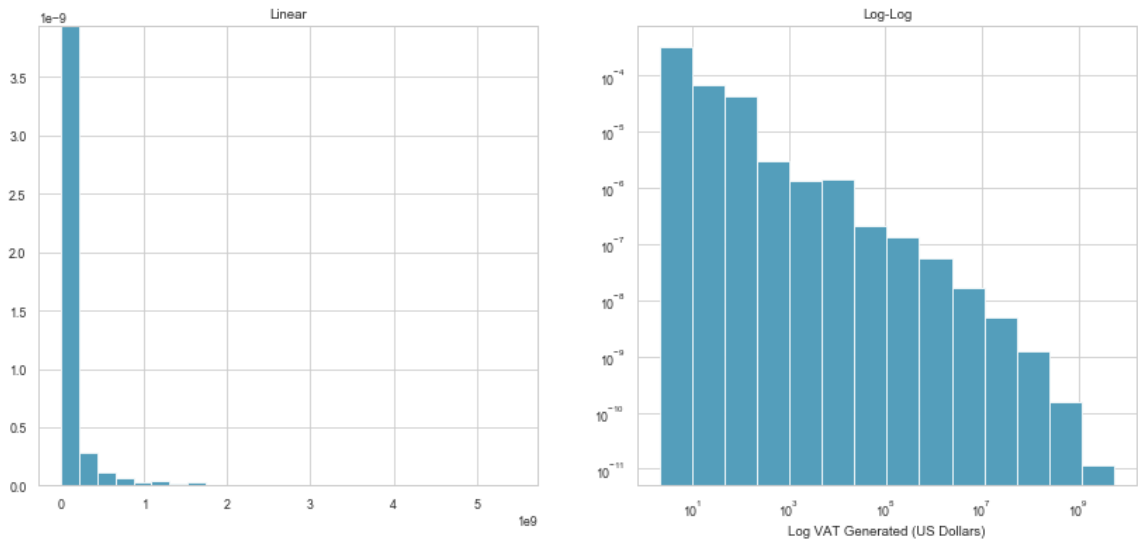
(a) Output [Scale: Square Root]



(b) Entities [Scale: Square Root]

Figure 5.7 illustrates the distribution of entities' output. In this figure we can appreciate that most of the output generated in the country is created by a small number of entities. In contrast, most of the existing entities produce limited output. This is a bad indicator for economic diversification, since only a limited number of firms (in a few sectors) have economic dominance. The observed distribution is not linear and resembles a pareto distribution⁵.

Figure 5.7: Entities' Output Distribution (2015)



5.3.2 Methodology

Industrial Relatedness in Cities

To start with the analysis, first one needs to assess how industries are related in a given city within the country. To accomplish this, I used an outcome-based measure of relatedness proposed by Hausmann et al. (2007b), Hausmann et al. (2007b), Hausmann et al. (2021). This method is agnostic about the specific inputs or complementarities that drive relatedness among activities (Hidalgo et al., 2018). For example, it has been applied to study the relatedness between products within countries (Hidalgo et al., 2007), products within regions (Boschma et al., 2013), technology classes within patents (Boschma, 2012), or co-occurrences of scientific knowledge (Boschma et al., 2014).

⁵The Pareto distribution is a skewed distribution with heavy, or 'slowly decaying' tails. This distribution is sometimes used to model that distribution of incomes (UCLA, 2007). In our case, this means that a high proportion of entities has low output while only few entities have very high output.

As a first step, I calculate the revealed comparative advantage (**RCA**) of every city within Ecuador for each year. This RCA measures whether the importance of an industry in a city is larger, as a share of all the industries, than the "average" city in a given country. In practical terms, this measurement helps to normalise the importance of an industry in a given city, based on the presence of other industries both in the city and throughout the country. Its effect is to reduce the noise of minor industries in a city.

The **RCA** of a city c over an industry i is computed as

$$\mathbf{RCA}_{c,i} = \frac{X_{c,i,t} / \sum_{c'} X_{c',i,t}}{\sum_{c'} X_{c',i,t} / \sum_{c',i'} X_{c',i',t}} \quad (5.1)$$

where $X_{c,i,t}$ is the industry size i (measured by its total output or by the number of entities⁶ that constitute it) by city c during year t . When the **RCA** is larger than 1, it indicates that a city has *comparative advantage* over an industry. Drawing on these results, I constructed a matrix that connects each city to their main industries that are present. Entries in the matrix are 1, if the city c has an industry i with RCA greater than 1, and 0 if otherwise. Formally, this was defined as the $M_{c,i}$ matrix, where:

$$\mathbf{M}_{c,i} = \begin{cases} 1 & \mathbf{RCA}_{c,i} \geq 1 \\ 0 & \mathbf{RCA}_{c,i} < 1 \end{cases} \quad (5.2)$$

Using the RCA results, the proximity (similarity) between the existing industries was also estimated. Formally, the proximity ϕ between industries i and i' is the minimum of the pairwise conditional probability⁷ of the presence of an industry in a city given that another industry is present in the same city Hidalgo et al. (2007).

$$\phi_{i,i'} = \min \{P(\mathbf{RCA}_{\mathbf{x}_i} | \mathbf{RCA}_{\mathbf{x}_{i'}}), P(\mathbf{RCA}_{\mathbf{x}_{i'}} | \mathbf{RCA}_{\mathbf{x}_i})\} \quad (5.3)$$

As a result, the measure is symmetric and $\phi_{i,i'} \in [0, 1]$. The values oscillate between 0 and 1, so a value of 0 indicates that the two industries never co-occurred within the same city,

⁶According to Hausmann et al. (2021), using the population of a location as a measure of industry size rather than its total production eliminates the impact of price and output fluctuations of one industries on the values of other industries.

⁷As suggested by Hausmann et al. (2021), we obtained comparable results using other similarity measures: Pearson correlation and cosine similarity.

while a value of 1 indicates that the two industries systematically co-occur. In other words, this means that if two industries have a proximity of 0.4 then there is a 40% chance that a random city with **RCA** over one industry has **RCA** over both industries. Table 5.1 shows the the proximity between industries based on their co-location. It can be observed that most of the extractive and manufacturing industries are associated with retail and food services.

ind.i	ind.j	similarity
Retail sale in non-specialized stores with food...	Real estate activities on a fee or contract basis	0.986175115
Restaurants and mobile food service activities	Retail sale in non-specialized stores with food...	0.986175115
Real estate activities on a fee or contract basis	Freight transport by road	0.976635514
Retail sale in non-specialized stores with food...	Wholesale of textiles, clothing and footwear	0.97235023
Real estate activities on a fee or contract basis	Restaurants and mobile food service activities	0.97235023
Freight transport by road	Retail sale in non-specialized stores with food...	0.97235023
Real estate activities on a fee or contract basis	Wholesale of textiles, clothing and footwear	0.971962617
Retail sale of hardware, paints and glass in specialized stores	Freight transport by road	0.971698113
Retail sale of hardware, paints and glass in specialized stores	Maintenance and repair of motor vehicles	0.971563981
Restaurants and mobile food service activities	Retail sale of hardware, paints and glass in specialized stores	0.967741935
Construction of buildings	Freight transport by road	0.966981132
Freight transport by road	Wholesale of textiles, clothing and footwear	0.966981132
Maintenance and repair of motor vehicles	Wholesale of textiles, clothing and footwear	0.966824645
Renting and leasing of other machinery, equipment and tangible goods	Retail sale of food in specialized stores	0.965853659
Retail sale in non-specialized stores with food...	Retail sale of hardware, paints and glass in specialized stores	0.963133641

Table 5.1: Top Industrial Proximity (Co-location) in 2015

Next, I combined proximity $\phi_{i,i'}$ between industries with the economic specialisation of cities (the economic activities that exist in a city) to construct a city-industry level variable that indicates how close a potential new industry is to the existing economic structure of a given city (Boschma et al., 2014). This variable is known as the Relatedness Density $\mathbf{RD}_i^c$ and combines the proximity between industries with the particular economic specialisation of cities. It indicates how cognitively close (how similar is the productive knowledge) a potential new industry is to the preexisting set of industries a given city is specialised in.

The Relatedness Density $\mathbf{RD}_i^c$ of an industry i to the expertise of city c in time t was computed by summing the industrial proximity i to all the industries that can be found in a city c , divided by the sum of the relatedness of an industry to all the other industries in t , and is given by the following formula:

$$\mathbf{RD}_i^c = \frac{\sum_{i'} M_{c,i'} \phi_{i,i'}}{\sum_{i'} \phi_{i,i'}} \quad (5.4)$$

The city-industry relatedness $\mathbf{RD}_i^c$ can be interpreted as the percentage of related industries found in a city. By construction, the relatedness variable lies then between 0 and 1.

City Relatedness Based on Industries

Similarly to the previous subsection, the city relatedness $\mathbf{CR}_c^i$ for every year is computed using the following formula:

city-i	city-j	similarity
Tena	Pastaza	0.594827586
Echeandia	Chillanes	0.587301587
San Miguel De Los Bancos	Pedro Vicente Maldonado	0.569444444
Guachapala	El Pan	0.566666667
San Juan Bosco	Logrono	0.555555556
La Troncal	El Triunfo	0.555555556
Pastaza	Morona	0.550847458
Santa Elena	La Libertad	0.545454545
Shushufindi	La Joya De Los Sachas	0.544117647
Logrono	Huamboya	0.542857143
Junin	Flavio Alfaro	0.542372881
Nabon	Limon - Indanza	0.54
Limon - Indanza	Carlos Julio Arosemena Tola	0.54
Urdaneta	Baba	0.537037037
Pindal	Pablo Sexto	0.535714286

Table 5.2: Top Cities Proximity (based on industrial composition) in 2015

$$\mathbf{CR}_c^i = \frac{\sum_{c'} M_{c',i} \phi_{c,c'}}{\sum_{c'} \phi_{c,c'}} \quad (5.5)$$

This $\mathbf{CR}_c^i$ can interpreted as how cities are related, based on their industrial profile. By construction, the relatedness variable lies between 0 and 1. This measurement is crucial to provide a more accurate measurement of the capabilities available in a city, because it helps us to correct the information that $\mathbf{RD}_i^c$ provides and vice-versa.

For cities, and in line with what Hausmann et al. (2014) explain, this requires the calculation of the average ubiquity of the economic activities that are taking place in that location and the average diversity of economic activities that those cities have. For industries, this requires calculating the average diversity of the economic activities that take place in the city and the average ubiquity of other industries that those cities have. This will be captured by the econometric model that it will be explained below.

Industrial Relatedness Based on Clients and Suppliers in Cities

To assess the industrial relatedness based on clients (RCL) and suppliers (RSP), I followed a similar logic to the one explained above (industrial relatedness in cities) with some modifications. In the original RCA, all the economic activities that are present in a location are taken into account. However, to evaluate the RCA for clients or for suppliers, I selected the entities that are acting like clients or suppliers for each city, respectively. Based on this, the values were computed as follows:

$$\mathbf{RCL}_{c,i} = \frac{CL_{c,i,t}}{\sum_{c',i'} CL_{c',i',t}} \quad (5.6)$$

$$\mathbf{RSP}_{c,i} = \frac{SP_{c,i,t}}{\sum_{c',i'} SP_{c',i',t}} \quad (5.7)$$

Using the previous values (RCL and RSP), the proximities calculating the respective conditional probabilities, were estimated:

$$\phi_{CL_{i,i'}} = \min \{P(\mathbf{RCLx}_i | \mathbf{RCLx}_{i'}), P(\mathbf{RCLx}_{i'} | \mathbf{RCLx}_i)\} \quad (5.8)$$

Table 5.3 shows the similarity between industries taking into account the clients in 2015, using the previous formula. As it can be observed, the industries that co-locate vary in comparison to the previous metrics.

ind.i	ind.j	similarity
Retail sale in non-specialised stores with food...	General public administration activities	0.870056497
Regulation of the activities of providing health care, education...	Activities of trade unions	0.858208955
Manufacture of corrugated paper and paperboard	Manufacture of soap and detergents, cleaning and polishing preparations...	0.857142857
Retail sale in non-specialised stores with food...	Architectural and engineering activities	0.853107345
Public order and safety activities	Activities of trade unions	0.843283582
Architectural and engineering activities	General public administration activities	0.84180791
Retail sale of hardware, paints and glass in specialised stores	Retail sale in non-specialised stores with food...	0.83908046
Fund management activities	Manufacture of wines	0.833333333
Service activities incidental to water transportation	Activities of collection agencies and credit bureaus	0.833333333
Data processing, hosting and related activities	Renting and leasing of recreational and sports goods	0.833333333
Residential nursing care facilities	Activities of collection agencies and credit bureaus	0.833333333
Life insurance	Retail sale of games and toys in specialised stores	0.833333333
Service activities incidental to water transportation	Residential nursing care facilities	0.833333333
Manufacture of other rubber products	0.833333333	
Wholesale of food, beverages and tobacco	Retail sale of pharmaceutical and medical goods, cosmetic...	0.828947368

Table 5.3: Top Industrial Proximity (Clients) in 2015

$$\phi_{SP_{i,i'}} = \min \{P(\mathbf{RSPx}_i | \mathbf{RSPx}_{i'}), P(\mathbf{RSPx}_{i'} | \mathbf{RSPx}_i)\} \quad (5.9)$$

In table 5.4, it can be observed the similarity between industries taking into account the suppliers in 2015, using the previous formula. Compared to the previous metrics, the industries that co-locate vary, but the similarity between industries is more accurate. We confirmed this by paying attention to the detailed description of the industries.

Finally, the industrial relatedness, based on clients and suppliers in a city, was assessed following a similar procedure as the one explained previously, using the equations:

ind_i	ind_j	similarity
Retail sale in non-specialised stores with food...	Restaurants and mobile food service activities	0.986175115
Retail sale in non-specialised stores with food...	Real estate activities on a fee or contract basis	0.97235023
Maintenance and repair of motor vehicles	Retail sale of hardware, paints and glass in specialised stores	0.966346154
Maintenance and repair of motor vehicles	Retail sale of food	0.966183575
Retail sale in non-specialised stores with food...	Construction of buildings	0.963133641
Retail sale in non-specialised stores with food...	Freight transport by road	0.963133641
Real estate activities on a fee or contract basis	Freight transport by road	0.962085308
Retail sale of hardware, paints and glass	Freight transport by road	0.961904762
Other retail sale of new goods	Architectural and engineering activities	0.96097561
Renting and leasing of other machinery, equipment	Retail sale of food in specialised stores	0.96097561
Other retail sale of new goods	Retail sale of food	0.96097561
Restaurants and mobile food service activities	Construction of buildings	0.958525346
Real estate activities on a fee or contract basis	Restaurants and mobile food service activities	0.958525346
Construction of buildings	Real estate activities on a fee or contract basis	0.95754717
Freight transport by road	Construction of buildings	0.95754717

Table 5.4: Top Industrial Proximity (Suppliers) in 2015

$$DCI_i^c = \frac{\sum M_{Cl_{c,i}} \phi_{Cl_{i,i'}}}{\sum \phi_{Cl_{i,i'}}$$

$$DSp_i^c = \frac{\sum M_{Sp_{c,i}} \phi_{Sp_{i,i'}}}{\sum \phi_{Sp_{i,i'}}$$

Industry appearances

In the previous sections, I calculated different combinations of proximities and densities, which are key variables for the coming analysis. In this subsection, I explain how to identify the appearances of industries in locations following the methodology used by (Hausmann et al., 2021).

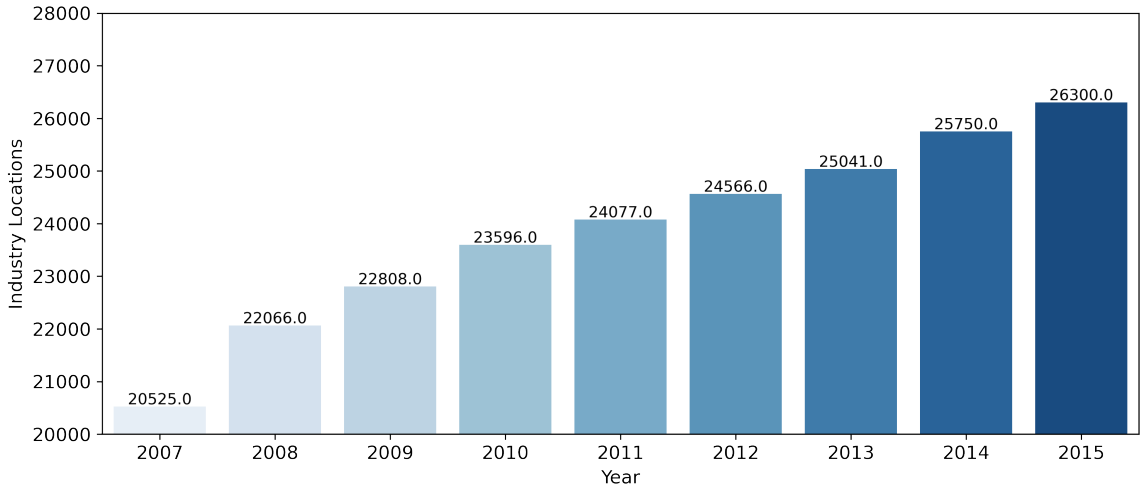
The first step was to identify the industry-locations that are present. Because the dataset used in this chapter reports the number of firms and individuals that are present in a location, this identification was straightforward: an industry is present in a location if at least one firm or individual is reported to exist there. Formally, this signal is captured by the binary presence variable M_{il,t_0} :

$$\mathbf{M}_{il,t_0} = \begin{cases} 1 & y_{il,t_0} \geq 1 \\ 0 & y_{il,t_0} = 0 \end{cases} \quad (5.10)$$

where, as before, y_{il,t_0} is the number of firms in industry i and location l in year t_0 . In this notation, I refer to an industry location as *present* when $M_{il,t_0} = 1$ and *absent* when

$M_{il,t_0} = 0$. As the result we have a matrix per each year, where the rows are every industry that exist in the country, the columns are all the Ecuadorian municipalities, and the values are: 1 if the industry exists in a location, 0 if otherwise. In figure 5.8 it can be observed the increase of 5775 industries/location between 2007 to 2015 in the country. This is the total number of appearances of industries in all the country's municipalities.

Figure 5.8: Number of Industry-Location Per Year



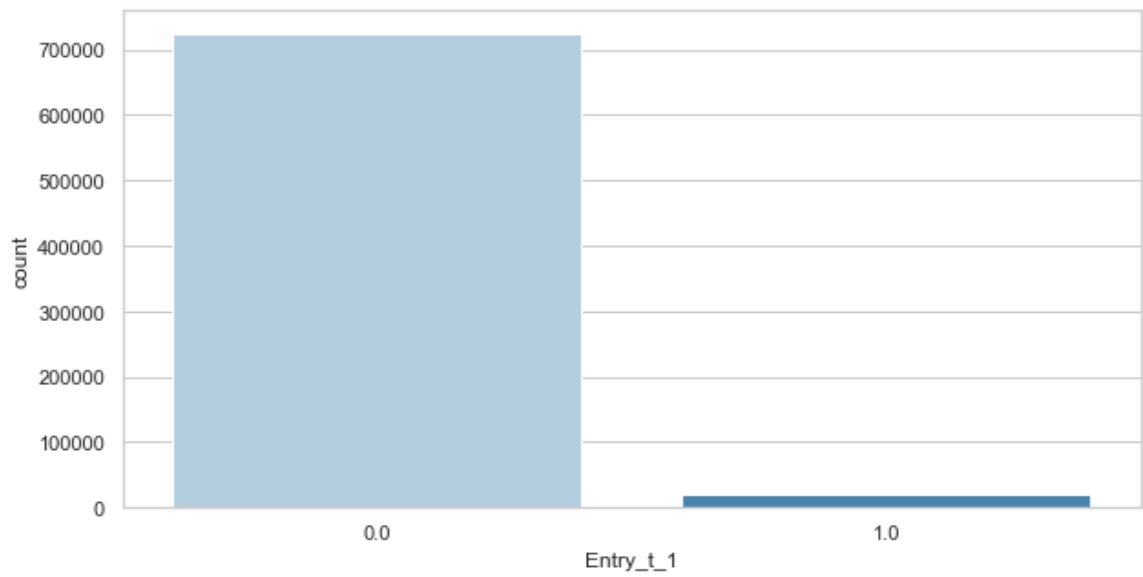
Using the previous matrix, I created a set of different "entry" matrices to identify the appearance of industries. Following a similar logic, an appearance between years t_0 and t_1 is defined as $M_{il,t_0} = 0 \rightarrow M_{il,t_1} = 1$. To do this, I used 2007 as a base year ($t_0 = 2007$) and the remaining years (t_1 to t_n) are used for the analysis. As a result, we have several matrices with the intervals of one, three and five years. For instance, for the matrix $M_{il,t_{2008-2009}}$, the rows are the industries, the columns are the locations, and the values are $M_{il,t_{2008-2009}} = 0$ if there is no appearance of a new industry in a location, and $M_{il,t_{2008-2009}} = 1$ if there is a pioneering (not existing) industry in a location.

5.3.3 Empirical Model

Data pre-processing

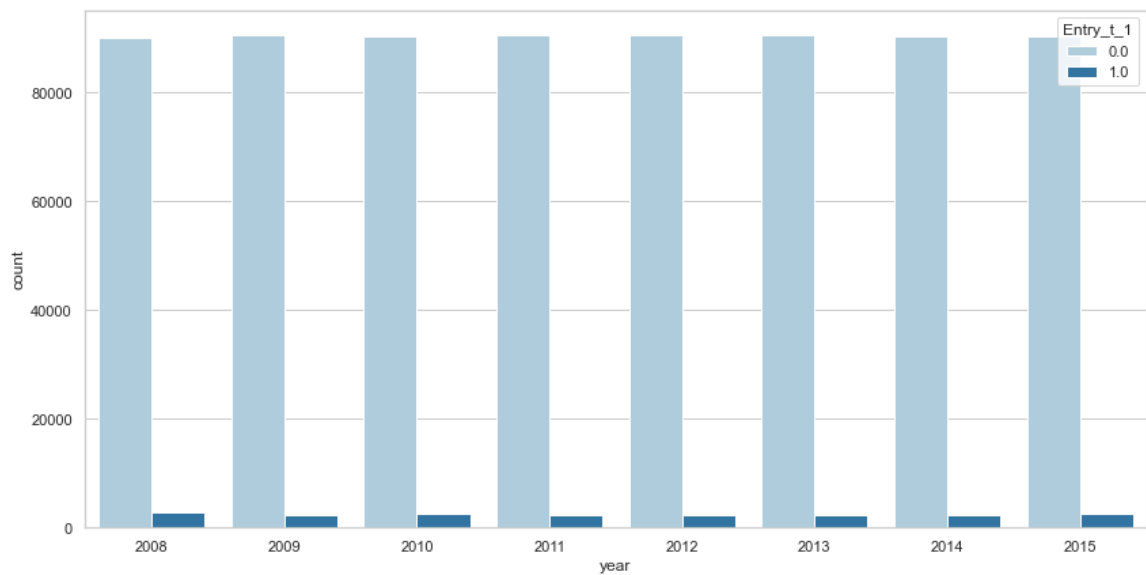
Using the previous measurements, I assessed the likelihood of industries that did not exist in a given place and time to emerge in a given city. However, before proceeding to the model details, there was the need to first assess the ratio of subscriptions (new industries) / non-subscriptions (existing industries) in the data set.

Figure 5.9: Total class subscription



Total number of existing entities (light blue) vs Total number of pioneer entities (dark blue)

Figure 5.10: Class subscription by year



Total number of existing entities per year (light blue) vs Total number of pioneer entities per year (dark blue)

As shown in figures 5.9 and 5.10, the classes are imbalanced. The ratio of non-subscription to subscription instances is 97:3, which means that most of the firms already existed and just a few are pioneers ⁸. To overcome this issue, I used a Synthetic Minority Oversampling Technique (SMOTE) to increase our sample size ⁹. This technique works by creating synthetic samples from the minor class (non-subscription) and randomly chooses one of the k-nearest-neighbours and create a similar, but randomly tweaked, new observation (Chawla et al., 2002). In our particular case, it was specified that cities, years and industrial classification are categorical data. In the following tables, we can observe a description of the dataset previous and after the application of this SMOTE technique. More details are discussed in the next sub-section.

Data, previous SMOTE

Number of no subscription, previous SMOTE	723150
Number of subscription, previous SMOTE	19410
Proportion of no subscription, previous SMOTE	0.97
Proportion of subscription, previous SMOTE	0.26

Data, after SMOTE

Length of over-sampled data	1007630
Number of no subscription in over-sampled data	503815
Number of subscription	503815
Proportion of no subscription data in over-sampled data	0.50
Proportion of subscription data in over-sampled data	0.50

Model Implementation

Following the methodology proposed by Boschma et al. (2014), a logistic model was used to regress the emergence of an industry ($Entry_{c,i,t}$)¹⁰ on their degree of relatedness with the industrial portfolio of cities, which is captured by the relatedness densities as previously calculated. It is worth recalling that there is a debate in the literature on whether one should use non-linear specifications (Logistic, Probit) or linear probability (OLS) models for this kind of analysis (Osborne, 2014). In this case, logarithmic models were selected.

⁸There is an imbalance between the classes in the dataset due to a large difference between the number of instances belonging to each class (Chawla et al., 2002)

⁹This is important because an imbalanced dataset, often causes the model to develop a certain bias towards the majority class (Chawla et al., 2002)

¹⁰This variable was computed as explained in the previous section (entry matrix)

Nevertheless, to check the robustness of the findings, logistic and OLS were also tested, which can be found on the *Appendix B*. The model is written as follows:

$$Entry_{c,i,t} = \beta_1 Density_{i,c,t_0} + \beta_2 City_{c,i,t_0} + \beta_3 Controls_{i,c} + \phi_c + \psi_i + \varepsilon_{i,c} \quad (5.11)$$

where the dependent variable $Entry_{c,i,t}$ takes the value of 1 when an industry i in a city c did not existed in the base year $t_0 = 2007$, but appears in years $t1 = 2008, t3 = 2010, t5 = 2012$. As independent variables, the individual densities previously calculated, $Density_{i,c}$ in the base year t_0 were used. The city relatedness $City_{c,i}$ in year t_0 , as well as *population density* and *productivity per capita* in the same year were used as controls. The same logic was used to assess the industrial disappearance likelihood $Exit_{c,i,t}$.

The independent variables, which were formalised in section 5.3.2, to be tested in the model are:

- a. Relatedness density (RD) - Equation (5.4)
- b. Relatedness suppliers (RSP) - Equation (5.7)
- c. Relatedness clients (RCL) - Equation (5.6)

In addition, ϕ_c is an city-fixed effect ψ_i is an industry-fixed effect, and $\varepsilon_{i,c}$ is the regression residual. Therefore, the baseline model was used in a two-way fixed effect model, to take into account omitted variable bias at the city and industry levels. The ϕ_c and ψ_i fixed effects that are directly estimated by including dummy variables for each city, industry for the period of time in concern.

Finally, to extend the econometric model and assess the potential that the density variables have on predicting the emergence of pioneering industries in regions, I propose a machine learning (ML) - logistic model. This approach, which is both statistical as computational, has been recognised by many scholars as a new toolbox to address the old challenges of prediction in econometric research (Mullainathan and Spiess, 2017; Athey, 2019; Judge, 2019). It has also been argued that it is bringing substantial changes to how empirical work is conducted, for instance, when dealing with large and complex datasets (as the one used in this research), and many ML methods simplify a variety of arbitrary choices analysts needed to make (Athey, 2019). Its appeal is based on its ability to uncover generalizable patterns by discovering complex structures that were not specified in advance (Mullainathan and Spiess, 2017; Athey, 2019; Athey and Imbens, 2019; Judge, 2019).

In this chapter, I used a supervised machine learning model, which seeks functions that predict well out of a given sample. More specifically, I look to predict the emergence of a pioneering industry y in a city from its observed characteristics x (which are the densities also used in the previous model) based on a sample of n industries (y_i, x_i) . In this case, the algorithm takes a loss function $L(\hat{y}, y)$ as an input and search for a function $\hat{f}$ that has low expected prediction loss $E_{(y,x)}[L(\hat{f}(x), y)]$ on a new data point from the same distribution.

To run this model, the dataset was randomly split into two parts. One part of it was used to train the model, whereas the another was used to test the prediction of the model. At this point, it is important to note that the oversampling previously described (SMOTE technique) only affected the training data (used for creating the synthetic observations), so the testing dataset is not affected (none of this information was used to create the synthetic observations) and it will be used to validate the accuracy of the model. The results will be presented in the next section.

5.4 Findings: Relatedness and Pioneering Industries' Emergence

The results of the regression analyses explained above are shown in table 5.5 for the emergence of new industries. At first glance, in Model 1, there is a significant relation between Relatedness Density (co-location relatedness) and the entry of pioneering industries in a place for 2007-2008 (t_1), 2007-2010 (t_3) and 2007- 2012 (t_5). The coefficient can be interpreted as the higher the RD , the more likely pioneering industries are to appear in a city. Therefore, our RD is positively associated with the emergence of new economic activities in these cities. This finding is coherent with the existing literature on the topic, which explains that the RD does signal the emergence of pioneering industries in a given place. Furthermore, when we look at the different periods of analysis, we can observe decreases in the RD coefficients (from 3.447 to 3.216), as well as in the explanatory power of the model ($PseudoR^2 = 0.296 - 0.220$). This means that RD does signal the emergence of pioneering industries, but its effect reduces over time.

As it was stated in the literature review, firms and industries do not operate in isolation. Rather, they are usually part of broader production networks. In order to understand which networks are the most relevant to explain the emergence of a pioneering industry, I expanded the previous co-location approach and analysed separately the networks of suppliers and clients for every industry in a given location. In this sense, two more variables were included and tested separately: the Suppliers Density DSP and the Customer Density DCL .

VARIABLES	t1 (2007-2008)			t3 (2007-2010)			t5 (2007-2012)		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Relatedness Density (RD)	3.447*** (0.0707) [31.41]			3.313*** (0.0670) [27.46]			3.216*** (0.0669) [24.94]		
Relatedness Suppliers (RSP)		3.157*** (0.0723) [23.5]			2.933*** (0.0684) [18.79]			2.847*** (0.0683) [17.23]	
Relatedness Clients (RCL)			-0.0557* (0.0283) [0.946]			-0.152*** (0.0276) [0.859]			-0.193*** (0.0276) [0.824]
City Relatedness (CR)	0.985*** (0.0360) [2.678]	1.108*** (0.0368) [3.029]	2.433*** (0.0262) [11.39]	0.883*** (0.0345) [2.418]	1.052*** (0.0352) [2.863]	2.357*** (0.0250) [10.55]	0.868*** (0.0345) [2.382]	1.033*** (0.0352) [2.810]	2.319*** (0.0249) [10.17]
Population Density	-0.00608** (0.00208) [0.994]	-0.00502* (0.00203) [0.995]	0.000246 (0.00177) [1]	-0.000978 (0.00425) [0.999]	-0.000408 (0.00417) [1.00]	-0.00135 (0.00409) [0.999]	0.0219 (0.0405) [1.022]	0.0241 (0.0396) [1.024]	0.207 (0.127) [1.228]
Production per Capita	-1.992** (0.677) [0.136]	-1.868** (0.658) [0.155]	-1.645* (0.683) [0.193]	-1.890 (1.755) [0.151]	-1.863 (1.818) [0.115]	-1.353 (1.184) [0.259]	-5.600 (9.681) [0.0037]	-6.423 (9.725) [0.00162]	20.59 (19.97) [0.0087]
City F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R ²	0.296	0.293	0.257	0.234	0.231	0.192	0.220	0.216	0.179
N	122919	122919	116317	121751	121751	114682	119713	119713	113185

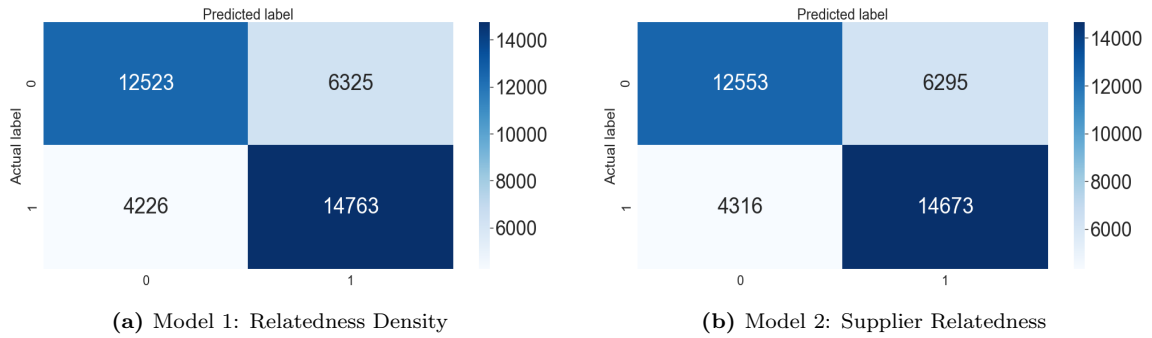
Standard errors in parentheses, odd ratios in square brackets
* p<0.05 ** p<0.01 *** p<0.001

Table 5.5: Logistic Regressions Results (Entry)

The results of the model that uses *DSP* to explain the emergence of pioneering industries are lower than the *RD* - their *pseudoR*² decreased from 0.296 to 0.293, and both variables are statistically significant. Moreover, the results show that the suppliers density has a lower coefficient, so it could have a slightly smaller influence on future paths of diversification than the relatedness density: the *DSP* has a coefficient of approximately 3.157 whereas the *RD* has a coefficient of 3.447. A similar variance occurs in the case of their odds ratios. When we consider the time periods *t1*, *t3* and *t5*, the coefficients and *pseudoR*² decrease similarly to the previous model, meaning that the density reduces its relevance with time.

The model also shows that the density based on consumers *DCL* is not as relevant as the density based on suppliers *DSP*, so its explanatory power to signal the entry of pioneering industries is lower than the density of suppliers. Both variables are statistically significant (but *DCL* is less significant than *DSP*), with a *PseudoR*² = 0.257 and 0.293, and coefficients of -0.005 and 3.157, respectively. Furthermore, as in the previous model, the coefficients and *pseudoR*² also decreases as time advances. In this sense, it can be argued that the location of the suppliers is a better indicator about the emergence of pioneering industries in a given location than the location of the clients.

Figure 5.11: Confusion Matrix



As explained at the beginning of this chapter, a machine learning (ML) model was created to complement the econometric analysis, in order to assess whether the variables could predict future paths of diversification. The results of the ML - Logistic Model are shown in figures 5.11 and 5.12. Herein, it is possible to observe the predictability differences between the relatedness based only on co-location and the relatedness based on the supply chain.

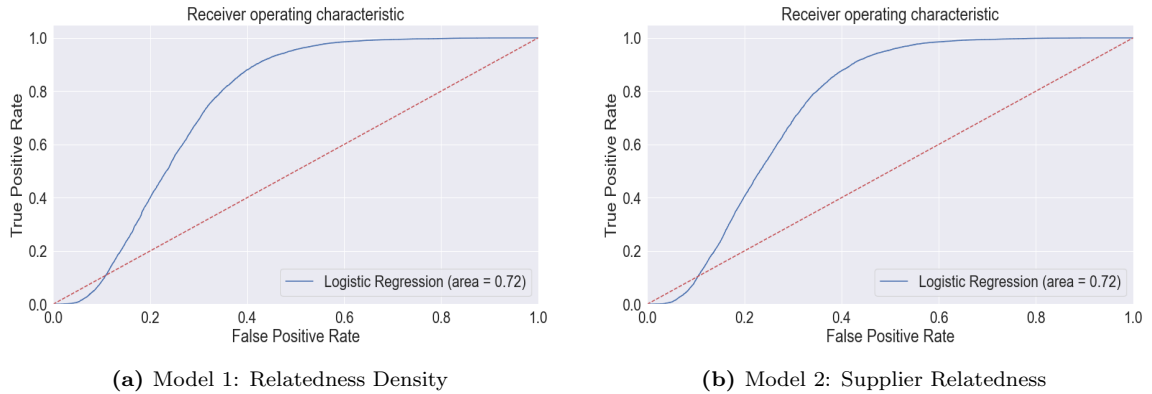
In the figures we can observe the predictability differences between the relatedness based only on co-location and the relatedness based on the suppliers' network. As it can be seen, figure 5.11 shows the confusion matrices of the two models ¹¹. While the first matrix is telling us that our ML model has $12523 + 14763$ correct predictions and $6325 + 4226$ incorrect predictions, the second matrix is telling us that it has $12553 + 14673$ correct predictions and $6295 + 4316$ incorrect ones. This clearly signals that the relatedness based on suppliers have more predictive power for the emergence of pioneering firms than the relatedness density.

In figure 5.12, we can observe the receiver operating characteristic curve¹² (ROC), which is a common tool used with binary classifiers. The dotted line represents the ROC curve of a purely random classifier; a good classifier stays as far away from that line as possible (i.e., towards the top-left corner). At the bottom-right of the figures there is the precision of our models. The left figure illustrates the relatedness based on co-location, which has a precision of 0.72, meaning that this variable could predict 72% the cases of pioneering industries emerging in cities. The figure on the right shows that the relatedness based on the suppliers has a similar precision (0.72), meaning that this variable could predict by 72% of the cases of pioneering industries emerging in cities.

¹¹These matrices describe the performance of the classification model used. More particularly, they display the relations between the true negatives (up-left), true positives (down-right), false positives (up-right) and false negatives (low-left) of the model.

¹²ROC curve, is a graphical plot [one of the main results of a ML classification method] that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied.

Figure 5.12: Area Under the Curve



5.4.1 Crude Oil and Mining Industry Emergence vs Manufacturing

Previously, we observed that the relatedness density is an important indicator for the emergence of new industries in a location. The density of suppliers does a better job signalling the appearance of pioneering industries in a location than clients' density. However, we know that every industry has its own special features. For instance, agriculture differs from manufacturing in terms of its ubiquity, location, type of providers and clients, and production processes. So, the oil and mining industry.

In this subsection, I argue that relatedness density and the variables analysed previously have a different effect on the emergence of specific pioneering sectors in a location. More specifically, I argue that different densities (relatedness, suppliers and clients) have a distinct effect on the emergence of new oil-related activities compared to manufacturing. To test this hypothesis I followed the methodology used previously in this chapter and focused on the extractive industry¹³ (metal mining and crude oil extraction) and manufacturing separately.

Table 5.6 shows the models for the extractive industry (oil and mining) and manufacturing for the periods t1 (2007-2008) and t5 (2007-2012) and follow the same structure of the results above. At first glimpse, the results do not contradict the main findings in this chapter: the related density is indeed a good indicator to signal the emergence of a pioneering industry, even if we analyse these economic sectors separately. Also, the density of suppliers is a better variable to signal the emergence of a new firm, than the density of clients (with some caveats that will be analysed later) and the importance of the densities analysed declined over time.

¹³The analysis focuses on the extractive industries as a whole because mining and oil extraction share some similarities on their production processes, as well as locations. Also, the number of entities associated to the oil industry is small and the results would not be representative

In contrast to the previous results, there are several differences between the industries analysed. First of all, the $pseudoR^2$ for the oil and mining industry is higher than manufacturing and than the overall analysis. So, the RD is a better explanatory variable for the emergence of pioneering industries in oil and mining sector, which means that the presence of extractive firms in a place is more important for this particular sector than for the manufacturing one. From a different angle, these results show that a smaller density of extractive industries for the emergence of new and related extractive firms, while a larger density of manufacturing firms is required for the emergence of pioneering manufacturing activities.

An interesting finding than can be also observed in these results concerns the importance of the clients. While the presence of clients in the manufacturing sector is not relevant for the emergence of similar sectors (it could even have a negative impact), the presence of clients in the extractive industry is almost as important as the suppliers, at least for $t1$.

VARIABLES	Oil and Mining (2007-2008)			Oil and Mining (2007-2012)			Manufacturing (2007-2008)			Manufacturing (2007-2012)		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Rel. Density (RD)	3.620*** (0.791) [37.34]			3.462*** (0.762) [31.89]			2.436*** (0.161) [11.43]			2.621*** (0.138) [13.74]		
Rel. Suppliers (RSP)		3.500*** (0.818) [33.12]			3.559*** (0.804) [35.14]			2.369*** (0.169) [10.69]			2.320*** (0.148) [10.18]	
Rel. Clients (RCL)			3.086*** (0.473) [21.89]			1.810*** (0.396) [6.108]			-0.625*** (0.0812) [0.535]			-0.223** (0.0743) [1.250]
City Relatedn. (CR)	1.054* (0.412) [2.869]	1.100** (0.424) [3.003]	1.328*** (0.310) [3.773]	1.339*** (0.385) [3.815]	1.277** (0.405) [3.584]	2.108*** (0.255) [8.23]	1.991*** (0.0863) [7.325]	2.031*** (0.0889) [7.619]	3.431*** (0.0720) [30.91]	2.023*** (0.0733) [7.564]	2.170*** (0.0775) [8.756]	3.046*** (0.0591) [21.03]
Pop. Density	-0.001 (0.001) [0.999]	-0.0007 (0.001) [0.999]	-0.002 (0.001) [0.997]	0.0007 (0.001) [1.001]	0.0008 (0.0012) [1.001]	0.0007 (0.0012) [1.001]	0.00006 (0.0004) [1]	0.0003 (0.0004) [1]	0.002*** (0.0004) [1.003]	-0.001 (0.0007) [0.999]	-0.0006 (0.0007) [0.999]	0.001 (0.0006) [1.001]
Prod. per Capita	-2.207* (0.972) [0.11]	-2.000* (0.957) [0.135]	-3.279** (1.042) [0.0377]	6.509*** (1.574) [671]	6.579*** (1.578) [719.9]	7.420*** (1.688) [1668.9]	-4.178*** (0.487) [0.0153]	-4.113*** (0.487) [0.0164]	-4.130*** (0.498) [0.0161]	-0.854** (0.331) [0.426]	-0.841* (0.332) [0.431]	-0.731* (0.338) [0.481]
City F.E.	Yes	Yes	Yes	Yes	Yes	Yes	YES	YES	YES	YES	YES	YES
Industry F.E.	Yes	Yes	Yes	Yes	Yes	Yes	YES	YES	YES	YES	YES	YES
Pseudo R ²	0.453	0.450	0.486	0.514	0.513	0.511	0.418	0.417	0.398	0.353	0.351	0.322
N	915	915	883	1265	1265	1245	30778	30778	29119	35969	35969	34080
Standard errors in parentheses, odd ratios in square brackets * p<0.05 ** p<0.01 *** p<0.001												

Table 5.6: Logistic Regressions - Oil and Mining Industry / Manufacturing (Entry)

Finally, when the machine learning model was tested to assess the potential of the model to predict the emergence of new industries, results show that the model's precision for the extractive sector is 0.81, while the precision for manufacturing is 0.76. The precision is higher than the overall model, but it is clear that the model has a higher predictive power

for the extractive industry. However, it is important to bear in mind that the sample size is smaller for this industry than to manufacturing. These differences can be clearly observed in the respective confusion matrices in figure 5.14.

Figure 5.13: Confusion Matrix

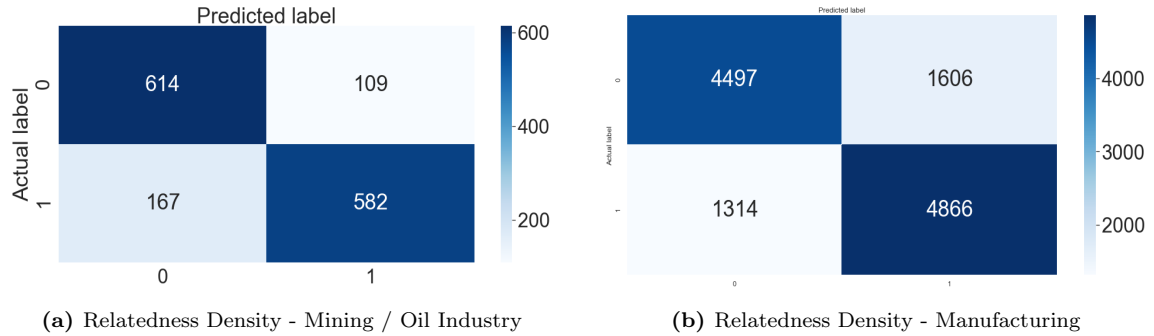
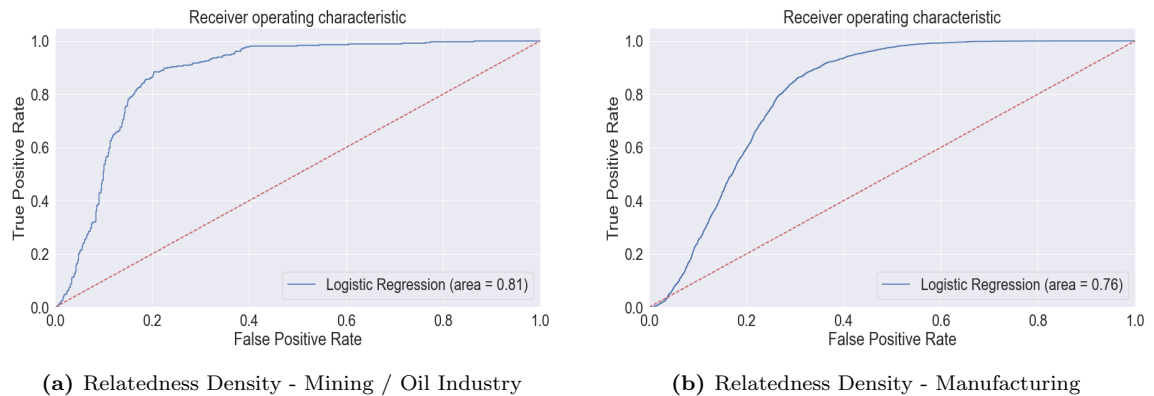


Figure 5.14: Area Under the Curve



5.5 Discussion

As it was explained in the literature review, as far as the author is aware, there are no studies that have explored productive diversification of oil exporting economies and regional diversification jointly. On one hand, existing literature on natural resource dependency has focused on the challenges that oil exporting economies have faced, having failed to shed light on how to overcome their dependency on few commodities and diversify their productive bases. On the other hand, the literature on regional diversification has not focused on countries that heavily rely on primary exports. In this chapter, I argued that

understanding regional productive dynamics are a key element to shed new light to the diversification debate in oil exporting economies.

Many years of research on regional diversification have found that exists a strong relationship between *how related economic activities in a place are* and *the emergence of new economic activities in related sectors*. This relation has been called "The Principle of Relatedness" by many economic geographers¹⁴ (Hidalgo et al., 2018). Due to data constraints, this research has focused mainly on the idea of relatedness based on co-location of economic activities and has ignored the fact that most of the economic activities located in a place are also part of wider networks of production and consumption.

In this chapter, based on one of the most complete datasets that encompasses all the firms (their industrial classification and locations) and the commercial relations between all of them, I explored the idea of co-location relatedness and the findings are coherent with the existing literature on this topic. This approach was further developed by incorporating information about production and consumption networks of each industry¹⁵. Results show that including information about these networks helps to better explain the emergence of pioneering industries in a place and signal more clearly future paths of economic diversification, as confirmed by the prediction power of the machine learning model (and the difference observed between the role that suppliers and clients play). Furthermore, the results suggest that knowing the location of clients does not help much to better understand the emergence of new industries, which can be counter intuitive. One may expect that clients are a key factor for the appearance of new industries, although these results show that clients may not be that relevant, as compared with the location and concentration of existing suppliers.

Despite the good results of our models, we need to bear in mind of several limitations that need to be addressed in future work. First, there is a bias in the firms' location and a lack of information on firms' establishments (specially for bigger firms). To have a more accurate picture of regional dynamics, these should be included. The Ecuadorian Government has implemented a new system of electronic billing which has information at establishment level, since 2017. Replicating this research with this data could generate new insights on regional diversification.

¹⁴César A. Hidalgo, Pierre-Alexandre Balland, Ron Boschma, Mercedes Delgado, Maryann Feldman, Koen Frenken, Edward Glaeser, Canfei He, Dieter F. Kogler, Andrea Morrison, Frank Neffke, David Rigby, Scott Stern, Siqi Zheng, Shengjun Zhu.

¹⁵Our literature review suggested that existing research on relatedness has not yet made a distinction between which firms are producers or which are consumers, nor explored their specific roles on diversification. This information is crucial because both producers and clients generate different incentives in an economy and their location could explain better some diversification dynamics

The previously described findings also have a shortcoming based on a non-explicit assumptions in the relatedness literature, in that it does not differentiate economic sectors and treats them as equals, despite the fact that each economic sector has its own characteristics. For instance, agricultural activities are strongly linked with specific locations (depending on the crops produce) and can be more widely spread in a territory. Usually, the growth of crops and their final consumption are linked with few intermediate nodes. On the contrary, manufacturing firms are usually less linked with a territory, since their production does not necessarily rely on land. Also, the production of complex goods requires longer and more diverse production chains. Therefore, the relatedness literature must pay more attention to the characteristics of specific industries.

The literature review of this chapter also pointed out that there are limited studies that have looked at and compared the influence that relatedness has on different industries. Therefore, in this chapter I explored the difference that exists between the extractive industry (oil and mining) and manufacturing. As it was reported above, there are some commonalities and differences between them, the most significant of which are related to the role of clients in each economic sector. Furthermore, based on the results, it can be argued that other economic activities would have other dynamics. For instance, the agricultural sector relies heavily on the resources available in a given location, while manufacturing might rely more on the proximity with physical infrastructure to produce and distribute, and the service industry might rely more on the location of the customers. The *Appendix C*¹⁶ contains the results of a brief analysis to address the aforementioned gaps by isolating three different industries to differentiate their nature and particular dynamics. The preliminary findings also show that each industry behaves differently as expected and the variables, which were previously assessed, vary depending on the type of economic activity. In the case of agriculture for instance, we observe that the *RD* effect is not significant, whereas the *DSP* is not only significant, but its effect is also more important when compared with manufacturing or commerce activities. Furthermore, for the case of commerce, the *DSP* in the same city matters more for predicting the emergence of pioneering related industries, in contrast with manufacturing or agriculture. More research is needed to explore these sectors in detail.

Finally, at the beginning of this chapter and throughout this thesis, it was stressed that the oil industry is one of the most important sources of income for Ecuador, as well as for many oil exporting economies, for obvious reasons. Therefore, the diversification of an economy also requires identifying the conditions needed to make firms more profitable, so they can generate more income for the country and eventually take over the extractive industries.

¹⁶This analysis was moved to the appendix because more tests are needed for making a stronger claim, however it clearly signals a path for future research.

Drawing on the work done by Mercedes Delgado (2018); Delgado et al. (2015, 2014), I started to explore these issues, as shown in the *Appendix D*, notably how geographical clustering and supply clustering foster industrial productivity. Using dynamic panel models, I explored how the variables calculated in this chapter influence the output of a firm. The preliminary results can establish the foundation for continuing this work.

5.6 Conclusions

This chapter had three main goals. The first was to explore two strands of literature that to date have not been analysed together: regional diversification and natural resources dependency. The second was to explore if the relatedness principle based on co-location holds, having tested if it is a driver for regional diversification in the context of an oil-exporting developing country. Finally, I contribute to the literature by exploring how production and supplier' networks influence the emergence of pioneering industries in a given location.

Our analysis confirms previous research in that relatedness based on co-location do hold for understanding the appearance of new industries in a place of an oil exporting economy, so the relatedness principle can be applied when analysis the emergence of pioneering industries in this particular context. This research also confirms that the existing industrial capabilities in a place are important to set or guide a diversification path. The results of this chapter provide the foundation to identify which new economic activities a country could develop to gradually move away from the resource sector.

When previous analyses of co-location relatedness were extended and information about the networks of suppliers and consumers was incorporated, it was found that the industrial relatedness based on the supply chain has a lower explanatory power than co-location. Also, when its predictive power was tested, the accuracy of the prediction of pioneering industries' emergence decreases when only the supply network is taken into account. However, when the consumption side (the networks of clients) was analysed, the results of the model were less relevant to the ones based only on co-location or based on the suppliers. Therefore, clients do not play an important role for regional diversification, as do supply networks.

We could observe how co-location and the networks of suppliers and clients influence the emergence of pioneering industries for the extractive industry and manufacturing. In this case, it was found that industrial relatedness based on co-location and based on the networks of suppliers have similar effects but different magnitude, depending on the industry. Our findings show that a smaller density of extractive firms could have a stronger effect on the emergence of similar industries in the region, while a larger density of manufacturing firms

would be needed for pioneering manufacturing industries to emerge in a place. Therefore, it could be more difficult to diversify towards manufacturing than towards other extractive industries. On the other hand, the networks of clients are more relevant to the extractive industries than to manufacturing.

Finally, it is important to note that these are not the only factors that drive diversification. Institutional players, from governments to private investors, whose decisions will be key for driving such processes of diversification. Similarly, as it was explained in the literature review, to foster this diversification, one also needs to take into account other elements related to knowledge creation, investment mobilisation, market formation and technology legitimisation that can be mobilised from inside or anchored from outside the region (Binz et al., 2016). The previous insights generated shed new light for a post-oil future in similar countries, but additional considerations are needed to drive these economies out of oil dependence.

Chapter 6

Conclusions and Implications

6.1 Thesis Overview

In this chapter, I bring together the analyses undertaken in the preceding chapters and discuss the contributions of the thesis as a whole. The chapter begins by summarising the motivation and objectives of this thesis. The following section is used to reflect upon the main findings of this research in relation to the three guiding research questions reiterated below. Based on these findings, I consider the policy implication, focusing mainly on diversification issues, and suggest pathways for future research.

As stated in the introductory chapter of this thesis, oil and mineral export-dependency has been a long-standing issue for many developing countries. Numerous studies have shown that there is a tendency of resource-rich countries to under-perform economically compared to resource-poor ones. To overcome these problems, many scholars have argued that the latter must diversify their productivity and significantly reduce their dependency on the primary sectors. Studies have also shown that there are many benefits of diversifying an oil-exporting economy. For example, Auty (2001), Lederman and Maloney (2006) and Ross (2013) explained that having a wider variety of products to export protects these countries against the effects of commodity prices' volatility (specially in the case of crude oil). Similarly, productive diversity could help create new jobs that are less reliant on the resource sector and thus on the rents created by natural resource extraction.

Notwithstanding the importance of this topic, little is known about how productive diversification can be achieved. This thesis addressed this issue and aimed to shed new light on crude oil export dependency and economic diversification. Herein, this thesis argued that one of the fundamental constraints for diversifying oil exporting economies pertains to the ways in which the economic structures oil-exporting countries are configured, as hypothesised by existing literature. More specifically, this thesis argued that understanding the

productive structures of these countries could help to explain why they have limited capacity to generate new economic activities, which are based on the productive knowledge they generate and accumulate. Furthermore, improving the understanding of these structures could provide insights to guide policy-making and investment decisions for diversification pathways.

To support the aforementioned arguments, this thesis is grounded on several strands of literature, which includes structuralism and economic growth, endogenous growth theory and the debates related with regional productive diversification. This literature informs us that the productive structures of a country are a clear determinant for their growth opportunities, and these structures are heavily linked with the knowledge and technology available and developed in a specific place, and that these processes of technological change are heavily path- and place- dependent. In summary, this means that a historical economic structure of a given place usually limits its development opportunities and is bounded by knowledge and technology that the society has accumulated or developed over time. These considerations are the foundations of this thesis and have been developed in the respective empirical chapters. Hereunder, I will summarised the guiding questions of this thesis and their contribution to existing literature.

6.2 Guiding questions and relevance to the literature

Underlying the main arguments of this thesis, I formulated three main research questions to guide this research:

- a. What is the productive structure of an oil exporting economy?
- b. How is this productive structure affected by a resource boom?
- c. How can this resource dependency be overcome?

These questions dealt with the specific components that build on each other to support the main argument of this thesis. The first two questions help us to understand the structural dependency of an oil-exporting country and how its economic structure has constrained the development of new productive activities. More specifically, the first question, which was addressed in Chapter 3, allowed us to comprehend how an oil exporting economy is configured. To accomplish this task, I proposed a network-based methodology to extract the structural information of input-output (I/O) data of a given country (Ecuador, in this case), as a way to study the productive structure of an oil exporting country. This methodology allowed us to assess the importance of different industries in a country, as measured by industrial centrality, which provides an approximation of its productive structure. It is worth

recalling that Input-Output analysis is a good tool for identifying key sectors based on the linkages between industries. However, this kind of identification cannot be performed when data are incomplete, especially in cases where only supply or use tables are available. The proposed methodology addressed this limitation by using the tools developed by network science. More specifically, it calculates the network centrality for different industries in cases when such cases are prevalent. By the end of this chapter, we could observe how each industry is linked with the rest of the economy and its relative importance for promoting growth.

The second question explored the effect that a resource boom has on the productive structure that was assessed previously. As explained before, I hypothesised that the structural impact of this boom is minimal because the oil industry operates as an enclave - it is weakly linked with the rest of the economy and its knowledge spill-overs are limited. This question was addressed in Chapter 4, where I explored the structural effect that the 1972 oil boom had on the productive structure of Ecuador. To do so, I assessed changes in industrial centralities from 1965 to 1985, as per calculated in Chapter 3. In particular, I argued that the economic shock generated by windfalls of oil resources had minimal impacts on the productive structure of an oil-dependent country. To test this hypothesis, I used a synthetic control method to identify a possible ‘structural’ crowding-out effect of the non-traded sectors on the economy (the Dutch Disease ‘structural’ effect). However, in this chapter we could observe changes in the centrality of different industries, but assessing the impact of the oil boom could not be undertaken due to methodological limitations.

The last question was a natural progression of the previous analyses and deals with the challenges of diversification. This question starts with the analysis of the current productive structure of the study case and based on the relatedness principle, assessed the likelihood of new industries to emerge in the country. Altogether, these questions helped us to understand how the productive structure of a country could limit its capacity to generate new economic activities, and considering it may set a path for future diversification. Also, these questions are a step forward to address the research gap identified by Michael Ross (2019) in relation with the literature on productive diversification of oil-exporting economies. This question was addressed in Chapter 5 and analysed how current concentrations of industries in a given place influences the emergence of “pioneering” economic activities (i.e., the “Principle of Relatedness”, which signals future paths for diversification within a country). Furthermore, how production and consumer networks influence the emergence of pioneering activities were examined. The extractive (oil and mining sector) and manufacturing was analysed in more detail.

In the following section, I will summarise the main findings and conclusions for each chapter.

6.3 Main findings and conclusions

Previous research has shown that there is a positive relationship between economic growth and how industries are interconnected within an economy (Kubo et al., 1986). In their findings, Kubo et al. (1986) argued that a higher concentration of manufacture goods, which are the biggest consumers of intermediate inputs (i.e., with higher industrial centralities), is key for fostering sustainable economic growth. In this research, particularly in Chapter 3, it was observed that the industries that generate more income for Ecuador (i.e., extraction or direct use of natural resources such as agriculture, mining or oil) are the ones that have the lowest centralities. This means that the country relies mostly on industries which are relatively isolated from the rest of the economy and are not big consumers of intermediate inputs, thus potentially be one of the main constraints for growth in Ecuador.

The results obtained in this chapter also showed a clear dichotomy commonly experienced in developing countries that rely mainly on crude oil exports. On the one hand, the oil industry generated approximately 57.6% of the [exports] income of Ecuador and represented 13.6% of its GDP in our study period, which makes it the most important industry in the country in absolute terms (BCE, 2019). Nevertheless, when its importance is measured based on how each sector is connected in the Ecuadorian economy (centrality), as suggested by Meller and Marfan (1981), the oil industry has the lowest centrality compared to other sectors. This finding has been confirmed by previous literature, which states that crude oil has fewer linkages and shares the fewest common characteristics with other products in an economy Ahmadov (2014) Lederman and Maloney (2012).

In Chapter 4, we observed the differences that exist when we analysed each industry individually based on aggregate measurements in comparison with a network approach. Under the traditional aggregate measurements (aggregated outputs of each sector), the importance of a sector is linked with the magnitude of intermediate consumption and how it changes. However, when we looked at the network measures (centralities), which portrayed a different picture, since their importance is measured in terms of how connected they are. In this sense, one of the most important takeaways of this chapter is that the centrality of the primary sectors, which generates most of the income for Ecuador, did not change over time, in contrast with the aggregate measurements that are constantly increasing.

Furthermore, in this chapter we observed that the centrality of crude oil grew significantly some years before the oil boom. Literature from this period explains that this change was

caused by the investments triggered by Texaco's oil discoveries in the country (Belli, 1984; Vos, 1987; Larrea, 2013). According to Gelb (2011), these effects are expected because the sector usually gains in importance during the construction phase of the infrastructure needed for extracting the resource. Moreover, we also observed that the centrality of manufacturing, machinery, construction and transportation increased in the years that followed the boom. Existing historical literature suggests that these changes were due to industrial policy and ISI reforms implemented by the Ecuadorian Government after the oil boom.

In Chapter 4, we also observed that the amount of links that the oil industry has with other primary products is minimal and it was expected that the oil boom would not have a strong effect on the other non-oil traded products. On the other hand, there are few sectors that are more connected with the oil sector, such as the manufacturing and construction industries. Since these sectors depend directly on the oil revenues inflows, the oil boom and the variation of oil prices would affect them directly. However, we could not test the effects of the oil boom in the industrial centralities, as explained in the respective chapter. The synthetic control method used was not fruitful due to data constraints and the lack of useful variables to build the needed controls.

One of the main findings of Chapter 5 also confirmed previous research. Relatedness based on co-location did hold true for understanding the appearance of new industries in a place of an oil exporting economy. The relatedness principle can thus be applied when analysing the emergence of pioneering industries in this particular context. This research also confirmed that the existing industrial capabilities in a place are important to set or guide diversification. The results of this chapter provided the foundation to identify which new economic activities could be developed to gradually move away from the resource sector.

When extending previous analyses of co-location relatedness and incorporating information about the networks of suppliers and consumers, our results suggest that the industrial relatedness based on the suppliers has not a better explanatory power than just co-location to predicting emergence of new economic activities. Also, when its predictive power was tested, the accuracy of the prediction of pioneering industries' emergence decreased when information regarding the suppliers was considered. However, when the consumption side (i.e., the networks of clients) was analysed, the results of the model were significantly smaller to the ones based only on co-location or based on suppliers only. Therefore, in terms of regional diversification, clients do not seem to play a role as significant as supply chains or firms co-location.

In this chapter, it was tested how the relatedness density (co-location) and the density of suppliers and clients affected the emergence of pioneering firms in the extractive industry

(oil and mining), in comparison to manufacturing. I found that the previous findings hold but with important differences. For the oil/mining sector the density of clients was more important for our predictions, while clients were not relevant for manufacturing. Also, I found that new industries related to the oil and mining sector to emerge, a smaller density of these firms is required in a given place, in comparison to manufacture. This means that it is easier to have new extractive firms in a place with a smaller agglomeration of pre-existing extractive firms, in comparison to the agglomeration of pre-existing manufacturing firms needed for the emergence of new manufacturing firms. In other words, a place with a higher density of extractive firms has the propensity to create similar firms whereas to diversify manufacturing, more efforts would be required.

As a general conclusion, it can be argued that the productive structure of a country does matter for its economic development and could limit its diversification potential. We observe that the Ecuadorian economy relies mostly on industries with the lowest degree of connectivity, which do not have the capacity to stimulate other sectors effectively. This situation has not changed much since the 1972 oil boom. Furthermore, we observed that the productive configuration of an economy could limit the productive knowledge that an economy can acquire or develop, and future development paths depend on this configuration.

The Ecuadorian productive structure relies heavily on crude oil. This commodity has indeed been central for providing income for the country and has been key for its growth. However, the oil industry has not stimulated the emergence of new economic sectors, not has it created knowledge spill-overs that would help the country to diversify. This phenomenon is well known in the literature, since the oil industry operates in countries as enclaves¹ and strongly limits "learning by doing".

Regarding economic diversification, the current concentration of productive activities could signal the emergence of new economic activities in the country, as explained in Chapter 5. Nevertheless, it is important to note that these are not the only factors that drive diversification. Institutional players, from governments to private investors, whose decisions are key for driving such processes of diversification, also play a crucial role. In addition, as it was explained in the literature review, fostering diversification requires taking into account other elements related to knowledge creation, investment mobilisation, market formation and technology legitimisation that can be mobilised from inside or anchored from outside the country (Binz et al., 2016).

¹The oil industry has been considered by many scholars as an economic enclave, which means that the industry is disconnected and delinked from the rest of the national economy (Bloch and Owusu, 2012). More generally, an enclave industry is associated with a lack of productive, physical backward and forward linkages (Bloch and Owusu, 2012; Hirschman, 1958)

The final contribution of this thesis pertains to the data generated during this research. For the first two analytical chapters, I digitised all data contained in the National Accounts since 1965 to 1995, which is going to be made publicly available for download by the Ecuadorian Central Bank. For chapter 5, I cleaned and organised the most comprehensive dataset of firms and transactions of the country. For the purposes of data processing, I created several visualisation tools² that would allow researchers to explore this data easily. As part of this effort, I created an interactive tool to explore all the production chains in the country, which can be found here: [http : //production.data593.com/prod_index.html](http://production.data593.com/prod_index.html).

6.4 Implication for future research

As I was developing this research, I identified many areas for future work and I already started working on few of them (see appendixes C and D). Below, I briefly explain some of these areas.

Firstly, in this thesis I focused only on one country, which makes it difficult to generalise the main findings. However, as explained in the beginning of this thesis, the current consensus among experts is that studying resource-rich countries is largely context-specific and has to be done case by case (Papyrakis, 2017). Therefore, since the findings of this study are coherent with existing research, the methodology could be applied to other oil exporting economies, but the synthetic control method needs to be used differently and it would require more efforts to identify more useful controls. This could yield new insights of specific countries and we could test if these thesis' findings can be generalised.

Secondly, one of the limitations of this study (for the first chapters) is data aggregation, which was already mentioned in Chapters 3 and 4. Studying only 30 aggregated sectors and products gives us a glance of the reality and masks a lot of internal productive dynamics. Therefore, it is recommended to test the proposed methodologies using data at a lower level of aggregation. This could give us new and complementary insights into the structural changes caused by an oil boom. Also, for the case of Ecuador, there is the need to link the historical data with current data. As it was stated in Chapter 3, it was not possible to use a longer data frame, since the country stopped publishing national accounts between 1996 and 2006.

For future research, I also suggest exploring how sector and product centralities change due to commodity prices fluctuations. The proposed methodology can be used to study the effects of the sudden increases and declines of oil prices experienced in the last decade and

²www.pastudillo.com/ecuador/ password: dataec2020

see the structural effects on the economy. The proposed methodology could also be used to examine how productive structures have reacted to different kinds of external events, beyond resources booms, as long as relevant controls are identified.

As it was explained in Chapter 5, *Appendix C* contains the results of a preliminary analysis to explore three different industries in isolation (agriculture, manufacturing and commerce). The preliminary findings show that each industry behaves differently, as expected, and the variables, which were previously assessed, vary depending on the type of economic activity. In the case of agriculture, for instance, we observe that the *RD* (Relatedness Density) effect on the emergence of new industries is not significant, whereas the *DSP* (Relatedness of Suppliers) is not only significant, but its effect is also more important when compared with manufacturing or commerce activities. Furthermore, in the case of commerce, the *DSP* in the same city matters the most for predicting the emergence of pioneering related industries, in contrast with manufacturing or agriculture. More research is needed to explore these sectors in detail.

Finally, at the beginning of this chapter and throughout this thesis, it was stressed that the oil industry is one of the most important sources of income for Ecuador, as well as for many oil exporting economies, for obvious reasons. Therefore, the diversification of an economy also requires identifying the conditions needed to make firms more profitable, so that they can generate more income for the country and eventually take over the extractive industries. Drawing on the work done by Mercedes Delgado (2018); Delgado et al. (2015, 2014), I started to explore some of these issues, notably how geographical clustering and supply clustering foster industrial productivity, as shown in *Appendix C*. Using dynamic panel models, I explored how the variables calculated in this chapter influence the output of a firm. The preliminary findings shows that there is a positive effect of clustering and output, in accordance with the literature. However, these models should be tested in more detail and it is recommended to run more comprehensive robustness analysis.

6.5 Implication for public policy

As stated throughout this thesis, there is a tendency of resource-rich countries to underperform economically compared to resource-poor ones. Ecuador is a clear example of this. The country's economy has historically relied on few commodities and its economic crises have been linked with variations of their international prices (Acosta, 2009a; Correa, 2012). As I write this section, the Ecuadorian government announced strong economic measures³

³These measures include public spending reduction, tax reforms and a compulsory unique payment from people who own cars that are valued higher than US\$ 20,000.

to deal with the sudden drop of oil price experienced by March 2020 in relation to the Covid-19⁴, and a new recession is expected (Orozco, 2020). Moreover, as it was stated previously in the thesis, a recent study suggested that crude oil production in Ecuador would reach its peak between 2014 and 2025 and that the country could become a net oil importer by 2035 (Espinoza et al., 2019). Oil trade balance could decrease from a current positive value of around 2 billion USD to incur deficits of 0.6 – 16.7 billion USD in 2035 (Espinoza et al., 2019). As such, it is urgent for the Ecuadorian Government and the private sector to take concrete actions to diversify the country's productivity and significantly reduce their dependency on the primary sectors.

As a first step to diversify the Ecuadorian economy, policy makers need to understand that possible development paths depend on a country's past. As David (2001), Page et al. (2006), and others have stated, development processes in general are heavily path- and place- dependent, since they rely on the experience that an economy has accumulated over time. Likewise, these processes rely on the society's capacity to absorb and develop new knowledge and know-how (Hausmann et al., 2014, 2007b). These considerations are particularly important when we are in the context of oil exporting economies. The reason is that the economic configurations in these countries have historically relied on rents produced by natural resources and the skills these societies have accumulated are usually rent-seeking, as argued by Cockx and Francken (2016), Wadho (2014) and Gylfason (2001). In this sense, policy makers need to generate the necessary incentives to develop new productive skills in the country, and increase the density of the sectors that need to be developed. For instance, the results of this thesis suggest that if new manufacturing activities are to be developed, the density of this sector should be significantly higher than the density of the extractive sector. In other words, if Ecuador wants to diversify towards manufacturing, it has to promote the creation of new manufacturing firms (i.e., increase the density) with updated skills, as a way to gain experience and create know-how.

As explained in the previous chapters, extractive industries, which are the most important generators of income of a primary economy, generate limited knowledge spill-overs and they usually require low quantities of high-skilled labour. Therefore, these economies do not have enough incentives for knowledge creation in areas different from the resource sector. In this sense, policy makers need to create the needed incentives to attract investment (or devote government investment) to industries that generate knowledge gains and that could interact with more sectors of the Ecuadorian economy. According to the literature and this thesis,

⁴By March 9, crude oil prices crashed "the most in 29 years as Saudi Arabia and Russia vowed to pump more in a battle for market share just as the coronavirus [Covid-19] spurs the first decline in demand since 2009" (Murtaugh et al., 2020).

Ecuador should prioritise investment in sectors that have higher industrial centralities, such as manufacturing or services.

A few years ago, there was an ambitious - but short-lived - plan to diversify the Ecuadorian economy. This plan set six "strategic" industries that the country should pursue: petrochemical, steel, copper refining, aluminium, shipyards and paper pulp (Vicepresidency of Ecuador, 2014; SENPLADES, 2013). Such industries were prioritised based on several consultancies that tried to imitate industrialisation paths that other countries followed in the 1970s. However, as it can be observed, none of these industries existed in Ecuador⁵. The plan was not considering the experience and know-how that already existed in Ecuador, which contradicts most of the literature regarding regional diversification. It is worth recalling that the only attempt to achieve diversification goals was the preparation of the terrain to build a new refinery, which costed around US\$ 1,501 million. The country, however, never found sources of financing to build such refinery. As we could observe in this thesis, research has shown that regional diversification should be conducted based on the existing capabilities and experience that a place has already acquired. This allows to explore new productive paths, which are usually related with the existing industries in a given place (i.e., Principle of Relatedness). Therefore, based on the analysis conducted in Chapter 5, the model developed signals the pioneering industries that Ecuador could develop at a county level. These results could provide guidance on the diversification paths that the country could follow. A summary of these industries can be found in *Appendix E*.

It is important to recall that the previous recommendations are proposed mostly from the knowledge and know-how perspective in order to guide diversification. However, as stated in Chapter 5, all the efforts have to be accompanied with other policies, such as investment mobilisation, market formation and technology legitimisation that can be mobilised domestically or anchored from outside the country (Binz et al., 2016). The government needs to provide investment to foster some industries, as well as the respective fiscal incentives to private companies to explore new productive venues. Above all, however, the government needs to recognise the urgency to reduce the dependency on crude oil.

6.6 Limitations of this research

The generalisability of the results from this thesis is subject to certain limitations. For instance, as explained in the beginning of this document, the current consensus among experts is that studying resource-rich countries is largely context-specific and has to be

⁵The oil industry in Ecuador is mainly focused on crude oil extraction and refining. However, there are not limited experience on petrochemistry.

done case by case (Papyrakis, 2017). In this particular case, this thesis only examined Ecuador because of data availability and accessibility, and the results obtained might be only applicable to this context. However, since our findings are coherent with existing research, most of the methodologies used could be applied to other oil exporting economies and tested if the findings for Ecuador could be found in other countries, with the exception of the synthetic control method. In this case, new controls need to be found and new analysis must be done.

Despite the fact that most of the proposed methodologies could be use and tested in different countries, finding and accessing historical data, as the ones used in this thesis, could be difficult. The data required by these methodologies could be difficult to find, because it is usually in central banks archives and ministries of economy, and not in a digital format. Also, there are countries that may not have these records or the data sets could be inaccessible, since they are proprietary or not publicly available. It is important to recall that the data used for Chapter 5 was obtained after several months of negotiation and under a non-disclosure agreement.

Another limitation of this study, in particular for Chapters 3 and 4, is the level of aggregation of the data. Studying only 30 aggregated sectors and products only gives us a glance of the reality, but it masks a lot of internal productive dynamics. Also, it is important to recall that this thesis focuses only on the formal economy, that represents almost half of the Ecuadorian real economy. For the study of economic diversification, Chapter 5 used a new and detailed dataset at firm level, which is by far more disaggregated but still does not covers the informal economy.

Although I was aware of these limitations, one of the main motivations underlying this research was to understand developing economies that have data deficiencies. As far as I am aware, these data constraints are common in most developing countries for the period of this study.

Another limitation of Chapter 3 and 4 pertains to the time frame analysed. In this study, we observed how the centrality of different industries changed between 1965 and 1985. However, we do not know how they evolved later on. As previously mentioned, we could not use Ecuador for a longer data frame, since the country stopped publishing national accounts in 1996 and only continued after 2006. When this new data was published, the methodologies, the base years and the aggregation level changed. In addition, the country changed its currency in 2000, which means that the IO data prior to 1994 and after 2006 are incompatible.

A reasonable approach to tackle the previous limitation could be to extend this study to other countries that have data at a lowest level of aggregation. This could have given us new insights into the structural changes caused by an oil boom. Also, we could conduct this type of study for countries that experienced resource booms in the last 20 years.

For future analysis, we could also explore how sector and product centralities change due to commodity prices fluctuations. We could use the proposed methodology to study the effects of the sudden increases and declines of oil prices experienced in the last decade and see the structural effects on the economy. The proposed methodology could also be used to examine how productive structures have reacted to different kinds of external events, beyond resources booms.

Finally, it is important to note that the scope of this study only focuses on the structural aspects of diversification and its association with knowledge, which are not the only factors that drive diversification. Institutional players, from governments to private investors, whose decisions are all key to drive such processes of diversification. Similarly, as it was explained in the literature review, to foster this diversification, one also needs to take into account other elements related to knowledge creation, investment mobilisation, market formation and technology legitimisation that can be mobilised from inside or anchored from outside the region (Binz et al., 2016). The previous insights generated shed new light for a post-oil future in similar countries, but additional considerations are needed to drive these economies out of oil dependence.

Appendix A

Bipartite Network Projections

In table 3.5, industries (X-nodes) and products (Y-nodes) should be the same type in figure 1 (i.e., industries (X-nodes) with industries (X-nodes) or products (Y-nodes) with products (Y-nodes)).

Figure 1: Bipartite to Unipartite Projection

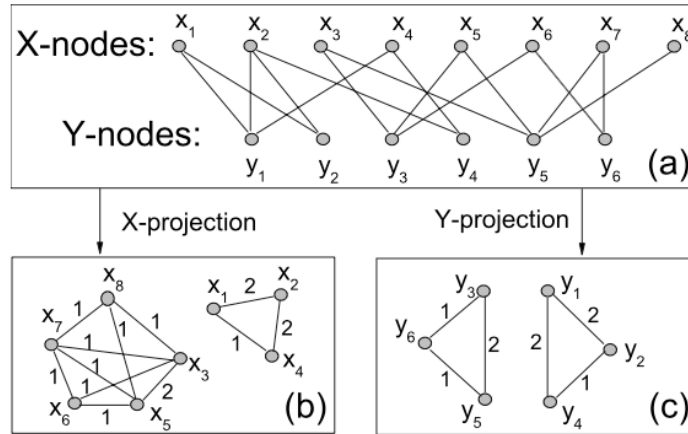


Illustration of a bipartite network (a), as well as its X projection (b) and Y projection (c). The edge weight in (b) and (c) is set as the number of common neighbors in Y and X, respectively (Zhou et al., 2007).

Newman (2010) contends that when we form a one-node projection, each group in the bipartite network results in a cluster of nodes in the one-mode projection, which are all connected to each other, as can be seen in figure 1. For instance, if a group contains four members in the bipartite network, then each of those members are connected to each of the others in the one-node projection by virtue of common membership to that group (Newman, 2010). He further explains that the projection can be mathematically written in terms of the incidence matrix B . The product $B_{ki}B_{kj}$ will be 1 if only if i and j both belong to the same group k in the bipartite network. Thus, the total number P_{ij} of groups to which both i and j belong is:

$$P_{ij} = \sum_{k=1}^g B_{ki} B_{kj} = \sum_{k=1}^g B_{ki}^T B_{kj}$$

where P_{ij} is new matrix projection (unipartite network), B_{ij} is the incidence matrix (Use Tables), g is the number of groups (sectors or products) and k is a single group. Newman (2010) explains that B_{ki}^T is an element of the transpose B^T of B . The $n \times n$ matrix $P = B^T B$ is similar to an adjacency matrix for the weighted one-mode projection onto the n vertices. Its off-diagonal elements are equal to the weights in that network, the number of common groups shared by each node pair. P is not quite an adjacency matrix, however, since its diagonal elements are non-zero, even though the network itself, by definition, has no self-edge. The diagonal elements have the following values:

$$P_{ii} = \sum_{k=1}^g B_{ki}^2 = \sum_{k=1}^g B_{ki}$$

where we have made use of the fact that B_{ki} only takes the values 0 or 1. Thus P_{ii} is equal to the number of groups to which vertex i belongs.

Thus to derive the adjacency matrix of the weighted one-mode projection, we calculated the matrix $P = B^T B$ and set the diagonal elements equal to zero. Finally, to derive the adjacency matrix of the unweighted projection, we took the adjacency matrix of the weighted version and replaced every non-zero matrix element with a 1 (Newman, 2010).

The other one-mode projection, onto the groups, can be represented by a $g \times g$ matrix $P' = B B^T$, whose off-diagonal element P'_{ij} gives the number of common members of groups i and j , and whose diagonal element P'_{ii} gives the number of members of group i (Newman, 2010). This projection was done with all the Use Tables of Ecuador since 1965 to 1995, using the NetworkX package in Python. This resulted in two unipartite networks: one for products and one for sectors of industries, as shown in table 1 and table 2, respectively.

The one-node projection is often useful and widely employed, and its simplest method involves projecting the bipartite network onto an un-weighted network, which does not take into account the topology of the network or the frequency of sharing a connection to the elements of the opposing set (Zhou et al., 2007). Moreover, this projection discards a lot of the information present in the structure of the original network and hence it is a less powerful representation of the data (Newman, 2010; Zhou et al., 2007). However, in order to get a more informative projection, the use of some weighing method is required, which

Table 1: Unipartite Network - Industries

		Industries				
		1	2	3	...	30
Industries	1 - Banana coffee cocoa	0	0	0	...	1
	2 - Other agricultural products	0	1	1	...	1
	3 - Animal products	0	0	1	...	1
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	30 - Hotels, bars and restaurants	0	1	0	...	1

could capture the links' weights (between two nodes in the projected network) (Zhou et al., 2007).

Table 2: Unipartite Network - Products

		Products				
		1	2	3	...	30
Products	1 - Banana coffee cocoa	0	1	0	...	0
	2 - Other agricultural products	0	0	0	...	0
	3 - Animal products	0	1	0	...	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	30 - Hotels, bars and restaurants	1	1	1	...	0

So far, the result we have is a unipartite network without links' weight. However, as it was discussed before, in order to have a more comprehensive picture of the network, the corresponding weights were calculated. Three different projections that infer weights statistically, for both the industries and the products, will be explored below: simple weighed collaboration, jaccard or overlap projections.

Simple weighed projection

The simple weighted projected graph is the projection of the bipartite network B_{ij} onto the specified nodes (could be the products or the industries) with weights representing the number of shared neighbours or the ratio between actual shared neighbours Borgatti and Halgin (2011); Zhou et al. (2007). The nodes retain their attributes and are connected in the resulting graph if they have an edge to a common node in the original graph. See figure 2a for the projection on 'products' and figure 2b for the projection on 'sectors'.

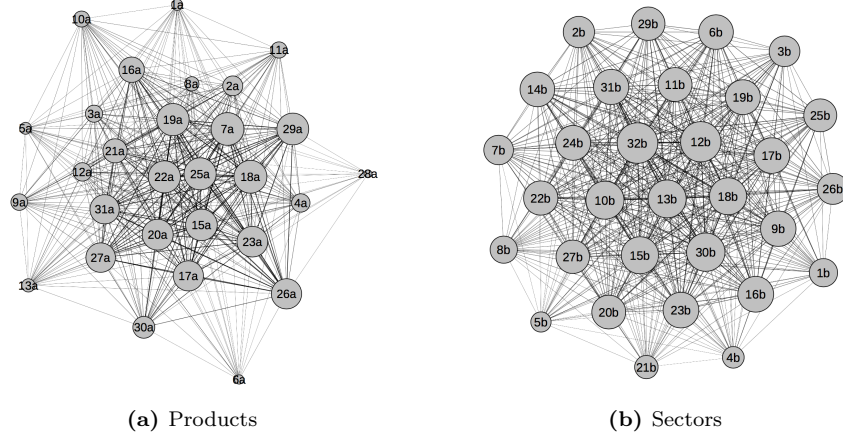


Figure 2: Weighed Projection

Collaboration projection

Collaboration projection is one of the most common projections and it is based on what is called a collaboration network. This means that generally defined as a network of actors connected by a (any) common collaboration act (Zhou et al., 2007). This is a projection of the bipartite network B_{ij} onto the specified nodes with weights assigned using Newman's collaboration model:

$$w_{ij} = \sum_{k=1} \frac{\delta_i^k \delta_j^k}{n_k - 1}$$

where i and j are nodes from the same bipartite node set, and k is a node of the opposite node set. The value n_k is the degree of node k in the bipartite network and δ_i^k is 1 if node i is linked to node k in the original bipartite graph or 0 otherwise. The nodes retain their attributes and are connected in the resulting graph if they have an edge to a common node in the original bipartite graph (NetworkX, 2015a; Newman, 2001). See figure 3a for the projection on 'products' and figure 3b for the projection on 'sectors'.

Jaccard or overlap projection

Similar to the previous cases, the weights represent the Jaccard index between the neighbourhoods of the two nodes in the original bipartite network (Borgatti and Halgin, 2011):

$$w_{i,j} = \frac{|N(j) \cap N(i)|}{|N(j) \cup N(i)|}$$

The nodes retain their attributes and are connected in the resulting graph if they have an edge to a common node in the original bipartite graph (NetworkX, 2015b). See figure 4a

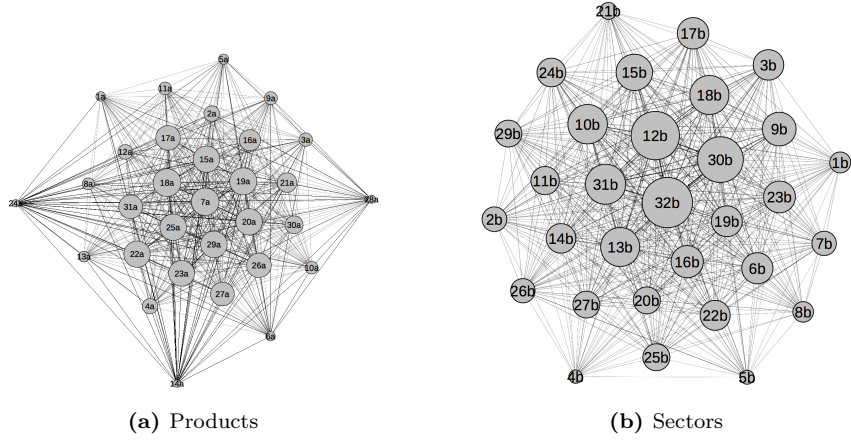


Figure 3: Collaboration Projection

for the projection on 'products' and figure 4b for the projection on 'sectors'.

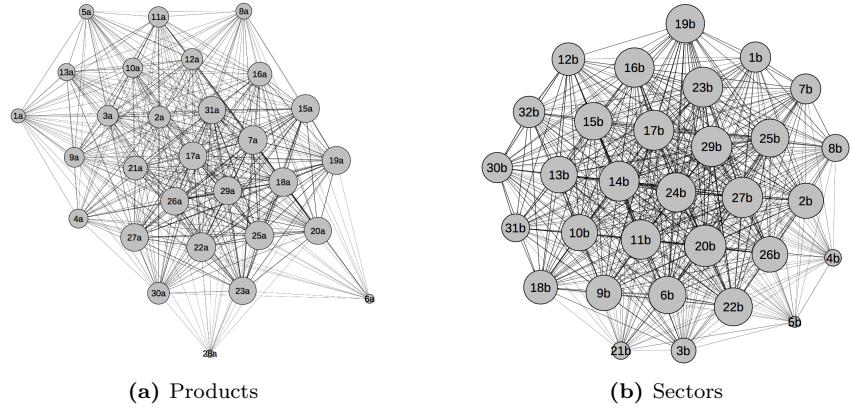


Figure 4: Jaccard Projection

Table 3, illustrates the different properties resulting from the three projections. For the analysis, we will use only the Simple Weighted projection, as the results are coherent with the Jaccard projection.

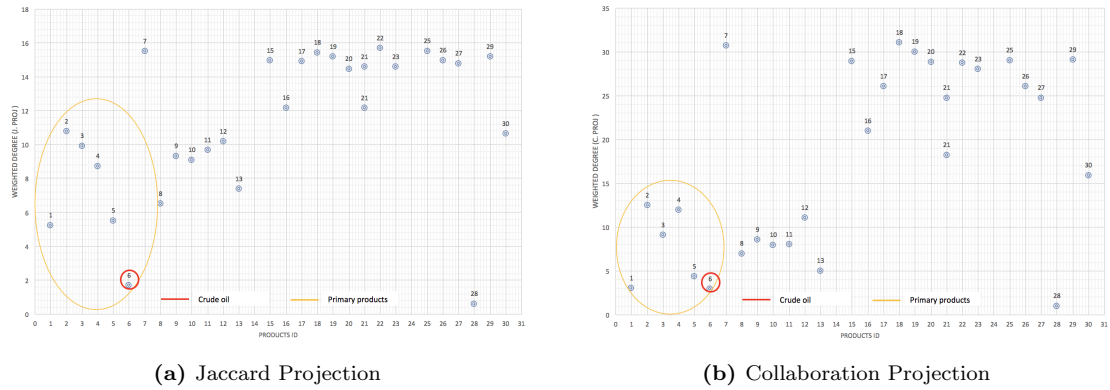
Table 3: Projections' network properties

	Simple (prod)	Simple (sect)	Collab (prod)	Collab (sect)	Jacc (prod)	Jacc (sect)
Average degree	13.214	15	15	15	13.31	15
Network diameter	2	1	1	1	2	1
Graph density	0.475	0.5	1	0.5	0.475	0.5
Modularity	0.031	0.008	0.035	0.055	0.016	0.025
Connect. compnts.	1	1	1	1	1	1
Avg. clust. coef.	0.484	0.5	0.5	0.5	0.484	0.5
Avg. path length	1.035	1	1	1	1.035	1

Once these projections were done, the sectors' and products' strengths were calculated based on their centrality. To do so, we explored two different types of network properties: weighted centrality and eigenvector centrality.

Figure 5 shows that the results are coherent when we use the Jaccard and Collaboration Projections, with some differences in the magnitude of weighted degree. As such, we can use either projection and we will get similar results in terms of the importance of the products analysed.

Figure 5: Products relevance, based on weighted centrality



Appendix B

Comparison Between Models: Logistic, Probit and Linear Probability (OLS)

In this appendix we can observe the comparison between all the parameters of the logistic, probit and linear probability regressions done in Chapter 5. Figure 6 shows the performance of different models and the associated normalised results. Herein we can observe the numeric correlation between them. Table 4 sums the parameters of different models: the resulting parameters on the left side, and the normalised parameters on the right side. The latter were used for the graphic comparison previously shown.

Figure 6: Comparison of Normalised Regression Parameters

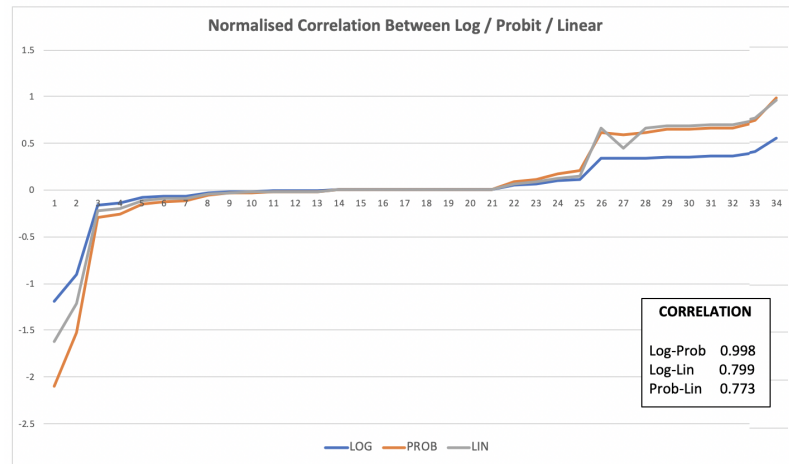


Table 4: Comparison of Regression Parameters

PARAMETERS			NORMALISED PARAMETERS		
LOG	PROB	LIN	LOG	PROB	LIN
7.585	4.042	1.18	-1.19353278	-0.902670724	0.476717771
5.756	2.778	0.766	-0.905731665	-0.620390715	0.309462553
1.043	0.588	0.176	-0.164120592	-0.131313802	0.071103667
0.9	0.514	0.15	-0.141618919	-0.114787915	0.060599717
0.508	0.29	0.0863	-0.079936012	-0.06476361	0.034865037
0.423	0.242	0.0716	-0.066560892	-0.054044116	0.028926265
0.401	0.229	0.0679	-0.063099096	-0.051140919	0.027431472
0.176	0.107	0.03	-0.027694366	-0.023895539	0.012119943
0.122	0.0716	0.0192	-0.019197231	-0.015989912	0.007756764
0.0851	0.053	0.0143	-0.013390856	-0.011836108	0.005777173
0.0743	0.047	0.0125	-0.011691429	-0.010496171	0.005049976
0.0689	0.0438	0.0116	-0.010841715	-0.009781538	0.004686378
0.0573	0.0374	0.0097	-0.009016405	-0.008352272	0.003918782
0.000153	0.000085	0.0000253	-2.40752E-05	-1.89824E-05	1.02212E-05
0.000144	0.0000817	0.0000243	-2.2659E-05	-1.82455E-05	9.81715E-06
0.000121	0.000069	0.0000205	-1.90399E-05	-1.54093E-05	8.28196E-06
0.000119	0.000067	0.0000203	-1.87252E-05	-1.49626E-05	8.20116E-06
0.0000979	0.000055	0.0000168	-1.5405E-05	-1.22828E-05	6.78717E-06
0.000078	0.000045	0.0000135	-1.22736E-05	-1.00495E-05	5.45397E-06
-0.000236	-0.000136	-0.0000376	3.71356E-05	3.03719E-05	-1.51903E-05
-0.00416	-0.00189	0.000376	0.000654594	0.00042208	0.000151903
-0.302	-0.169	-0.054	0.047521015	0.037741552	-0.021815898
-0.407	-0.218	-0.0652	0.064043222	0.048684369	-0.026340677
-0.609	-0.343	-0.106	0.095828802	0.076599718	-0.0428238
-0.729	-0.41	-0.128	0.114711325	0.091562345	-0.051711758
-2.137	-1.251	0.111	0.336266256	0.279376812	0.04484379
-2.144	-1.154	-0.35	0.337367736	0.257714502	-0.141399339
-2.154	-1.261	0.109	0.33894128	0.28161004	0.044035794
-2.247	-1.323	0.0933	0.353575235	0.295456053	0.037693024
-2.276	-1.339	0.0885	0.358138511	0.299029218	0.035753833
-2.279	-1.342	0.0876	0.358610574	0.299699186	0.035390235
-2.282	-1.342	0.0872	0.359082637	0.299699186	0.035228635
-2.55	-1.458	0.0611	0.401253604	0.32560463	0.024684285
-3.435	-1.909	-0.0548	0.540512208	0.426323209	-0.022139096

Appendix C

Comparison of Regression Models: Agriculture, Manufacture and Commerce

To compare the different industries, I extracted from the main dataset the industries subject to this analysis based on their ISIC Classification at one digit level. I run the model as per specified below. It is the same model used in Chapter 5 but applied to every individual economic sector: Agriculture, Manufacture and Commerce.

$$Entry_{c,i,t} = \beta_1 Density_{i,c,t-1} + \beta_2 City_{c,i,t-1} + \beta_3 Controls_{i,c,t-1} + \phi_c + \psi_i + \alpha_t + \varepsilon_{i,c,t} \quad (1)$$

The preliminary findings show that each industry behaves differently, as would be expected and the variables, which were previously assessed, vary depending on the type of economic activity. In the case of agriculture for instance, we observe that the *RD* effect is not significant. In comparison, the *DSP* is not only significant, but its effect is more important when compared with manufacturing or commerce activities. Furthermore, in the case of commerce, the *DSP* in the same city matters more for predicting the emergence of pioneering related industries, in contrast with manufacturing or agriculture. More research is needed to explore these sectors in detail.

Table 5: Logistic Regressions Per Industry

VARIABLES	Agric. 1 <i>Entry_{c,i,t}</i>	Ind. 1 <i>Entry_{c,i,t}</i>	Comm. 1 <i>Entry_{c,i,t}</i>	Agric. 2 <i>Entry_{c,i,t}</i>	Ind. 2 <i>Entry_{c,i,t}</i>	Comm. 2 <i>Entry_{c,i,t}</i>	Agric. 3 <i>Entry_{c,i,t}</i>	Ind. 3 <i>Entry_{c,i,t}</i>	Comm. 3 <i>Entry_{c,i,t}</i>
Relatedness Density (RD)	-0.476 [0.622] (0.298)	-1.030*** [0.357] (0.149)	0.365** [1.440] (0.131)						
Relatedness Suppliers (RSP)				6.064*** [430.2] (0.238)	4.282*** [72.41] (0.112)	5.242*** [189.1] (0.108)			
RSP Same City							3.381*** [29.40] (0.264)	2.687*** [14.69] (0.113)	6.479*** [651.4] (0.179)
Population Density	0.000257* [1.000] (1.02e-4)	0.000406*** [1.000] (3.71e-5)	0.0000136 [1.000] (3.40e-5)	-0.000426*** [1.000] (1.27e-4)	0.0000804* [1.000] (3.27e-5)	-0.000190*** [1.000] (3.38e-5)	0.000242* [1.000] (0.95e-4)	0.000314*** [1.000] (3.38e-5)	0.0000575 [1.000] (3.28e-5)
Production Per Capita	0.0413 [1.042] (0.0999)	0.324*** [1.383] (0.0412)	-0.145*** [0.865] (0.0327)	-0.0687 [0.934] (0.106)	0.161*** [1.175] (0.0421)	-0.202*** [0.817] (0.0333)	0.0673 [1.070] (0.100)	0.279*** [1.322] (0.0412)	-0.114*** [0.892] (0.0328)
City Relatedness (CR)	3.127*** [22.81] (0.314)	13.88*** [1069852.8] (0.280)	-3.463*** [0.0313] (0.0925)	-0.137 [0.872] (0.275)	7.096*** [1207.0] (0.239)	-5.427*** [0.00439] (0.0861)	2.202*** [9.044] (0.244)	12.03*** [168240.4] (0.194)	-3.412*** [0.0330] (0.0674)
Cons	-3.645*** (0.212)	-4.116*** (0.0899)	0.0936 (0.0682)	-4.389*** (0.220)	-4.079*** (0.0897)	-0.465*** (0.0701)	-3.774*** (0.213)	-4.037*** (0.0890)	-0.125 (0.0689)
City F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R2	0.0300	0.114	0.0709	0.0705	0.133	0.0982	0.0385	0.120	0.0891
N	45746	172444	73614	45746	172444	73614	45746	172444	73614

Odds ratios in square brackets - Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.0014$

Appendix D

Geographical agglomeration and firm output

A central concern of Economic Geography has been to understand why related economic activities agglomerate in a small number of locations (Delgado, 2018; Delgado et al., 2015, 2014) and how these agglomerations are important for industrial and firm performance and output.

With the models proposed below, I attempt to test the effects that industrial clustering have on firm output and if the effects hold by analysing the related activities not only within a location, but also within the whole production networks, regardless of their location. In other words, the proposed models explore the effects that relatedness by co-location and relatedness based on the supply chain and clients (which were calculated in Chapter 5) have on the performance of firms, measured by their total output. More specifically, using dynamic panel models, I look at the effects that three relatedness variables (relatedness based on co-location, relatedness based on suppliers, and clients - located within and outside a region) have on the output of individual firms. The preliminary results establish the grounds for continuing this work in the future. It is important to notice that this model needs to be further calibrated and requires several robustness checks.

$$Output_{f(t)} = Output_{f(t-1)} + Output_{f(t-2)} + RD_{i(t-1)}^c + RD_{c(t-1)}^i + DCI_{i(t-1)}^c + DSp_{i(t-1)}^c$$

Table 6: Dynamic Panel Models - Firm Output

VARIABLES	Model 1 <i>Output_{f(t)}</i>	Model 2 <i>Output_{f(t)}</i>	Model 3 <i>Output_{f(t)}</i>	Model 4 <i>Output_{f(t)}</i>	Model 5 <i>Output_{f(t)}</i>	Model 6 <i>Output_{f(t)}</i>	Model 7 <i>Output_{f(t)}</i>
Relatedness Density (RD)	0.003 (0.021)						
Relatedness Suppliers (RSP)		0.054*** (0.02)					
RSP Same City			0.161*** (0.016)				
RSP Different City				0.109*** (0.015)			
Relatedness Clients (DCI)					0.042*** (0.013)		
DCI Same City						0.046*** (0.012)	
DCI Different City							0.02 (0.013)
Firm Output _(t-1)	0.356*** (0.003)	0.356*** (0.003)	0.341*** (0.004)	0.356*** (0.003)	0.356*** (0.003)	0.355*** (0.003)	0.356*** (0.003)
Firm Output _(t-2)	0.035*** (0.001)	0.035*** (0.001)	0.033*** (0.002)	0.035*** (0.001)	0.035*** (0.001)	0.035*** (0.001)	0.035*** (0.001)
City Relatedness (CR)	0.053** (0.021)	0.052** (0.021)	0.077*** (0.025)	0.052** (0.021)	0.052** (0.021)	0.054** (0.021)	0.052** (0.021)
Observations	1,866,164	1,866,164	1,209,183	1,866,164	1,864,874	1,864,384	1,864,867
Number of firms	539,684	539,684	402,386	539,684	539,483	539,370	539,480
p-value AR(2)	0.00104	0.00105	0.134	0.00111	0.00105	0.00102	0.00105
p-value Hansen test	0	0	0	0	0	0	0
Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1							

References

- Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program. *Journal of the American statistical Association*, 105(490):493–505.
- Abadie, A., Diamond, A., and Hainmueller, J. (2011). Synth: An r package for synthetic control methods in comparative case studies. *Journal of Statistical Software*, 42(13).
- Abadie, A., Diamond, A., and Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2):495–510.
- Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. *The American Economic Review*, 93(1):113–132.
- Abdon, A. and Felipe, J. (2011). The product space: What does it say about the opportunities for growth and structural transformation of Sub-Saharan Africa? *Levy Economics Institute Working Papers*, 670.
- Acemoglu, D. (2003). Labor-and capital-augmenting technical change. *Journal of the European Economic Association*, 1(1):1–37.
- Acemoglu, D. (2009). *Introduction to modern economic growth*. Princeton University Press, Princeton ; Oxford.
- Acemoglu, D., Carvalho, V. M., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The Network Origins of Aggregate Fluctuations. *Econometrica*, 80(5):1977–2016.
- Acemoglu, D. and Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6):1488–1542.
- Acosta, A. (2006). *[Brief Economic History of Ecuador]*. Corporación Editora Nacional, Quito, 2nd. edition.
- Acosta, A. (2009a). *Ecuador: un país maniatado frente a la crisis?* Friedrich Ebert Stiftung.

- Acosta, A. (2009b). [*The curse of plenty*]. Abya-Yala, Quito.
- Acosta, A. (2011). [The hydrocarbon law’s reform and renegotiation of oil contracts]. *La Tendencia*, 11:95–103.
- Ahmadov, A. (2014). Blocking the pathway out of the resource curse: What hinders diversification in resource-rich developing countries? *Global Economic Governance (GEG) Working Paper No*, 98.
- Akerlof, G. and Nordhaus, W. D. (1967). Balanced growth-a razor’s edge? *International Economic Review*, 8(3):343–348.
- Albeaik, S., Kaltenberg, M., Alsaleh, M., and Hidalgo, C. A. (2017). Measuring the knowledge intensity of economies with an improved measure of economic complexity. *arXiv preprint arXiv:1707.05826*.
- Alfaro, L., Chor, D., Antras, P., and Conconi, P. (2019). Internalizing global value chains: A firm-level analysis. *Journal of Political Economy*, 127(2):508–559.
- Alshamsi, A., Pinheiro, F. L., and Hidalgo, C. A. (2018). Optimal diversification strategies in the networks of related products and of related research areas. *Nature Communications*, 9(1):1328.
- Anand, A. and Walsh, I. (2016). Should knowledge be shared generously? Tracing insights from past to present and describing a model. *Journal of Knowledge Management*, 20(4):713–730.
- Arkolakis, C., Ramondo, N., Rodríguez-Clare, A., and Yeaple, S. (2018). Innovation and production in the global economy. *American Economic Review*, 108(8):2128–73.
- Armendáriz, B. and Larraín, F. (2017). *The Economics of Contemporary Latin America*. MIT Press.
- Arrow, K. J. (1971). The economic implications of learning by doing. In *Readings in the Theory of Growth*, pages 131–149. Springer.
- Arrow, K. J. and Nelson, R. (1962). The rate and direction of inventive activity. *Economic welfare and the allocation of resources for invention*, 609.
- Asheim, B. T., Boschma, R., and Cooke, P. (2011). Constructing regional advantage: Platform policies based on related variety and differentiated knowledge bases. *Regional studies*, 45(7):893–904.

- Ashraf, N., Calvin, C., Delfino, A., Glaeser, E., and Swanson, N. (2017). Urban density, trust, and knowledge sharing in lusaka, zambia.
- Athey, S. (2019). The impact of machine learning on economics. In Goldfarb, A., Gans, J., and Agrawal, A., editors, *The economics of artificial intelligence: An agenda*, pages 507–547. University of Chicago Press.
- Athey, S. and Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11:685–725.
- Atkeson, A. and Burstein, A. T. (2010). Innovation, firm dynamics, and international trade. *Journal of political economy*, 118(3):433–484.
- Auty, R. (2002). *Sustaining development in mineral economies: the resource curse thesis*. Routledge.
- Auty, R. and Pontara, N. (2008). A dual-track strategy for managing Mauritania’s projected oil rent. *Development Policy Review*, 26(1):59–77.
- Auty, R. M. (1994). Industrial policy reform in six large newly industrializing countries: The resource curse thesis. *World development*, 22(1):11–26.
- Auty, R. M. (2001). *Resource abundance and economic development*. Oxford University Press.
- Auty, R. M. and Furlonge, H. I. (2019). *The Rent Curse: Natural Resources, Policy Choice, and Economic Development*. Oxford University Press, USA.
- Auzer, K. A. (2017). *Institutional Design and Capacity to Enhance Effective Governance of Oil and Gas Wealth: The Case of Kurdistan Region*. Springer.
- Balland, P.-A., Boschma, R., Crespo, J., and Rigby, D. L. (2019). Smart specialization policy in the European Union: Relatedness, knowledge complexity and regional diversification. *Regional Studies*, 53(9):1252–1268.
- Balland, P.-A., Boschma, R., Frenken, K., et al. (2020a). Proximity, innovation and networks: A concise review and some next steps. Technical report, Utrecht University, Department of Human Geography and Spatial Planning
- Balland, P.-A., Jara-Figueroa, C., Petralia, S. G., Steijn, M. P., Rigby, D. L., and Hidalgo, C. A. (2020b). Complex economic activities concentrate in large cities. *Nature Human Behaviour*, pages 1–7.

- Baquero, D. (2017). Oil and economic growth: addiction to shocks? The Ecuadorian case (1972-2014). Unpublished thesis.
- Barabasi, A.-L. (2016). *Network Science*. Cambridge University Press, Cambridge, UK.
- Baronett, S. (2015). *Logic, 3er Edition*. Oxford University Press.
- Barrat, A., Barthélemy, M., Pastor-Satorras, R., and Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11):3747–3752.
- BBC (2016). Falling oil prices: How are countries being affected? <http://www.bbc.co.uk/news/world-35345874>. Accessed on 2016-09-14.
- BCE (1981). *National Accounts 1972-1981*, volume 1. Banco Central del Ecuador.
- BCE (1982). *[National Accounts 1965-1971]*, volume 2. Banco Central del Ecuador.
- BCE (1983). *[National Accounts of Ecuador - Methodology]*, volume 3. Banco Central del Ecuador.
- BCE (1989). *[National Accounts of Ecuador - 1979-1988]*, volume 3. Banco Central del Ecuador.
- BCE (2019). Monthly statistical information no. 2014 - December 2019. Database 1, Banco Central del Ecuador.
- Beaudry, C. and Schiffauerova, A. (2009). Who’s right, Marshall or Jacobs? The localization versus urbanization debate. *Research policy*, 38(2):318–337.
- Beinhocker, E. D. (2006). *The origin of wealth: Evolution, complexity, and the radical remaking of economics*. Harvard Business Press.
- Belli, P. (1984). Ecuador: an agenda for recovery and sustained growth. *Washington DC: World Bank*.
- Benites, L. (2005). *Ecuador: Drama and Paradox*. Casa de la Cultura Ecuatoriana.
- Beyers, W. B. (1976). Empirical identification of key sectors: some further evidence. *Environment and Planning A*, 8(2):231–236.
- Binz, C., Truffer, B., and Coenen, L. (2016). Path creation as a process of resource alignment and anchoring: Industry formation for on-site water recycling in Beijing. *Economic Geography*, 92(2):172–200.

- Blas, J. (2018). The U.S. just became a net oil exporter for the first time in 75 years. *Bloomberg*.
- Bloch, R. and Owusu, G. (2012). Linkages in Ghana’s gold mining industry: Challenging the enclave thesis. *Resources Policy*, 37(4):434–442.
- Blöchl, F., Theis, F. J., Vega-Redondo, F., and Fisher, E. O. (2011). Vertex centralities in input-output networks reveal the structure of modern economies. *Physical Review E*, 83(4):046127.
- Bonilla, H., Chaunu, P., Halperin, T., Vilar, P., Spalding, K., and Hobsbawm, E. J. (1982). *La independencia del Peru*. Instituto de Estudios Peruanos.
- Borgatti, S. P. and Halgin, D. S. (2011). Analyzing affiliation networks. *The Sage handbook of social network analysis*, 1:417–433.
- Bornhorst, F., Gupta, S., and Thornton, J. (2008). Natural resource endowments, governance, and the domestic revenue effort: Evidence from a panel of countries. Report, International Monetary Fund.
- Boschma, R. (2017). Relatedness as driver of regional diversification: A research agenda. *Regional Studies*, 51(3):351–364.
- Boschma, R., Balland, P.-A., and 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.
- Boschma, R., Coenen, L., Frenken, K., and Truffer, B. (2017). Towards a theory of regional diversification: Combining insights from evolutionary economic geography and transition studies. *Regional Studies*, 51(1):31–45.
- Boschma, R. and Frenken, K. (2011). Technological relatedness, related variety and economic geography. In Cooke, P., Asheim, B., Boschma, R., Martin, R., Schwartz, D., and T-dtling, F., editors, *Handbook of regional innovation and growth*, chapter 14, page 187. Edward Elgar Publishing.
- Boschma, R., Minondo, A., and Navarro, M. (2013). The emergence of new industries at the regional level in Spain: A proximity approach based on product relatedness. *Economic geography*, 89(1):29–51.
- Boschma, R. A. (2012). Competitiveness of regions from an evolutionary perspective. In *Regional Competitiveness*, pages 17–30. Routledge.

- Bourguignon, F. (2004). The poverty-growth-inequality triangle. *Poverty, Inequality and Growth*, 69:342674–1206111890151.
- Bouttell, J., Craig, P., Lewsey, J., Robinson, M., and Popham, F. (2018). Synthetic control methodology as a tool for evaluating population-level health interventions. *J Epidemiol Community Health*, 72(8):673–678.
- Brenes, E. R., Camacho, A. R., Ciravegna, L., and Pichardo, C. A. (2016). Strategy and innovation in emerging economies after the end of the commodity boom—insights from Latin America. *Journal of Business Research*, 69(10):4363–4367.
- Britto, G., Romero, J. P., Freitas, E., and Coelho, C. (2019). The great divide: economic complexity and development paths in brazil and the republic of korea. *Cepal Review*.
- Bruton, H. (1989). Import substitution. In Chenery, H. B. and Srinivasan, T., editors, *Handbook of Development Economics*, chapter 30, pages 1602–1641. Elsevier.
- Budina, N., Pang, G., and van Wijnbergen, S. (2006). Nigeria: Dutch disease or debt overhang? Diagnosing the past. *Lessons for the Future. The World Bank, Washington, DC*.
- Bulmer-Thomas, V. (2003). *The economic history of Latin America since independence*. Cambridge University Press.
- Bussell, J. (2018). Import substitution industrialization. <https://www.britannica.com/topic/import-substitution-industrialization>. Accessed on 2020-02-28.
- Bustos, S., Gomez, C., Hausmann, R., and Hidalgo, C. A. (2012). The dynamics of nestedness predicts the evolution of industrial ecosystems. *PloS one*, 7(11):e49393.
- Caldarelli, G. (2007). *Scale-free networks: complex webs in nature and technology*. Oxford University Press.
- Caldarelli, G., Cristelli, M., Gabrielli, A., Pietronero, L., Scala, A., and Tacchella, A. (2012). A network analysis of countries’ export flows: firm grounds for the building blocks of the economy. *PloS one*, 7(10):e47278.
- Caldecott, B., Tilbury, J., and Carey, C. (2014). Stranded assets and scenarios. Technical report, Smith School of Enterprise and the Environment, University of Oxford.
- Canuto, O. (2014). The commodity super cycle: is this time different? *Economic Premise*, 150:1–3.

- Carvalho, V. M. and Voigtländer, N. (2014). Input diffusion and the evolution of production networks. Technical report, National Bureau of Economic Research.
- CD (2020). Natural person. <https://dictionary.cambridge.org/dictionary/english/natural-person>. Accessed on 2020-02-26.
- Chang, H.-J. (2002). *Kicking away the ladder: development strategy in historical perspective*. Anthem Press.
- Chauvin, J. P., Glaeser, E., Ma, Y., and Tobio, K. (2017). What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States. *Journal of Urban Economics*, 98:17–49.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357.
- Chenery, H. (1979). *Structural change and development policy*. Oxford University Press.
- Chenery, H. B. (1975). The structuralist approach to development policy. *The American Economic Review*, 65(2):310–316.
- Chenery, H. B. and Watanabe, T. (1958). International comparisons of the structure of production. *Econometrica: Journal of the Econometric Society*, pages 487–521.
- Chesnais, F. (1993). The French national system of innovation. *National innovation systems: A comparative analysis*, pages 192–229.
- Cline, W. R. (1982). Mexico’s crisis, the world’s peril. *Foreign Policy*, 1(49):107–118.
- CLS (2020). Legal information institute. https://www.law.cornell.edu/wex/natural_person. Accessed on 2020-02-26.
- Cobb, C. W. and Douglas, P. H. (1928). A theory of production. *The American Economic Review*, 18(1):139–165.
- Cockx, L. and Francken, N. (2016). Natural resources: A curse on education spending? *Energy Policy*, 92:394–408.
- Coe, N. M. and Yeung, H. W.-C. (2015). *Global production networks: Theorizing economic development in an interconnected world*. Oxford University Press.
- Collier, P. and Goderis, B. (2009). Structural policies for shock-prone developing countries. *Oxford Economic Papers*, 61(4):703–726.

- Collier, P. and Page, J. (2009). Breaking in and moving up: New industrial challenges for the bottom billion and the middle-income countries. *Industrial Development Report*, UNIDO.
- Comin, D. A., Dmitriev, M., and Rossi-Hansberg, E. (2012). The spatial diffusion of technology. Technical report, National Bureau of Economic Research.
- Corden, W. M. and Neary, J. P. (1982). Booming sector and de-industrialisation in a small open economy. *The economic journal*, 92(368):825–848.
- Correa, R. (2012). *Ecuador: de Banana Republic a la no república*. Debate, Quito.
- Cowen, M. and Shenton, R. (2010). The invention of development. In Gasper, D. and Lera St. Clair, A., editors, *Development Ethics*, chapter 1, pages 3–33. Routledge.
- Cristelli, M., Gabrielli, A., Tacchella, A., Caldarelli, G., and Pietronero, L. (2013). Measuring the intangibles: A metrics for the economic complexity of countries and products. *PloS one*, 8(8):e70726.
- Cristelli, M., Tacchella, A., and Pietronero, L. (2015). The heterogeneous dynamics of economic complexity. *PloS one*, 10(2):e0117174.
- Cust, J. F. and Mihalyi, D. (2017). Evidence for a presource curse? oil discoveries, elevated expectations, and growth disappointments. *Oil Discoveries, Elevated Expectations, and Growth Disappointments (July 10, 2017)*. *World Bank Policy Research Working Paper*, (8140).
- Darwent, D. F. (1969). Growth poles and growth centers in regional planning: a review. *Environment and Planning A*, 1(1):5–31.
- David, P. A. (2001). Path dependence, its critics and the quest for ‘historical economics’. *Evolution and path dependence in economic ideas: Past and present*, 15:40.
- Dées, S., Gasteuil, A., Kaufmann, R., and Mann, M. (2008). Assessing the factors behind oil price changes. *Working Paper Series*.
- Delgado, M. (2018). Firms in context: Internal and external drivers of success. In Clark, G. L., Feldman, M. P., Gertler, M. S., and Wojcik, D., editors, *The New Oxford Handbook of Economic Geography*, chapter 17, page 324. Oxford University Press.
- Delgado, M. (2020). The co-location of innovation and production in clusters. *Industry and Innovation*, pages 1–29.

- Delgado, M., Porter, M. E., and Stern, S. (2010). Clusters and entrepreneurship. *Journal of economic geography*, 10(4):495–518.
- Delgado, M., Porter, M. E., and Stern, S. (2014). Clusters, convergence, and economic performance. *Research policy*, 43(10):1785–1799.
- Delgado, M., Porter, M. E., and Stern, S. (2015). Defining clusters of related industries. *Journal of Economic Geography*, 16(1):1–38.
- Diao, X., McMillan, M., and Rodrik, D. (2017). The recent growth boom in developing economies. *HKS Working Paper*.
- Dimand, R. W. and Spencer, B. J. (2009). Trevor swan and the neoclassical growth model. *History of Political Economy*, 41(Suppl.1):107–126.
- Diodato, D., Neffke, F., and O’Clery, N. (2018). Why do industries coagglomerate? How Marshallian externalities differ by industry and have evolved over time. *Journal of Urban Economics*, 106:1–26.
- Djimeu, E. W. and Omgba, L. D. (2019). Oil windfalls and export diversification in oil-producing countries: Evidence from oil booms. *Energy Economics*, 78:494–507.
- Dobrynskaya, V. and Turkisch, E. (2010). Economic diversification and Dutch disease in Russia. *Post-Communist Economies*, 22(3):283–302.
- Domar, E. D. (1946). Capital expansion, rate of growth, and employment. *Econometrica, Journal of the Econometric Society*, pages 137–147.
- Dutt, A. K. (2019). Structuralists, structures, and economic development. In Nissanke, M. and Ocampo, J. A., editors, *The Palgrave Handbook of Development Economics*, pages 109–141. Springer.
- ECLAC (1992). *Social equity and changing production patterns: an interpretation of the integrated approach*. Economic Commission for Latin America and the Caribbean.
- Erten, B. and Ocampo, J. A. (2013). Super cycles of commodity prices since the mid-nineteenth century. *World Development*, 44:14–30.
- Espinoza, V. S., Fontalvo, J., Martí-Herrero, J., Ramírez, P., and Capellán-Pérez, I. (2019). Future oil extraction in Ecuador using a Hubbert approach. *Energy*, 182:520–534.
- EuroStat (2016). Statistics explained: Intermediate consumption. http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Intermediate_consumption. Accessed on 2017-04-27.

- EUSTAT (2020). Leontief inverse matrix. <https://en.eustat.eus/documentos/elem15552/definicion.html>. Accessed on 2020-03-02.
- Feldman, M. P. and Audretsch, D. B. (1999). Innovation in cities: Science-based diversity, specialization and localized competition. *European economic review*, 43(2):409–429.
- Felipe, J. (2012). *Inclusive Growth, Full Employment, and Structural Change: Implications and Policies for Developing Asia*. Anthem Press.
- Feser, E., Renski, H., and Goldstein, H. (2008). Clusters and economic development outcomes: An analysis of the link between clustering and industry growth. *Economic Development Quarterly*, 22(4):324–344.
- FitzGerald, V. (1998). La cepal y la teoría de la industrialización. *Revista de la CEPAL*.
- Galor, O., Aghion, P., and Durlauf, S. N. (2005). From stagnation to growth: Unified growth theory. In Aghion, P. and Durlauf, S., editors, *Handbook of Economic Growth, Vol. 1*, page 171–293. Elsevier.
- Galor, O. and Weil, D. N. (2000). Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond. *American economic review*, 90(4):806–828.
- Gao, J., Jun, B., Pentland, A., Zhou, T., Hidalgo, C. A., et al. (2017). Collective learning in China’s regional economic development. *arXiv preprint arXiv:1703.01369*.
- García, J. M. R. (2001). Scientia potestas est—knowledge is power: Francis Bacon to Michel Foucault. *Anglia-Zeitschrift für Englische Philologie*, 119(1):1–19.
- Garud, R. and Karnøe, P. (2003). Bricolage versus breakthrough: distributed and embedded agency in technology entrepreneurship. *Research policy*, 32(2):277–300.
- Garud, R., Kumaraswamy, A., and Karnøe, P. (2010). Path dependence or path creation? *Journal of Management Studies*, 47(4):760–774.
- Gelb, A. (2011). Economic diversification in resource rich countries. In Sy, M. A. N., Arezki, M. R., and Gylfason, T., editors, *Beyond the curse: policies to harness the power of natural resources*, chapter 4, pages 55–80. International Monetary Fund.
- Gelb, A. and Grasmann, S. (2010). How should oil exporters spend their rents? *Working paper*.
- Glaeser, E. (2011a). Cities, productivity, and quality of life. *Science*, 333(6042):592–594.
- Glaeser, E. (2011b). *Triumph of the city: How urban spaces make us human*. Pan Macmillan.

- Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., and Shleifer, A. (1992). Growth in cities. *Journal of political economy*, 100(6):1126–1152.
- Glaeser, E. L. and Kerr, W. R. (2009). Local industrial conditions and entrepreneurship: how much of the spatial distribution can we explain? *Journal of Economics & Management Strategy*, 18(3):623–663.
- Gomez-Lievano, A., Patterson-Lomba, O., and Hausmann, R. (2018). Explaining the prevalence, scaling and variance of urban phenomena. *Nature Energy*, pages 1–9.
- González, J. and Cabrera, D. (2018). Análisis del sector informal. *Observatorio de la Economía Latinoamericana*, March.
- Gruss, B. (2014). *After the boom—commodity prices and economic growth in Latin America and the Caribbean*. Number 14-154 in IMF Working Paper. International Monetary Fund.
- Gylfason, T. (2001). Natural resources, education, and economic development. *European economic review*, 45(4-6):847–859.
- Haldane, A. G. and May, R. M. (2011). Systemic risk in banking ecosystems. *Nature*, 469(7330):351–355.
- Hansen, G. D. and Prescott, E. C. (2002). Malthus to solow. *American Economic Review*, 92(4):1205–1217.
- Haraguchi, N., Cheng, C. F. C., and Smeets, E. (2017). The importance of manufacturing in economic development: Has this changed? *World Development*, 93:293–315.
- Harrod, R. F. (1939). An essay in dynamic theory. *The economic journal*, 49(193):14–33.
- Hartmann, D. (2014). *Economic complexity and human development: how economic diversification and social networks affect human agency and welfare*, volume 110. Routledge.
- Hartmann, D., Guevara, M. R., Jara-Figueroa, C., Aristarán, M., and Hidalgo, C. A. (2017a). Linking economic complexity, institutions, and income inequality. *World Development*, 93:75–93.
- Hartmann, D., Jara-Figueroa, C., Guevara, M., Simoes, A., and Hidalgo, C. A. (2017b). The structural constraints of income inequality in Latin America. *arXiv preprint arXiv:1701.03770*.
- Harvey, D. I., Kellard, N. M., Madsen, J. B., and Wohar, M. E. (2010a). The Prebisch-Singer hypothesis: four centuries of evidence. *The review of Economics and Statistics*, 92(2):367–377.

- Harvey, D. I., Kellard, N. M., Madsen, J. B., and Wohar, M. E. (2010b). The prebisch-singer hypothesis: four centuries of evidence. *The review of Economics and Statistics*, 92(2):367–377.
- Hassink, R., Isaksen, A., and Trippel, M. (2019). Towards a comprehensive understanding of new regional industrial path development. *Regional Studies*, 53(11):1636–1645.
- Hausmann, R. and Hidalgo, C. A. (2011). The network structure of economic output. *Journal of Economic Growth*, 16(4):309–342.
- Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., Simoes, A., and Yildirim, M. A. (2014). *The atlas of economic complexity: Mapping paths to prosperity*. MIT Press.
- Hausmann, R., Hwang, J., and Rodrik, D. (2007a). What you export matters. *Journal of economic growth*, 12(1):1–25.
- Hausmann, R., Klinger, B., et al. (2007b). The structure of the product space and the evolution of comparative advantage. Technical report, Center for International Development at Harvard University.
- Hausmann, R., Klinger, B., and Lopez-Calix, J. (2010). Export diversification in Algeria. In Lopez-Calix, J. R., Walkenhorst, P., and Diop, N., editors, *Trade Competitiveness of the Middle East and North Africa*, chapter 4, pages 63–103. World Bank.
- Hausmann, R. and Rigobon, R. (2003). An alternative interpretation of the ‘resource curse’: Theory and policy implications. Technical report, National Bureau of Economic Research.
- Hausmann, R., Stock, D. P., and Yıldırım, M. A. (2021). Implied comparative advantage. *Research Policy*, page 104143.
- Helm, D. (2016). The future of fossil fuels: is it the end? *Oxford Review of Economic Policy*, 32(2):191–205.
- Hémous, D. and Olsen, M. (2014). The rise of the machines: Automation, horizontal innovation and income inequality. *CEPR Discussion Paper No. DP10244*.
- Herfindahl, O. C. (1950). *Concentration in the steel industry*. PhD thesis, Columbia University.
- Herrera, J. C. and Jarrin, S. (2013). *[Generation and distribution of oil revenues in 2012]*, volume 5 of *Lupa Fiscal*. Group Faro, Quito.
- Hesse, H. (2009). Export diversification and economic growth. *Breaking into new markets: emerging lessons for export diversification*, pages 55–80.

- Hidalgo, C. (2015). *Why information grows: The evolution of order, from atoms to economies*. Basic Books.
- Hidalgo, C. A. (2021). Economic complexity theory and applications. *Nature Reviews Physics*, pages 1–22.
- Hidalgo, C. A., Balland, P.-A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., Glaeser, E., He, C., Kogler, D. F., Morrison, A., et al. (2018). The principle of relatedness. In *International Conference on Complex Systems*, pages 451–457. Springer.
- Hidalgo, C. A. and Hausmann, R. (2009). The building blocks of economic complexity. *proceedings of the national academy of sciences*, 106(26):10570–10575.
- Hidalgo, C. A., Klinger, B., Barabási, A.-L., and Hausmann, R. (2007). The product space conditions the development of nations. *Science*, 317(5837):482–487.
- Hirschman, A. O. (1958). *The Strategy of Economic Development*. Yale University Press.
- Hirschman, A. O. (1980). *National power and the structure of foreign trade*, volume 105. Univ of California Press.
- Humphreys, M., Sachs, J., and Stiglitz, J. E. (2007). *Escaping the resource curse*. Columbia University Press.
- Hurtado, O. (2018). *Las costumbres de los ecuatorianos*. DEBATE.
- IEA (2012). *World Energy Outlook 2012*. OECD Publishing.
- Imbs, J. and Wacziarg, R. (2003). Stages of diversification. *The American Economic Review*, 93(1):63–86.
- Isaksen, A. (2014). Industrial development in thin regions: trapped in path extension? *Journal of economic geography*, 15(3):585–600.
- Jackson, M. O. (2010). *Social and economic networks*. Princeton university press.
- Jackson, M. O. (2014). Networks in the understanding of economic behaviors. *Journal of Economic Perspectives*, 28(4):3–22.
- Jara-Figueroa, C., Jun, B., Glaeser, E. L., and Hidalgo, C. A. (2018). The role of industry-specific, occupation-specific, and location-specific knowledge in the growth and survival of new firms. *Proceedings of the National Academy of Sciences*, 115(50):12646–12653.

- Jones, A. and Rice, N. (2011). Econometric evaluation of health policies. In Glied, S. and Smith, P. C., editors, *The Oxford handbook of health economics*. Oxford University Press, Oxford.
- Jones, C. I. (1999). Was an industrial revolution inevitable? Economic growth over the very long run. Technical report, National Bureau of Economic Research.
- Jones, C. I. (2019). Paul Romer: Ideas, nonrivalry, and endogenous growth. *The Scandinavian Journal of Economics*, 121(3):859–883.
- Jones, L. P. (1976). The measurement of hirschmanian linkages. *The Quarterly Journal of Economics*, 90(2):323–333.
- Judge, G. (2019). Combining the information from econometrics learning (el) and machine learning (ml). *Available at SSRN 3376953*.
- Jun, B., Alshamsi, A., Gao, J., and Hidalgo, C. A. (2019). Bilateral relatedness: knowledge diffusion and the evolution of bilateral trade. *Journal of Evolutionary Economics*, pages 1–31.
- Kaur, M. and Singh, L. (2016). Knowledge in the economic growth of developing economies. *African Journal of Science, Technology, Innovation and Development*, 8(2):205–212.
- Keister, L. A. (2003). Religion and wealth: The role of religious affiliation and participation in early adult asset accumulation. *Social Forces*, 82(1):175–207.
- Keller, W. and Yeaple, S. R. (2013). The gravity of knowledge. *American Economic Review*, 103(4):1414–44.
- Keynes, J. M. (2018). *The general theory of employment, interest, and money*. Springer.
- Kirman, A. (2016). Networks: A paradigm shift for economics? In Yan, B., Galeotti, A., and Rogers, B., editors, *The Oxford Handbook of the Economics of Networks*. Oxford University Press.
- Kogler, D. F., Rigby, D. L., and Tucker, I. (2013). Mapping knowledge space and technological relatedness in US cities. *European Planning Studies*, 21(9):1374–1391.
- Kubo, Y., De Melo, J., Robinson, S., and Syrquin, M. (1986). Interdependence and industrial structure. In Chenery, HB, S. R. and Syrquin, M., editors, *Industrialization and Growth: A Comparative Study*. World Bank.

- Kugler, M. (2018). The economics of ideas: Paul Romer, former Berkeley Economics Professor, receives the 2018 Nobel Prize. <http://https://www.econ.berkeley.edu/content/guest-post-economist-ideas-paul-romer-former-berkeley-economics-professor-receives-2018>. Accessed on 2020-04-10.
- Lange, M., Mahoney, J., and Vom Hau, M. (2006). Colonialism and development: A comparative analysis of Spanish and British colonies. *American Journal of Sociology*, 111(5):1412–1462.
- Larrain, J. (2013). *Theories of development: Capitalism, colonialism and dependency*. John Wiley & Sons.
- Larrea, C. (1992). The mirage of development: oil, employment, and poverty in Ecuador. *PhD diss., Graduate Programme in Social and Political Thought, York University, Ontario*.
- Larrea, C. (2004). Dolarización y desarrollo humano en Ecuador. *Iconos. Revista de Ciencias Sociales*, 19.
- Larrea, C. (2013). Extractivism, economic diversification and prospects for sustainable development in Ecuador. *Repositorio Digital UASB*.
- Lederman, D. and Maloney, W. (2012). *Does what You Export Matter?: In Search of Empirical Guidance for Industrial Policies*. World Bank Publications.
- Lederman, D. and Maloney, W. F. (2006). *Natural resources, neither curse nor destiny*. World Bank Publications.
- Lederman, D. and Maloney, W. F. (2007). Neither curse nor destiny: Introduction to natural resources and development. *Natural Resources*, 1:1–396.
- Leontief, W. W. (1986). *Input-output economics*. Oxford University Press on Demand.
- Levy, A. (1954). Economic views of Thomas Hobbes. *Journal of the History of Ideas*, pages 589–595.
- Lin, J. and Chang, H.-J. (2009). Should industrial policy in developing countries conform to comparative advantage or defy it? A debate between Justin Lin and Ha-Joon Chang. *Development policy review*, 27(5):483–502.
- Lin, J. Y. (2010). *New structural economics: A framework for rethinking development*. World Bank, Washington, DC.

- Lin, J. Y. (2017). Industrial policies for avoiding the middle-income trap: A new structural economics perspective. *Journal of Chinese Economic and Business Studies*, 15(1):5–18.
- Lin, J. Y. (2019). New structural economics: the third generation of development economics. *Asian Education and Development Studies*.
- Lin, J. Y. and Wang, Y. (2020). Seventy years of economic development: A review from the angle of new structural economics. *China & World Economy*, 28(4):26–50.
- Lipscomb, L., Ramakrishnan, U., Adler, G., Budina, N., and Li, X. (2017). Indonesia: Selected issues. Technical report, International Monetary Fund.
- Locke, J. (2016). *Second treatise of government and a letter concerning toleration*. Oxford University Press.
- Love, J. L. (2005). The rise and decline of economic structuralism in Latin America: New dimensions. *Latin American Research Review*, 40(3):100–125.
- Lucas, R. E. (2002). *Lectures on economic growth*. Harvard University Press.
- Lucas Jr, R. E. (1988). On the mechanics of economic development. *Journal of monetary economics*, 22(1):3–42.
- Lustig, N., Lopez-Calva, L. F., Ortiz-Juárez, E., and Monga, C. (2016). Deconstructing the decline in inequality in latin america. In *Inequality and growth: Patterns and policy*, pages 212–247. Springer.
- Marini, R. M. (2008). *América Latina, dependencia y globalización*. CLACSO.
- Marshall, A. (1920). *Principles of Economics, 8th Edition*. Palgrave Macmillan, NY.
- Martin, R. (2010). Roepke lecture in economic geography: Rethinking regional path dependence: beyond lock-in to evolution. *Economic geography*, 86(1):1–27.
- Martin, R. (2012). (re) placing path dependence: a response to the debate. *International Journal of Urban and Regional Research*, 36(1):179–192.
- Martin, R. and Sunley, P. (2006). Path dependence and regional economic evolution. *Journal of economic geography*, 6(4):395–437.
- Marx, K. (2019). *Capital: volume one*. Courier Dover Publications.
- Mateo, J. P. and Garcia, S. (2014). El sector petrolero en Ecuador 2000-2010. *Revista Problemas del Desarrollo*, 117(45):113–139.

- McMillan, M., Rodrik, D., and Sepulveda, C. (2017). Structural change, fundamentals, and growth: a framework and case studies. *NBER Working Papers Series*, 23378.
- Mealy, P. and Teytelboym, A. (2020). Economic complexity and the green economy. *Research Policy*, page 103948.
- Meller, P. and Marfan, M. (1981). Small and large industry: employment generation, linkages, and key sectors. *Economic Development and Cultural Change*, 29(2):263–274.
- Milanovic, B. (2013). Global income inequality in numbers: In history and now. *Global policy*, 4(2):198–208.
- Miller, R. E. and Blair, P. D. (2009). *Input-output analysis: foundations and extensions*. Cambridge University Press.
- Miller, R. G. and Sorrell, S. R. (2014). The future of oil supply. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372.
- Mirowski, P. (1991). *More heat than light: Economics as social physics, physics as nature’s economics*. Cambridge University Press.
- Missio, F., OREIRO, J. L., et al. (2015). The structuralist tradition in economics: methodological and macroeconomics aspects. *Brazilian Journal of Political Economy*, 35(2):247–266.
- Moggridge, D. E. (1980). *Keynes*. Springer.
- MRNNR (2013). Ronda sur-oriente - Ministerio de Recursos Naturales No Renovables. <http://www.rondasurorientegob.ec/wp-content/uploads/2013/08/RONDA-SURORIENTE-ECUADOR-ENGLISH-revision-enero-2013.pdf>. Accessed on 2014-07-31.
- Mullainathan, S. and Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2):87–106.
- Muneepeerakul, R., Lobo, J., Shuttters, S. T., Gómez-Liévano, A., and Qubbaj, M. R. (2013). Urban economies and occupation space: can they get “there” from “here”? *PloS one*, 8(9):e73676.
- Murshed, S. M. and Serino, L. A. (2011). The pattern of specialization and economic growth: The resource curse hypothesis revisited. *Structural Change and Economic Dynamics*, 22(2):151–161.

- Murtaugh, D., Longley, A., and Davalos, J. (2020). Oil plunges most since 1991 after producers embark on price war. <https://www.bloomberg.com/news/articles/2020-03-08/oil-in-freefall-after-saudis-slash-prices-in-all-out-crude-war>. Accessed on 2020-03-10.
- Neffke, F., Henning, M., and 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.
- NetworkX (2015a). Collaboration weighted projected graph.
- NetworkX (2015b). Overlap weighted projected graph.
- Newman, M. (2010). *Networks: an introduction*. Oxford university press.
- Newman, M. E. (2001). Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Physical review E*, 64(1):016132.
- Newman, M. E. (2004). Analysis of weighted networks. *Physical review E*, 70(5):056131.
- Noguera-Méndez, P., Semitiel-García, M., and López-Martínez, M. (2016). Estructura interindustrial y desarrollo económico. un análisis desde las perspectivas de redes e input-output. *El Trimestre Económico*, 83(331).
- Nordhaus, W. D. (1969). An economic theory of technological change. *The American Economic Review*, 59(2):18–28.
- North, D. C. (1995). The new institutional economics and third world development. In *The new institutional economics and third world development*, pages 31–40. Routledge.
- O’Boyle, L. (1983). Learning for its own sake: The German university as nineteenth-century model. *Comparative Studies in Society and History*, 25(1):3–25.
- Ocampo, J. A. (2014). Latin american structuralism and production development strategies. *Transforming Economies. Making Industrial Policy Work for Growth, Jobs and Development*. Geneva: ILO, pages 41–64.
- Ocampo, J. A., Rada, C., and Taylor, L. (2009). *Growth and policy in developing countries: a structuralist approach*. Columbia University Press.
- Opsahl, T., Agneessens, F., and Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social networks*, 32(3):245–251.
- Orozco, M. (2020). Recorte de gasto público y deuda plantea gobierno para afrontar escenario económico. <https://www.elcomercio.com/actualidad/recorte-gasto-publico-deuda-plantea.html>. Accessed on 2020-03-10.

- Osborne, J. W. (2014). *Best practices in logistic regression*. Sage Publications.
- Page, S. E. et al. (2006). Path dependence. *Quarterly Journal of Political Science*, 1(1):87–115.
- Papyrakis, E. (2017). The Resource Curse - What Have We Learned from Two Decades of Intensive Research: Introduction to the Special Issue. *The Journal of Development Studies*, 53(2):175–185.
- Papyrakis, E. and Pellegrini, L. (2019). The Resource Curse in Latin America. In Thompson, W. R., editor, *Oxford Research Encyclopedia of Politics*. Oxford University Press.
- Peet, R. and Hartwick, E. (2015). *Theories of development: Contentions, arguments, alternatives*. Guilford Publications.
- Perez, M. G. (2019). Ecuador fights banana fungus threatening America’s favorite fruit. <https://www.bloomberg.com/news/articles/2019-08-21/top-banana-shipper-fights-fungus-threatening-favorite-u-s-fruit>. Accessed on 2020-01-11.
- Pinheiro, F. L., Alshamsi, A., Hartmann, D., Boschma, R., and Hidalgo, C. A. (2018). Shooting high or low: Do countries benefit from entering unrelated activities? *arXiv preprint arXiv:1801.05352*.
- Porter, M. (2003). The economic performance of regions. *Regional studies*, 37(6-7):549–578.
- Porter, M. E. (2008). *On competition*. Harvard Business Press.
- Prebisch, R. (1950). The economic development of Latin America and its principal problems. Technical report, Naciones Unidas Comisión Económica para América Latina y el Caribe (CEPAL).
- Ramondo, N., Rodríguez-Clare, A., and Saborío-Rodríguez, M. (2016). Trade, domestic frictions, and scale effects. *American Economic Review*, 106(10):3159–84.
- Rasmussen, P. N. (1956). *Studies in inter-sectoral relations*, volume 15. E. Hark.
- Ravallion, M. (2004). Pro-poor growth: a primer. *World Bank Policy Research Working Paper*, 3242.
- Ricardo, D. (1891). *Principles of political economy and taxation*. G. Bell.
- Rodrik, D. (2006). What’s so special about China’s exports? *China & World Economy*, 14(5):1–19.

- Romer, D. (2019). *Advanced macroeconomics*. McGraw-Hill series in economics. McGraw-Hill, New York, fifth edition. edition.
- Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of political economy*, 94(5):1002–1037.
- Romer, P. M. (1990). Endogenous technological change. *Journal of political Economy*, 98(5, Part 2):S71–S102.
- Romer, P. M. (1994). The origins of endogenous growth. *Journal of Economic perspectives*, 8(1):3–22.
- Ross, M. L. (2013). *The oil curse: How petroleum wealth shapes the development of nations*. Princeton University Press.
- Ross, M. L. (2018). The politics of the resource curse. *The Oxford Handbook of the Politics of Development*, page 200.
- Ross, M. L. (2019). What do we know about export diversification in oil-producing countries? *The Extractive Industries and Society*, 6(3):792–806.
- Sachs, J. D. and Warner, A. M. (1995). Natural resource abundance and economic growth. Technical report, National Bureau of Economic Research.
- Sachs, J. D. and Warner, A. M. (2001). The curse of natural resources. *European economic review*, 45(4):827–838.
- Sachs, J. D. and Williamson, J. (1985). External debt and macroeconomic performance in latin america and east asia. *Brookings Papers on Economic Activity*, 1985(2):523–573.
- Schultz, S. (1977). Approaches to identifying key sectors empirically by means of input-output analysis. *The Journal of Development Studies*, 14(1):77–96.
- Schumpeter, J. A. (2013). *Capitalism, socialism and democracy*. Routledge.
- Sen, A. (1990). Development as capability expansion. *The community development reader*, pages 41–58.
- SENPLADES (2013). Transformación de la matriz productiva. https://www.planificacion.gob.ec/wp-content/uploads/downloads/2013/01/matriz_productiva_WEBtodo.pdf. Accessed on 2020-03-10.
- Singer, H. (1950a). Comments to the terms of trade and economic development. *Review of Economics and Statistics*, 40:84–89.

- Singer, H. W. (1950b). The distribution of gains between investing and borrowing countries. *The American Economic Review*, 40(2):473–485.
- Smith, A. (2010). *The Wealth of Nations: An inquiry into the nature and causes of the Wealth of Nations*. Harriman House Limited.
- Solow, R. M. (1956). A contribution to the theory of economic growth. *The quarterly journal of economics*, 70(1):65–94.
- Solow, R. M. (1957). Technical change and the aggregate production function. *The review of Economics and Statistics*, pages 312–320.
- Solow, R. M. (2007). The last 50 years in growth theory and the next 10. *Oxford review of economic policy*, 23(1):3–14.
- Stambuli, P. K. (1998). Causes and consequences of the 1982 world debt crisis. *Predoctoral Research Paper Department of Economics University of Surrey Guildford*, pages 1–35.
- Staudt, K. (1991). *Managing development*, volume 178. Sage.
- Stevens, P., Lahn, G., and Kooroshy, J. (2015). The resource curse revisited. *Chatham House Research Paper, London*.
- Swan, T. W. (1956). Economic growth and capital accumulation. *Economic record*, 32(2):334–361.
- Sy, M. A. N., Arezki, M. R., and Gylfason, T. (2011). *Beyond the curse: policies to harness the power of natural resources*. International Monetary Fund.
- Szirmai, A. (2013). Manufacturing and economic development. In Szirmai, A., Naudé, W., and Alcorta, L., editors, *Pathways to Industrialization in the Twenty-First Century: New Challenges and Emerging Paradigms*, chapter 2, page 53. Oxford University Press.
- Tabata, S. (2012). Observations on Russian exposure to the Dutch disease. *Eurasian Geography and Economics*, 53(2):231–243.
- Tily, G. (2009). John Maynard Keynes and the development of national accounts in Britain, 1895–1941. *Review of income and wealth*, 55(2):331–359.
- Tripl, M., Grillitsch, M., and Isaksen, A. (2018). Exogenous sources of regional industrial change: Attraction and absorption of non-local knowledge for new path development. *Progress in Human Geography*, 42(5):687–705.

- UCLA (2007). Ap statistics curriculum 2007 pareto. [http](http://wiki.stat.ucla.edu/socr/index.php/APstatisticsCurriculum2007pareto) : [//wiki.stat.ucla.edu/socr/index.php/APstatisticsCurriculum2007pareto](http://wiki.stat.ucla.edu/socr/index.php/APstatisticsCurriculum2007pareto). Accessed on 2020-02-24.
- UNSTAT (2018). Supply and use tables. [https](https://unstats.un.org/Unsd/nationalaccount/workshops/2018/beijing/S1sUTIntroductionUNSD.pdf) : [//unstats.un.org/Unsd/nationalaccount/workshops/2018/beijing/S1sUTIntroductionUNSD.pdf](https://unstats.un.org/Unsd/nationalaccount/workshops/2018/beijing/S1sUTIntroductionUNSD.pdf). Accessed on 2020-02-27.
- Utkovski, Z., Pradier, M. F., Stojkoski, V., Perez-Cruz, F., and Kocarev, L. (2018). Economic complexity unfolded: Interpretable model for the productive structure of economies. *PloS one*, 13(8).
- Vicepresidency of Ecuador (2014). Estrategia nacional para el cambio de la matriz productiva. <https://www.vicepresidencia.gob.ec/wp-content/uploads/2014/04/Estrategia.pdf>. Accessed on 2020-03-10.
- Vos, R. (1987). *Industrialización, empleo y necesidades básicas en el Ecuador*, volume 14. Corporación Editora Nacional, Quito.
- Wadho, W. A. (2014). Education, rent seeking and the curse of natural resources. *Economics & Politics*, 26(1):128–156.
- Warsh, D. (2007). *Knowledge and the wealth of nations: A story of economic discovery*. WW Norton & Company.
- Whittle, A. and Kogler, D. F. (2019). Related to what? Reviewing the literature on technological relatedness: Where we are now and where can we go? *Papers in Regional Science*.
- Williams, M. (2014). *The end of the developmental state?* Routledge.
- World Bank (2020). World bank country and lending groups. [https](https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups) : [//datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups](https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups). Accessed on 2020-04-05.
- Zambrano, E. (2016). External shocks on a dollarized economy: an empirical evidence from Ecuador. *Cuestiones Económicas*, 26(22):111–142.
- Zhou, T., Ren, J., Medo, M., and Zhang, Y.-C. (2007). Bipartite network projection and personal recommendation. *Physical Review E*, 76(4):046115.
- Zhu, S., He, C., and 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.