Technological Change and Uneven Economic Development

Sergio Gabriel Petralia

ISBN 978-90-6266-477-1

Cover design by Estefanía Gilges

TECHNOLOGICAL CHANGE AND UNEVEN ECONOMIC DEVELOPMENT

TECHNOLOGISCHE VERANDERING EN ONGELIJKE ECONOMISCHE ONTWIKKELING

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op maandag 20 oktober 2017 des ochtends te 10.30 uur

door

Sergio Gabriel Petralia

geboren op 29 oktober 1981 te Buenos Aires, Argentinië

Promotor: Prof.dr. R. Boschma

Copromotoren: Dr. P.A. Balland

Dr. A. Morrison

This thesis was partly accomplished with financial support from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).

Committee: Prof.dr. M. P. Feldman

Prof.dr. K. Frenken

Dr. C. A. Hidalgo

Dr. A. Nuvolari

Prof.dr. F. Lissoni

Prof.dr. F. Van Oort

Contents

1 Introduction **13**

2 Climbing the Ladder of Technological Development **27**

2.1 Introduction 29

2.2 Theoretical Background 32

2.2.1 On technological diversification and development 32

2.2.2 Technological Diversification, Relatedness, and Stages of Development 35

2.3 Data and Methods 39

2.3.1 Data Sources and Variables 39

2.3.2 Methodology 47

2.4 Results 50

2.5 Concluding Remarks 56
2.6 Appendix 60
2.6.1 Test for Sample Selection 60
2.6.2 Robustness of the Results: Dividing Countries into Developed and Developing Economies using World Bank's Classification . 64
2.6.3 Technologies Life Cycles and Barriers to Entry 68

3 Unveiling the geography of historical patents in the United States from 1836 to 1975 **71**
3.1 Background & Summary 73
3.2 Methods 77
3.2.1 Step 1: Finding candidate locations within patent documents 78
3.2.2 Step 2: Filtering correct locations 82
3.2.3 Step 3: Including location for 'unreadable' patent documents . 88
3.3 Code Availability 93
3.4 Data Records 93
3.5 Technical Validation 95
3.5.1 Predicting model performance 96
3.5.2 Geographical Distribution of the Final Sample 99
3.6 Usage Notes 105

4 From Ideas to Trade **107**
4.1 Introduction 109
4.2 Theoretical Framework 115
4.2.1 Model Setup 115
4.2.2 Technology: the role of T and θ 116
4.2.3 A small model with country-specific comparative advantage . . 119
4.2.4 Full model with country-specific comparative advantage 121
4.3 Data Description 125

4.3.1 Patent Data 126
4.4 Estimation Results 133
4.5 Concluding Remarks 138
4.6 Appendix 140
4.6.1 Approximation 140
4.6.2 Data 142
4.6.3 Additional Figures 143
4.6.4 Regression Results (Full Table) 143

5 Unravelling the Trail of a GPT: The Case of Electrical & Electronic Technologies from 1860 to 1930 145

5.1 Introduction 147
5.2 Data Sources 150
5.3 Toward a characterization of GPTs 157
5.3.1 Wide Scope for Improvement and Elaboration 160
5.3.2 Potential for Use in a Wide Variety of Products and Processes . 162
5.3.3 Strong Complementaries with Existing and New Technologies 166
5.4 A GPT at Work 169
5.5 Conclusions 182
5.6 Appendix 184
5.6.1 List of Key-Words Used in Section 5.3.2 184
5.6.2 Regression Table with Year Interactions 185
5.6.3 Controlling for Initial Characteristics of Places 188
5.6.4 Robustness to Alternative Estimation Methods for Dynamic Panels 192
5.6.5 Including Additional Endogenous Variables 197

6 Conclusion 201

6.1 Main Findings 203

6.2 Implications for Future Research 208
6.3 Policy Implications 211

Bibliography **217**

Nederlandse Samenvatting **234**

Acknowledgements **240**

Curriculum Vitae **246**

CHAPTER 1

INTRODUCTION

There is a growing global concern over the current disparities in wealth and income, as well as an urgent desire to understand why and how they came to be. In 2017, the World Economic Forum (WEF) ranked the rise in income disparities and the polarization of societies as the primary risk to the global economy. According to UNESCO, the richest 1% of the world's population now own as much as the rest of the world combined.

These disparities are, to a large extent, determined by the capacity of individuals to collectively accumulate knowledge and know-how. Even though technological innovation is the main driving force behind economic growth, industrial development, and the rise of living standards, only a handful of countries are actively developing new technologies. The United States (US), Western European countries, Japan and South Korea host a small fraction of the world's population but are responsible for most technological advances. This unequal distribution of innovative activities establishes the roles played by different countries in the global value chain. Countries that innovate are able to capture a larger share of the value added, while others are trapped in less-profitable activities.

This uneven distribution of technological capabilities is not a mere reflection of cross-country differences, as it also emerges when comparing innovative activities across regions within countries or across firms within sectors. For instance, in the US, the state of California received an average of more than 1 patent per thousand inhabitants in 2015, while the next most productive state (New York) received less than half of that. The state of Maine received barely more than 200 patents in 2015, comparable in absolute terms to all patents granted in the US to Mexican residents during that year. This means that, on average, the per-capita technological production in California was 6 times higher than in Maine and 205 times higher than in Mexico.

The innate disruptive nature of technological progress surfaced abruptly with recent technological trends and materialized in unprecedented income disparities, even within places. David Rotman notes that while the median income in Sillicon Valley (the epicenter of the latest technological revolution) reached $102,000 in 2015, considerably above the national median ($55,000), an estimated 30% of jobs in the region pay $16 an hour or less, with poverty rates of approximately 19% in Santa Clara County[1]. According to Piketty (2014), labor income inequality is “probably higher than in any other society at any time in the past, anywhere in the world”.

Be that as it may, to what extent can we claim that this is something new, something particular to recent technological trends? After all, technologies have had disruptive effects before. For instance, it is argued that the diffusion of the steam engine affected the entire spatial distribution of the US population, by allowing firms and people to relocate away from rivers (the primary source of mechanical power in the 19th century) to meet markets and benefit from agglomera-

[1]See MIT Technology Review, November 2014. I updated values whenever possible.

tion externalities (Rosenberg and Trajtenberg, 2004). Then, the disruptive effects of technological change may be common to some types of technologies and have effects that go beyond generating disparities in income, even transforming the demographics of places and, perhaps, other aspects of the economy. If these are the symptoms, what are the channels through which technological change operates? Moreover, is there a way to avoid these seemingly inevitable and undesired effects?

Even though there is an extensive literature analyzing the process of the accumulation of technological capabilities (see, among others, Bell and Pavitt (1992), Enos (1991), Lall (1992), Dahlman et al. (1987), King and Fransman (1984), Lee and Lim (2001), Kim (1999), Dosi et al. (1988), Malerba (1992), Patel and Pavitt (1997)), we still have a limited understanding of how countries accumulate new technological capabilities along different stages of their economic development. In fact, cross-country quantitative studies exploring patterns of technological diversification and specialization have been very limited and often restricted to the analysis of a handful of developed economies (see, for instance, Archibugi and Pianta (1994) and Cantwell and Vertova (2004)). As a result, we lack a robust and comprehensive body of evidence providing a general characterization of the types of technologies countries are more likely to produce, whether they tend to follow coherent patterns of technological specialization as they develop, and to what extent technological change is bounded to pre-existing technological capabilities.

Along with this lack of a comprehensive characterization of the general patterns of technological accumulation, theoretical models have also struggled to meaningfully incorporate the role of technological accumulation. This imposes an immediate limitation on the possibility of disentangling the mechanisms and channels that lead to uneven and disruptive economic consequences. For instance, general equilibrium trade models, such as Eaton and Kortum (2002), often relegate the role of technology to a parameter that is later collapsed and netted out during estimation or derived as a residual of the model. Incorporating a more prominent role for technological progress in these models can help to understand the disparate way in which countries integrate into external markets and global value chains.

To a similar extent, empirical studies assessing the impact of key technologies have been limited by the availability of data and the nature of the object of study. After all, the diffusion of technologies takes decades, and the most interesting cases happened long before data began to be collected systematically. For instance, the Corliss steam engine discussed in Rosenberg and Trajtenberg (2004) was patented in 1849[2], while Edison's carbon filament incandescent lamp mentioned in David (1990) dates back to 1880[3]. This is also true for the invention of integrated circuits, allegedly the engine of the Information & Communication Technologies (ICT) revolution, which can be traced back in the US to 1959[4].

This dissertation addresses most of these issues. It begins by studying countries' patterns of knowledge accumulation, providing a comprehensive characterization of how unequally distributed technological capabilities are and how obstacles to and opportunities for technological diversification change as countries develop. To overcome one of the main limiting factors of the literature, i.e. data availability, Chapter 3 proposes a statistical machine learning algorithm to identify the geographical locations of inventors within patent documents at the United States Patent and Trademark Office (USPTO) dating back to 1836 (the longest period considered to date). This unprecedented data initiative provides the opportunity to trace back the emergence, development, and diffusion of virtually all technologies for the past two centuries. The second part of the dissertation exploits this information to understand, first, the mechanisms through which technological change and knowledge accumulation operate to determine inter-

[2]See https://www.google.com/patents/US6162

[3]See https://www.google.com/patents/US223898

[4]See https://www.google.com/patents/US3138743. The first integrated circuit is attributed to Werner Jacobi (Siemens AG) in 1949 (https://www.google.com/patents/DE833366).

national trade patterns across countries. Finally, the last chapter studies how the early adoption of disruptive technologies, such as Electric & Electronic technologies at the beginning of the 20th century, can generate uneven and durable disparities in income and wages across places.

Specifically, Chapter 2 begins by studying countries' patterns of technological diversification and specialization along different stages of the development process, as reflected by their patenting activity at the USPTO. It relies on disaggregated data on patenting activity by type of technology for 65 countries and covering a period of 15 years (1993 to 2007) to estimate an econometric model that differentiates between diversification and specialization patterns. The empirical strategy is rich enough to capture general trends in terms of technological production (i.e. specialization patterns), and to single out factors affecting the emergence of new technologies (i.e. diversification patterns).

This chapter provides a richer and more comprehensive characterization of countries' patterns of technological development than those already available (Archibugi and Pianta, 1994; Cantwell and Vertova, 2004). It includes: a wider and more heterogeneous collection of countries, a novel characterization of technologies intended to capture their complexity and economic value, and a measure of cognitive proximity (or relatedness) among technologies as a key determinant of the likelihood of technological diversification.

These findings provide evidence regarding the importance of existing technological capabilities (Bell, 2009; Bell and Pavitt, 1992, 1997) and the relatedness of technologies (Boschma et al., 2014; Breschi et al., 2003; Jaffe, 1986; Neffke et al., 2011) in shaping possible paths of technological development. It shows that the likelihood of diversification is higher for those technologies that are related to countries' existing profile of competences and, more important, that this effect is stronger at earlier stages of development. Moreover, it provides evidence that countries tend to follow clear patterns of specialization along the development path, by moving toward more complex and valuable technologies. Overall, these findings are in line with and complement related evidence showing that well-performing countries tend to have a productive structure oriented toward the production of more sophisticated and valuable goods (Hausmann and Hidalgo, 2011; Hausmann et al., 2007; Hidalgo and Hausmann, 2009; Hidalgo et al., 2007; Lall, 2000).

As mentioned above, one of the main impediments to comprehensively study the emergence, development, and diffusion of technologies is data availability. This issue becomes evident when one considers the different empirical strategies pursued by scholars to measure the impact of key technologies. For instance, Hall et al. (2006) and Moser and Nicholas (2004) are forced to rely on the explanatory power of patent citations after 1975 to measure the impact of Electrical & Electronic (E&E) technologies 100 years earlier.

At present, few options exist for scholars seeking to analyze historical data linking the types of technologies invented to their place of invention. The primary source of information on the geography of knowledge production is the patent document. A patent provides exclusive intellectual property rights on an invention to its inventor (or assignee). In this way patents encourage the development of ideas. More precisely, the USPTO defines a patent as,"... the right to exclude others from making, using, offering for sale, selling or importing the invention". In exchange for such rights, the inventor (or assignee) is requested to provide detailed public disclosure of the patented invention. Public disclosure was designed to spur the diffusion of new ideas. Disclosure has also been key for academic researchers, as it provides a wealth of information on the business of science. All patents contain systematic information about the invention, the grant date, the name(s) of the inventor(s) and their home address(es), the name of the assignee and its business address, the date of application, the technological domains to which the patent applies, reference to prior academic publications and other patent documents on which the invention builds, and a brief abstract of the invention. This information is regularly used in economics, geography, and science and technology studies.

Although patent data are freely available at the USPTO Patent Full-Text and Image Database, they are not always available in a format that can be directly used for applied research. For some research questions, the raw data first have to be cleaned and processed (location disambiguation or inventor/assignee name disambiguation, for instance). A few structured, geo-referenced datasets have been developed in recent years. One of the most commonly used is the patent dataset of the National Bureau of Economic Research (NBER), which provides information on the state of residence of first inventor for patents from 1975 to 1999. Another widely used database for US patents is the Patent Network Dataverse, which provides longitude and latitude coordinates of inventor addresses for patents granted by the USPTO from 1975 to 2010. In a similar fashion, the REGPAT dataset of the Organisation for Economic Co-operation and Development (OECD) provides inventor locations (NUTS3 level for Europe, TL3 for other OECD countries) for patents filed with the European Patent Office (EPO) or the World Intellectual Property Organization (WIPO) from 1978 to 2011 (OECD, 2015).

However, these datasets only provide detailed geographical information on patents granted since 1975, the year when the USPTO began to electronically record patents. The main objective of Chapter 3 is to develop a well-structured, ready-to-use, comprehensive, and geo-referenced dataset of historical patents in the US covering the years 1836 to 1975 (HistPat). This database will provide geographical information (at the county level) on approximately 2.8 million patent documents (around 83% of all patents granted to US residents). The availability of this data creates an unprecedented opportunity to study the emergence, development, and diffusion of the most influential technologies in their historical context.

Chapter 4 exploits the wealth of information contained in HistPat to understand how technological change and the accumulation of knowledge and capabilities can generate disparities in how countries integrate into external markets and global value chains.

One of the oldest and most well-known theories of international trade, the Ricardian theory, highlights the role of technological dispersion as the key driver of bilateral trade. Differences in technological capabilities across sectors and across countries determine who exports what goods. Countries will benefit by specializing in those goods in which they have a comparative advantage and exchanging them for the other goods. One of the most influential Ricardian models was developed by Eaton and Kortum (2002), henceforth EK. In their paper, EK develop a general equilibrium Ricardian model with many countries and many goods that is able to capture how the opposing forces of technological change and geographic barriers affect bilateral trade. This seminal model, however, assumes that technological dispersion is the same for every country. In essence, this implies that technological knowledge is distributed in exactly the same way across industries for all countries.

Chapter 4 develops an extension of the EK model that allows the dispersion of technological activities to vary across countries. It provides a theoretical framework to understand how differences in the allocation of technological knowledge can affect trade patterns.

The main contribution of this chapter is to reconcile the theory with the empirics. On the one hand, most empirical studies of the effect of innovation on bilateral exports are specific to a country or a sector or are purely empirical. This means that while it has been shown that the gravity equation can be derived from various models, it is common practice to apply transformations to its canonical version as an ex post excercise. On the other hand, theoretical models often relegate the role of technology to a parameter that is later collapsed and netted out during estimation, or derived as a residual of the model. For instance, in Eaton and Kortum (2002) most of the technological parameters are collapsed into a dummy variable during estimation, invalidating any attempt to identify their impact on trade patterns. This is also a reflection of the current lack of appropriate measures of technological knowledge available to researchers.

In summary, Chapter 4 develops a variant of the EK model in which the process of innovation determines both the stock and dispersion of technological knowledge in each country. As in the EK model, a higher stock of technological knowledge and lower relative input costs foster exports. Unlike in the EK model, a country's overall comparative advantage (which determines overall exports) depends on technological dispersion. Intuitively, our model contains an additional term: the interaction between the within-country dispersion of technological knowledge and input costs. In particular, technological dispersion governs the advantage of having lower input costs. A lower dispersion benefits countries with lower input costs since their exports are determined by these and not technological differences. The opposite happens when technological dispersion is high, since exports for these countries are determined by differences in technology and not costs.

The empirical strategy to test the effect of technological innovation on bilateral exports relies on the historical data on foreign granted patents at the USPTO collected in HistPat and described in Chapter 3. It uses historical patent grants to construct measures of the stock and dispersion of technology by country and year. They are used to estimate the Kortum (1997) idea-generating model, which serves as the micro-foundation of EK.

Chapter 5 also exploits the database constructed in Chapter 3, here to measure the economic consequences of disruptive technological change. Even though it may appear intuitive and perhaps straightforward to relate technological change and income disparities, scholars have struggled to provide evidence of a causal relationship between technological change and income, wealth, or wage disparities. Two notorious strands of literature have addressed these issues in a comprehensive way, building solid theoretical grounds to understand the mechanisms behind this alleged relationship.

On the one hand, the literature on Skill-Biased Technical Change (SBTC) attributes the recent rise in wage inequality to the diffusion and proliferation of new and complex technologies. SBTC represents the idea that technological improvements can favor skilled over unskilled labor by increasing the former's relative productivity and, therefore, its relative demand. This theory suggests that new Information and Communication Technologies (ICTs) are complementary with skilled labor, thus placing technological change at the epicenter of the income distribution debate. A key problem for the SBTC hypothesis is that wage inequality seems to have stabilized after the 1990s, while ICTs continue to grow and improve at exponential rates.

On the other hand, a different but related strand of literature argues that the disruptive nature of particular technologies may explain sudden changes in the pace and direction of economic progress, which can generate durable disparities in per capita growth, income, and wages between those who adopt it and those who do not. In economics, these revolutionary technologies are referred to as "General Purpose Technologies" (GPTs). Widely known examples are the steam engine and electricity; ICTs are often mentioned as a contemporaneous example.

GPTs are characterized by possessing a wide scope for continuous improvement and elaboration, on the one hand, and high complementarity, on the other. The latter means that a GPT should be able to diffuse across a wide range of sectors, not only because it is used as an input in many different products and processes, but also because it is a technological complement of existing and new technologies. These characteristics are what make GPTs "engines of growth".

What connects the two strands of literature is the scattered empirical evidence supporting their claims, which creates a sort of empirical "puzzle" around them, mainly because these intuitive and theoretically sound arguments are unable to find support in data. When assesing the impact of a GPT on the economy, most of the evidence has been collected at an aggregated level (national); which imposes an immediate limitation on the possibility to univocally relate economic changes to the diffusion of a GPT. See, for instance, David (1990), Greenwood (1997), David and Wright (1999), Crafts and Mills (2004), Crafts (2004), and Jovanovic and Rousseau (2005). Additional evidence has been provided in the form of detailed historical accounts of the economic and societal changes generated by several GPT candidates throughout history (Lipsey et al., 2005). A valuable contribution toward providing evidence at a finer level of disaggregation is made by Rosenberg and Trajtenberg (2004). They use county-level information on the adoption of the Corliss steam engine (an alleged GPT) during the late nineteenth century and demostrate that it had a positive effect on population growth.

Therefore, there is a lack of comprehensive and disaggregated econometric evidence of a causal link between GPT adoption and growth, income, and wage disparities. Chapter 5 exploits the wealth of information contained in HisPat (developed in Chapter 3) to provide evidence of a GPT having a real impact on the economy. It combines economic and demographic data provided by the US Census Bureau and IPUMS (the Integrated Public Use Microdata Series) with detailed information on the geographical location, as well as the type of technological domain, of patents granted by the USPTO dating back to 1836.

Chapter 5 shows that the adoption of Electrical & Electronic (E&E) technologies had a positive effect on the wages and the income per capita growth of places after the 1900s. The empirical strategy relies on using measures of the adoption of E&E technologies prior the 1870s as an instrument to predict the adoption of E&E technologies between 1900 and 1930. It assumes that the early adoption of

E&E technologies (prior to the invention of the electric lightbulb or the establishment of the first power plant) were essential to the inventive structure of places 50 years later while not being correlated with the events that will determine growth and wages between 1900 and 1930.

In summary, this dissertation investigates the unequal and disruptive effects of knowledge accumulation. It begins by providing a general picture of the current state of the world's technological production and how unequally distributed technological capabilities are. After this first characterization, it focuses on one of the main drivers of cross-country disparities, i.e. their capacity to integrate into international markets. To do so it studies the mechanisms through which technological change and knowledge accumulation operate to determine international trade patterns across countries. The last chapter of this dissertation shifts its attention to the within-country effects of technological change by evaluating the uneven and durable disparities in income and wages generated by the diffusion of disruptive technologies, such as the invention of electricity.

CHAPTER 2

Climbing the Ladder of Technological Development

*This article has been produced in collaboration with P.A. Balland and A. Morrison. The PhD candidate is the first author of the article, which has been published in 2017 as "**Climbing the ladder of technological development. Research Policy, 46(5), 956-969**". We thank participants to the AAG conferences in Chicago (2015) and Tampa (2014), the organizers and participants of the Doctoral Workshop at Collegio Carlo Alberto, Torino (2015) and the DRUID academy 2015. We also thank Ron Boschma, the two anonymous referees and the Editor of Research Policy for their useful comments. Andrea Morrison acknowledges the support by The Netherlands Organisation for Scientific Research (NWO) [VIDI grant number 452-11-013].*

2.1 Introduction

Technological innovation is the main thriving force behind economic growth, industrial development, and the rise of living standards. However, only a handful of countries are actively developing new technologies. The United States, Western European countries, Japan and South-Korea host a small fraction of the world's population but are responsible for most technological advances. This unequal distribution of innovative activities sets the role played by different countries in the global value chain. Countries that innovate are able to capture a larger share of the value added, while others are trapped in less profitable activities. Climbing the ladder of economic development also requires climbing the ladder of technological development. But how do countries accumulate and develop their innovative capabilities? What kind of obstacles do they need to overcome? How could they identify opportunities to develop new areas of technological specialization?

These questions have attracted a lot of interest in the innovation literature. An extensive literature has analyzed the process of accumulation of technological capabilities in developing countries (see among others Bell and Pavitt (1992), Enos (1991), Lall (1992), Dahlman et al. (1987), King and Fransman (1984), Lee and Lim (2001), Kim (1999)). We have also a good understanding of patterns of sectoral and technological change (Breschi et al. (2000), Malerba and Orsenigo (1995)) and how their dynamics are shaped by cumulative and path dependent processes (Dosi et al. (1988), Malerba (1992), Patel and Pavitt (1997)).

Despite this extensive literature, we still have a limited understanding of how countries build new technological capabilities along the different stages of their economic development. In fact, cross-country quantitative studies exploring patterns of technological diversification and specialization have been very limited, and often restricted to the analysis of a handful of developed economies (see for instance Archibugi and Pianta (1994) and Cantwell and Vertova (2004)). As a result, we lack a robust and comprehensive bulk of evidence providing a general characterization of the type of technologies countries are more likely to produce, whether they tend to follow coherent patterns of technological specialization as they develop, and to what extent technological change is bounded to pre-existing technological capabilities.

This chapter will address these issues by analyzing countries' patterns of technological diversification and specialization along the development process, as reflected by their patenting activity at the United States Patent and Trademark Office (USPTO). We use disaggregated data on patenting activity by type of technology for 65 countries and covering a period of 15 years (1993 to 2007). We estimate an econometric model that differentiates between diversification and specialization patterns. In this way we are able to understand both, the general trends in terms of technological production (i.e. specialization patterns) and to single out factors affecting the emergence of new technologies (i.e. diversification patterns). We contribute to the literature by providing a richer and more comprehensive characterization of countries' patterns of technological development, which includes: a wider and more heterogeneous collection of countries, a novel characterization of technologies aimed at capturing their complexity and economic value, and a measure of cognitive proximity (or relatedness) among technologies as a key determinant of the likelihood of technological diversification.

Our findings provide evidence regarding the importance of existing technological capabilities (Bell, 2009; Bell and Pavitt, 1992, 1997) and relatedness among technologies (Breschi et al., 2003; Jaffe, 1986) in shaping possible paths of technological development.We show that the likelihood of diversification is higher for those technologies that are related to countries' existing profile of competences. Moreover, we show this effect to be stronger at earlier stages of development. On the other hand, we show that countries tend to follow clear patterns of specialization along the development path, by moving towards more complex and valuable technologies. Overall, our findings are in line, and complement related evidence showing that well-performing countries tend to have a productive structure oriented towards the production of more sophisticated and valuable goods (Hausmann and Hidalgo, 2011; Hausmann et al., 2007; Hidalgo and Hausmann, 2009; Hidalgo et al., 2007; Lall, 2000). The chapter is structured as follows: the next section presents the literature review and outlines the conceptual framework. In section 3 we illustrate the data and describe the methodology, while section 4 presents the results. The last section discusses the findings and sketches some policy implications.

2.2 Theoretical Background

2.2.1 On technological diversification and development

Within the innovation literature, country-level studies have focused on exploring patterns of technological specialization and/or diversification of advanced economies. For instance, Archibugi and Pianta (1994) found an inverse relationship between countries' technological size (measured as cumulative R&D expenditure) and the degree of sectoral concentration of technological activities. They covered the period 1975-1988 and used patent information for around a dozen of countries, mostly OECD members. Cantwell and Vertova (2004), Vertova (1999) and Vertova (2001) investigated patterns of technological specialization by looking at the patenting activity of a handful of developed economies between 1890 and 1990. They found a similar pattern regarding the relationship between countries' technological size and the degree of concentration in patenting activity, and additionally, that only few countries were able to specialize in fast-growing technological fields.

Besides the patent-based evidence, a more detailed overview of the topic has been provided by empirical studies using international trade data. For example, Lall (2000) explored export patterns of developing economies using bilateral trade data. He found that countries with an export portfolio oriented towards technology-intensive products tend to grow faster in the world trade. Similarly, Rodrik (2008) argued that a structural transformation in the export basket from traditional to non-traditional products constitutes the main engine of growth. Hausmann et al. (2007) developed an index to measure the quality of countries' export baskets and showed that countries specializing in products which lay higher on this quality spectrum tend to perform better. Moreover, Hidalgo et al. (2007), Hidalgo and Hausmann (2009), and Hausmann and Hidalgo (2011) found evidence that countries' export patterns become more sophisticated and complex as they develop. All in all, the above studies seem to agree on the fact that the distribution of the productive structure of well-performing countries tends to be biased towards the production of more sophisticated or/and valuable goods.

More recently, the role of relatedness among products and technologies and its effect on the diversification process of regions and firms has gained considerable attention, as reflected by the number and diversity of studies incorporating this concept (Frenken et al., 2007; Hidalgo and Hausmann, 2009; Saviotti and Frenken, 2008). The main idea behind the concept of relatedness is that firms' diversification possibilities (or regions/countries) are affected by the degree to which products or technologies are connected to one another, where the link between two technologies/products is usually measured as how much they share in terms of common scientific knowledge, technical principles, heuristics, and common needs in general. The concept of relatedness suggests that technological change may follow a path dependent process, in which production of new knowledge is bound to the existing knowledge (Dosi et al., 1988; Patel and Pavitt, 1997).

At country level, the pioneering study of Hidalgo et al. (2007) shows that countries are able to develop products which are close (in terms capabilities needed to produce them) to their current basket of products, providing evidence on the importance of product relatedness. Additionally, Saviotti and Frenken (2008) show that developing related products is beneficial in the short term, while long-term growth comes from the emergence of unrelated sectors.

At regional level, strong support has been found to the role of relatedness in driving either technological or sectoral development. For example, Boschma et al. (2014) and Rigby (2015) showed that technological relatedness was a crucial driving force behind technological change in U.S. cities. Colombelli et al. (2014) found that the development of new nanotechnologies is linked to the structure of the existing local knowledge base. Similarly, but focusing on industrial diversification of regions, Neffke et al. (2011), Boschma et al. (2014), and Essletzbichler (2015) showed that regions are more likely to enter into industries which are related to those already in place.

At firm level, results show that firms tend to follow coherent patterns of diversification. Jaffe (1986) and Breschi et al. (2003) found that firms' tend to diversify into groups of technological activities that share a common or complementary knowledge base. Yip (1982) studied firms' choices between internal development and acquisition and found that the likelihood of entry into new markets increases as those markets are more related to firms' own characteristics. MacDonald (1985) analyzed patterns of diversification within U.S. manufacturing firms, finding they were more likely to enter rapidly growing industries, and industries that were related to their primary activities through supply relationships or marketing similarities. Additionally, Teece et al. (1994) showed U.S. manufacturing firms maintain certain level of coherence while diversifying.

As shown above, robust evidence at both firm and regional level has convincingly shown the presence of a link between diversification and relatedness. However, comprehensive quantitative evidence at country level is lacking. The few existing studies reviewed above focus on product relatedness, while there are no country-level studies of technological diversification that have incorporated yet the role of relatedness.

In the next subsection we review the potential drivers and constraints for technological diversification/specialization, and the role of relatedness in this process. In doing so, we build a conceptual framework to understand the role of relatedness for technological diversification.

2.2.2 Technological Diversification, Relatedness, and Stages of Development

The process of technological accumulation observed at the country-level is a reflection of firms' capacity to accumulate and develop new technologies. Within the innovation literature, several factors have been acknowledged to affect firms' incentives to diversify or specialize in particular technologies. They include cases where diversification/specialization may be triggered by firms' attempts to move towards more profitable positions, which can be motivated either by inter-industry differences in the rates of return to R&D investment, as in the literature of technological opportunities (Jaffe, 1986; Klevorick et al., 1995; Laursen, 1999; Malerba, 2002), or by differences in the dynamism of the demand conditions, as in Schmookler (1966). Technological diversification can be also the result of firms' efforts to mitigate or avoid the effects of risk and volatility, which can negatively affect firms' productivity (Koren and Tenreyro, 2007). Additionally, entry barriers due to the requirement of high initial investments, in either technological or scientific knowledge, can deter diversification initiatives (Perez and Soete, 1988).

Besides these incentives, firm's potential for technological diversification depends heavily on its prior capabilities (Patel and Pavitt, 1997), which allows them to acquire, accumulate, and process the knowledge required for engaging in such a process. Understanding how the existing capabilities of a firm are related to the capabilities needed to develop new technologies is therefore crucial. We can identify two main channels through which this proximity affects firm's possibilities for technological diversification: economies of scope in the use of knowledge, and firms' absorptive capacity.

The economies of scope in the "use of one piece of knowledge" imply that the same type of knowledge could be used as an input in multiple technological fields (Penrose, 1959; Teece, 1982). Therefore, the more related two technological fields the bigger the share of common heuristics and scientific principles they rely on (Breschi et al., 2003), and consequently, the bigger the possibility to take advantage of the already acquired knowledge. For instance, economies of scope can affect barriers to entry by reducing investment costs, and the costs of acquiring the scientific and technological knowledge required to assimilate and carry out the innovation.

The concept of absorptive capacity refers to the fact that prior knowledge confers the ability to recognize the value of new information, assimilate it, and exploit it to commercial ends (Cohen and Levinthal, 1990). In particular, higher absorptive capacity allows for a better understanding of the challenges and opportunities for knowledge exploitation, and the benefits and costs associated with it. Firm's absorptive capacity affects its perception of the technological opportunities a given technology may offer, and its ability to form accurate expectations about the demand and risks associated with any diversification strategy. Hence, the more related a technology is with firms' absorptive capacity, the more likely it will accurately assess the benefits and costs associated with its adoption.

Based on the above theoretical arguments we set out our first research aim, focused on the role of relatedness in the process of technological diversification. We test to what extent technological capabilities within countries, and the relatedness among technologies, shape the possibilities of technological diversification. In particular, we aim at testing whether the likelihood of diversification into new technologies decays as the "technological distance" to existing capabilities increases. In line with the recent evidence on technological diversification at regional level we expect that pre-existing technological capabilities shapes patterns of technological diversification of a country.

The innovative process is not the outcome of an individual process of learning and capability accumulation, it is placed and determined within a larger system that supports and benefits from it (Edquist and Lundvall, 1993; Freeman, 1989; Lundvall, 2010; Metcalfe, 1995; Nelson and Rosenberg, 1993; Niosi et al., 1993; Patel and Pavitt, 1994).The national productive and innovative environment determines not only the opportunities and costs of diversifying/specializing into different technological activities, but also the way firms perceive opportunities and estimate costs. On the one hand, technological opportunities, risks, or demand conditions vary considerably across countries and technologies, creating differential incentives to diversify/specialize. For instance, Furman et al. (2002) find considerable differences across countries in terms of R&D productivity and type of inputs devoted to innovation, while Marin and Petralia (2015) show that sectors with high technological opportunities (TOs) in LACs differ greatly from those having high TOs in developed economies. Patel and Pavitt (1997) showed that the rate of technological accumulation is heavily affected by the country competitive environment. Lee and Lim (2001) pointed out the differences between developed and emerging economies, showing that the latter have higher propensity to innovate under specific technological regimes.

On the other hand, the scope and quality of capabilities available within a country affect firms' possibilities to take advantage of the economies of scale in the use of knowledge, or to appropriately assess the benefits and costs of new technologies. Even if the advantages of any particular technology are similar across countries, the lack of indigenous capabilities may shift costs of exploration upwards, as firms' should develop internally the entire set of competences otherwise available, among other options, via local technology alliances (see for instance Ahuja (2000), and Rowley et al. (2000)). The lack of capabilities may also affect how accurately benefits and costs of new technologies are estimated, to the point that possible development paths become unfeasible or unreasonable (given the perceived benefits and costs). As Nelson and Winter (1997) point out: "Real search processes take place in specific historical contexts, and their outcomes clearly depend in part on what those contexts contain in the way of problem solutions that are available to be 'found'."

Hence, technological diversity within and outside the firm may influence firms' capacity for combining and recombining their stock of existing knowledge, which may lead to new, and probably more valuable innovations (Fleming, 2001; Katila and Ahuja, 2002). Also, external knowledge may allow the firm to overcome lock-in traps (Levinthal and March, 1993; Levitt and March, 1988). There is of course the possibility to profit from valuable but 'distant' external knowledge, perhaps available in a different country, however, as Phene et al. (2006) show, firms have difficulties in absorbing and utilizing knowledge that is geographically distant. This is particularly true for developing economies, where the early experiences of international technological transfer showed that knowledge does not travel easily (Enos, 1991). Although the increasing codification of knowledge has certainly made the diffusion of technologies across borders easier, additional obstacles such as the growing complexity of the knowledge value chains still heavily hampers this process (Nelson, 2008). This also explains the renewed interest of scholars to the role of indigenous capabilities.

Based on above theoretical arguments we formulate our second set of research aims. First, we test whether the effect of technological relatedness decays as countries develop: advanced economies can rely on a more diverse and rich technological environment than developing ones, which allows them to incur into more distant re-combinations of knowledge. Second, we test whether countries systematically upgrade their technological structure following a coherent pattern of specialization, by increasing their participation in the production of more valuable and complex technologies as they develop.

2.3 Data and Methods

We start this section by describing the data sources and variables we will use through the analysis. Later, we explain our methodological approach, aimed at characterizing countries' patterns of specialization and diversification.

2.3.1 Data Sources and Variables

We use patent data as an indicator of innovative capabilities. Data on patenting activity was obtained from the "Patent Network Dataverse" developed by the Institute for Quantitative Social Science at Harvard University (Li et al., 2014) using original data from the USPTO. Patenting activity in the US has been used extensively in economics and innovation studies to address issues of global scope, this responds not only to the importance and size of the US technological market but also to the consistent and systematic way patents applications have been evaluated over the years; making data collected at the USPTO very suitable for comparisons, both across countries and time. We take advantage of the vast and rich information contained in patents regarding the technological domain (or technological class) patents belong to, and use it to construct variables aimed at capturing different aspects of technologies. We also make use of the Annual Survey of Manufactures (ASM) from the US census bureau to include economic measures such as the value added of sectors technological classes contribute to. Patents are widely used in the innovation literature because they provide a systematic and quantitative measure of new technological inventions. Nevertheless, the use of patent activity in the US as a measure of innovative capabilities is not exempt of criticisms; we acknowledge these limitations and also propose a way to deal with them in the methodological subsection.

Evaluating countries' technological trajectories requires being able to track and quantify countries' technological capabilities and their changes over time.

We measure patterns of specialization by computing countries' Revealed Technological Advantage (RTA) for each technological class (Soete and Wyatt, 1983). In particular,

$$RTA_{(c,j,t)} = \frac{Patents_{c,j,t}/\sum_j Patents_{c,j,t}}{\sum_c Patents_{c,j,t}/\sum_{c,j} Patents_{c,j,t}}$$

$$S_{(c,j,t)} = I[RTA_{(c,j,t)} > 1]$$

Where c stands for country, j for technological class, t for the time period (in three years intervals), and $I[.]$ represents the indicator function. We assign the nationality of a patent by looking at inventors' addresses; and consider a patent to be part of country c portfolio of competences whenever an inventor resides there. This index provides information on countries' patterns of technological specialization by comparing the share each technology represents in countries' own profile of patenting activity, relative the world average. Then the dependent dichotomic variable $S_{c,j,t}$ identifies technological classes where country c has a relatively high rate of patenting (i.e. RTA value above unity) when compared to the world average. We identify instances of diversification by considering those cases in which countries started patenting in particular technological domains. This can be done by focusing on cases where there was no patenting activity in the previous period, or at the beginning of the sample.

In order to construct the technology-level variables, we need to define the so-called Technological Space (TS). The TS was first addressed empirically by Jaffe (1986), who calculated relatedness among two given technologies by looking at how often they were used in combination with a third technology. In a similar manner, we construct the TS following the "product space" (PS) framework developed by Hidalgo et al. (2007). Within this framework, the TS can be seen as a network-based representation of the technological production, where nodes define technologies and ties among them indicate their degree of relatedness (see also Rigby (2015); and Boschma et al. (2014)). We identify 344 technologies using the USPTO patent classification and measure relatedness by counting co-occurrences of technologies (or technological classes) among patents. In particular, the degree of relatedness between technology i and j is measured as follows:

$$R_{i,j,t} = \frac{C_{ijt}}{\sqrt{S_{it}S_{jt}}}$$

Where $C_{i,j,t}$ counts the co-occurrences of technologies i and j, and S_i and S_j count the number of occurrences (size) of technologies at period t. This is often referred to as the cosine similarity measure and it has been widely applied in recent work (see Eck and Waltman 2009 for a detailed analysis). Therefore, the more often two technological classes appear together within the same patent, the more related they are, after controlling for the effect of size. This measure is therefore intended to capture the degree of common heuristics and scientific principles technologies share, by looking at how often they appear together in inventions.

We follow Hidalgo and Hausmann (2009) and Balland and Rigby (2017) to measure the complexity of technologies. The main idea is that, by analyzing the structure of the bipartite network connecting countries to the technologies they produce; complex technological structures can be characterized as those producing a wider range of exclusive technologies (i.e. non ubiquitous, produced by few countries). A country with a complex technological structure will not only produce technologies in many different technological domains, but they will do so in technologies requiring capabilities found only in a handful of countries. Therefore, the construction of an Index of Technological Complexity (ITC) requires combining information on both, the 2-mode degree distribution of a country (diversity) and the 2-mode degree distribution of the technologies it produces (ubiquity). We follow their 'method of reflections' and iteratively calculate:

$$k_{c,N} = \frac{1}{k_{c,0}} \sum_{t} M_{c,j} k_{j,N-1}$$

$$k_{j,N} = \frac{1}{k_{j,0}} \sum_{c} M_{c,j} k_{c,N-1}$$

The matrix $M_{c,j}$ takes value 1 if country c is a significant producer of technology j, and zero otherwise. We consider country c to be a significant producer of technology j if its $RTA > 1$. $k_{c,0}$ and $k_{j,0}$ measure levels of diversification of a country (the number of technologies produced by that country), and the ubiquity of a technology (the number of countries producing that technology). Each additional step incorporates feedback effects and produces more precise estimates of the knowledge complexity of countries by using information on the complexity of technologies they produce. By the same token, $k_{j,N}$ estimates the knowledge complexity of a technology using information on the complexity of countries that produce this technology. For a detailed description of the procedure and the properties of the indicator see Hidalgo and Hausmann (2009) and Balland and Rigby (2017)[1].

We also include three more variables: the Herfindhal concentration index (across countries) of each technological class, their size, and the value added of industries technological classes contribute to. The variable measuring class size aims at controlling for scale effects. Even though the level of patenting activity has been used to measure technological opportunities, as in Laursen (1999); here we don't go beyond any interpretation other than capturing differences in the propensity of patenting activity among technological classes. As it is custom-

[1]Values were scaled around the mean. We iterated the method of reflections 20 times. At the last step the correlation with the previous iteration was above 0.999.

Table 2.1: *Technology-level Variables*

Variable Name	Description	Source
Value Added	Value added of industries technologies contribute to (in billions of dollars)	ASM
Size	Number of patents within the technological class (in logs)	USPTO
Herfindhal Index	Herfindhal concentration index of patenting activity	USPTO
ITC	Index of Technological Complexity (as in Hidalgo et al. 2009)	USPTO

ary, we use the Herfindhal index as an indicator of the amount of competition among countries within a particular technological domain. The third variable (i.e. value added) aims at capturing the economic value of technological classes. We measure it by computing the value added of industries technological classes contribute to. Note that the USPTO provides a concordance linking technological classes to standard industrial classifications; which can be used to match technological classes with characteristics of the industries in the US[2]. We use this concordance to generate a weighted average of the value added technologies contribute to, using information of the value added by industry, more specifically:

$$VA_i = \sum_s W_{i,s} VA_s$$

Where i indexes technological classes and s industrial sectors. $W_{i,s}$ is a weighting matrix which assigns a weight proportional to the amount of patents within class i contributing to industry s. Table 2 below summarizes all the tech-level variables.

In addition to the variables described above, we include a measure providing information about the proximity of countries' existing capabilities to every technology. For each technological class, we quantify the degree to which countries' current technological production 'surrounds' that given technology. This measure uses information about the relatedness among technologies as well as

[2]Concordances can be found here: http://www.uspto.gov/web/offices/ac/ido/oeip/taf/data/

countries' profiles of indigenous capabilities to calculate, for any given technology, the proportion of related technological classes countries' shows patenting activity on. It varies from 0 to 1, with higher values indicating there are competences nearby a given technological domain (where distance is measured in terms of proximity within the TS), and it is calculated as follows[3] :

$$Density_{c,j,t} = \frac{\sum_i R_{i,j,t} X_{c,i,t}}{\sum_i R_{i,j,t}}$$

Where $X_{c,i,t}$ takes value 1 if country c shows patenting activity in technology i at time t, and 0 otherwise. This variable will be used to disentangle whether and how indigenous capabilities, as well as relatedness among technologies, affect possibilities for technological diversification. Table 3 below shows descriptive statistics of the aforementioned variables. After combining all different sources of information we end up with a sample of 65 countries covering a 15 years period, from 1993 to 2007.

Basic statistics show that only a 25% of the cases represent instances of diversification, in the way defined it earlier (i.e. when $S_{c,j,t} = 1$ with no prior patenting activity, in the table below $NPA_{c,j,t-1}$ identify countries with no prior patent activity in that technology while $PA_{c,j,t-1}$ identify those who had). Additionally, note that there exist a high correlation between the ITC and the size of technological classes, meaning that in order to appropriately capture any patterns involving the complexity of technologies it will be necessary to net out, or control for, the effect of size.

Let's consider a finer description of the data used in this study. Figure 2.1 below, characterizes the technological production of three selected countries at different stages of development; Argentina, Korea, and Germany. It shows, for

[3]This measure has been widely applied in recent studies, see Hidalgo et al. (2007), and Boschma et al. (2014) for instance.

Table 2.2: *Main Descriptive Statistics*

	Mean	SD	Min	Max
Specialization	0.21	0.41	0	1

	$PA_{c,j,t-1}$	$NPA_{c,j,t-1}$	Total
$S_{c,j,t} = 1$	0.748	0.252	1.00
$S_{c,j,t} = 0$	0.29	0.71	1.00

Technology-Level Variables

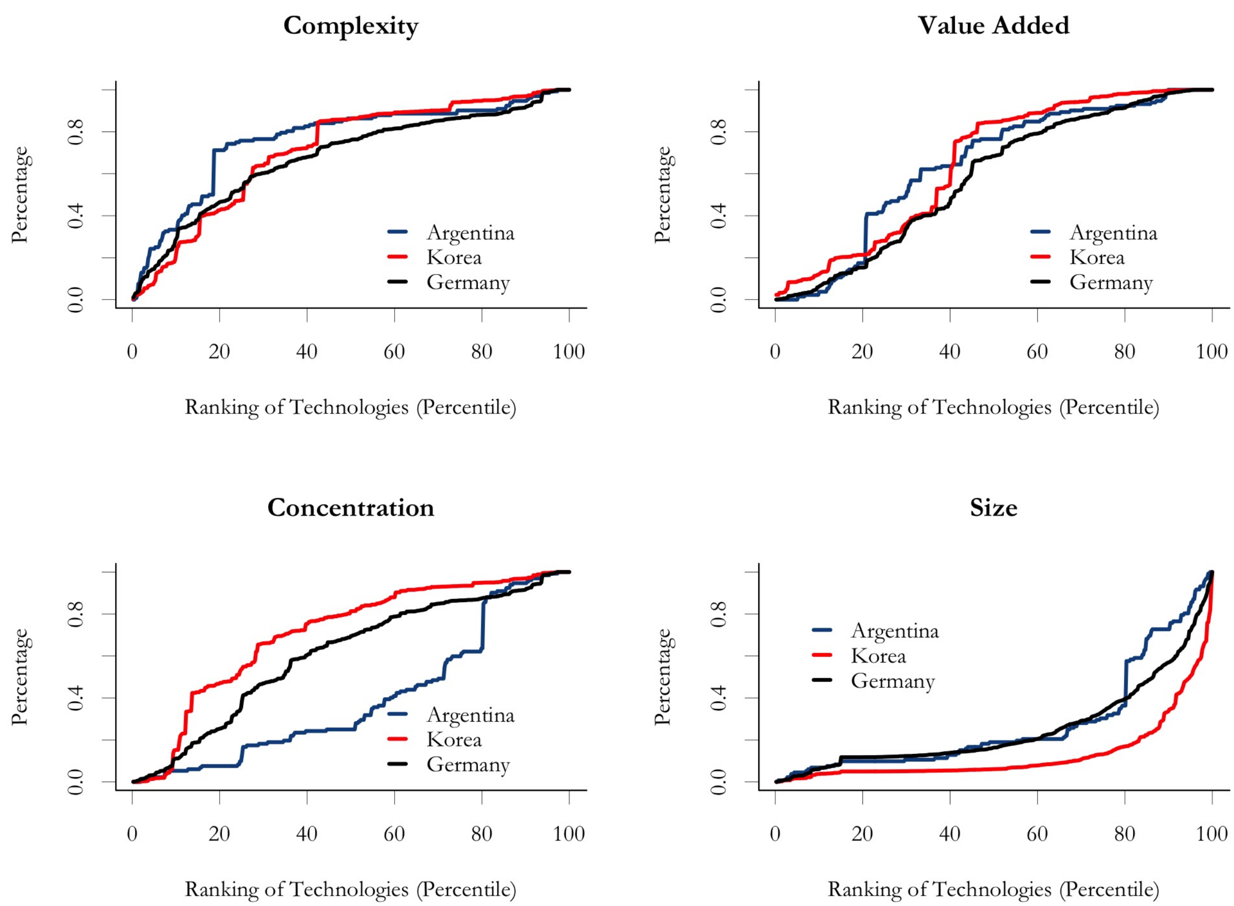
	Mean	SD	Min	Max	Correlation Table				
Value Added	79.34	40.47	5.53	184.7	1				
Log Size	5.68	1.53	0.51	9.52	-0.303	1			
Herfindahl	0.36	0.13	0.11	0.92	-0.104	-0.054	1		
ITC	0	1	-4.45	3.96	0.340	-0.660	-0.171	1	
Density	0.40	0.19	0	1.00	-0.120	0.017	0.133	-0.173	1

Number of Countries: 65
Number of Technologies: 344
Coverage: 1993-2007 (5 intervals of 3 years each)
Number of Observations: 92300

the main technological variables and last period of the sample, how countries are distributing their technological production.

All graphs show on the horizontal axis a ranking of technological classes ordered according to the variable in question. For instance, the first graph entitled 'Complexity' orders technologies from the less to the most complex along the horizontal axis. Vertical axes display the cumulative share of countries' technological production. Therefore, the first graph shows that Argentina accumulates almost 80% of its technological production among the 20% less complex technologies, while Korea and Germany concentrate only around 45%. A similar pattern characterizes the distribution of valuable technologies, where Argentina more than doubles the concentration of its technological production within the 20% less

Figure 2.1: *Characterising Countries' Technological Production*



valuable technologies (around 40%), when compared to Korea and Germany (less than 20%). The last two graphs also depict different patterns of technological specialization across countries. Remarkably, Korea shows a technological profile biased towards the production of technologies that are relatively big in size.

The next section describes the methodological approach, which among other things, aims at providing an answer to the question of whether these seemingly intuitive patterns hold on in general, and when possible sources of bias are controlled for.

2.3.2 Methodology

In this subsection we outline the methodological approach to address our main research aims. On the one hand we aim at characterizing countries' patterns of technological diversification by estimating how the likelihood of entering into new technological activities decays as the 'technological distance' to existing technological capabilities increases. Additionally, we evaluate whether this effect decays as countries develop. On the other hand, we test to what extent countries follow a coherent pattern of specialization, by increasing their participation in the production of more valuable and complex technologies as they develop.

To identify instances of diversification we restrict our sample those cases where a particular technology was not already produced in a country prior the period of analysis. By doing this, we are able to evaluate which factors affect the likelihood that technologies will emerge. As there is not a clear consensus on how to empirically identify instances of diversification we propose two different ways. First, restricting the sample to cases where there was barely or no activity at the beginning of the sample. In doing so, we can assess if technological production emerged in any of the subsequent periods. Second, we restrict the sample to cases where there was barely or no activity in the previous period, for each time period. This way we are able to consider only the previous period as a benchmark, where diversification instances should occur from one year to the next. We consider that a country has barely or no activity in patenting if RTA values are below 0.1 as in Hidalgo and Hausmann (2009)[4].

We use patented inventions at the USPTO as a proxy of innovative capabilities of countries, which is an imperfect measure, as it has been argued that: (a) They provide an incomplete characterization of the production of knowledge, mainly because firms may not patent and choose secrecy instead; (b) They may induce biases, as the rate of patenting differs greatly across sectors and time; (c) It is not

[4]Results are robust to changes in this cutoff.

possible to account for differences in the economic value of the inventions; (d) It is not clear whether patenting activity at the USPTO can be taken as an appropriate source of information to evaluate global patterns of knowledge production; as rates of participation may vary systematically across countries, leading to a misrepresentation of their competences.

Our methodological approach aims at both providing answers to our research objectives while overcoming the potential limitations of the data we work with. We propose to estimate two separate linear probability models (one to characterize diversification patterns and another to evaluate specialization patterns), which include tech-level dummies to overcome the potential biases that differences in patenting rates may have on our results (b). Consider the following specifications, with c indexing countries, j technological classes, and t time periods:

Diversification Equation (on the restricted sample):

$$S_{c,j,t} = \theta_1 Density_{c,j,t-1} + \theta_2 Density_{c,j,t-1} * GDP_{c,t} + \sum_k \beta_k T^k_{j,t} + \delta_c D_c + \delta_j D_j + \delta_t D_t + \epsilon_{c,j,t} \quad (2.1)$$

Specialization Equation:

$$S_{c,j,t} = \theta_1 Density_{c,j,t-1} + \sum_k \beta_k T^k_{j,t} + \sum_k \beta_k T^k_{j,t} * GDP_{c,t} + \delta_c D_c + \delta_j D_j + \delta_t D_t + \epsilon_{c,j,t} \quad (2.2)$$

Where $S_{c,j,t}$ identifies technological classes country c has specialized in the production of (having an RTA above unity), $GDP_{c,t}$ accounts for the GDP per capita of countries[5], $Density_{c,j,t-1}$ measures country c technological proximity to technology j, $T^k_{j,t}$ for $k \in 1-4$ contain all technology-level variables described in Table 2, $D_i for i : c, j, t$ represent a set of country, technology, and time variables respectively, and $\mu_{c,j,t}$ is the error term. Note that we account for differences in the

[5]Constant (based in 2005) and PPP adjusted measures of GDP per capita were obtained from the World Development Indicators (WDI) database provided by the World Bank.

economic value of the inventions (c) by including among the tech-level variables the value added of sectors technological classes contribute to.

Note that the first equation aims at identifying whether the effect of having competences in related technologies is higher when entering into a new technological domain and additionally, whether this effect changes as countries develop. Tech-level variables along with country, technology, and time fixed effects are consider as control variables in this equation.

There are two key differences in the specialization equation. First, we aim at evaluating the distribution of the technological production, disregarding whether the production of that technology is new to the country or not (in this case we use the entire sample). Therefore we explore specialization profiles according to characteristics of the technologies, which are interacted with GDP per capita of countries to evaluate whether countries tend to follow coherent patterns of specialization as they develop. Second, unlike in the previous equation, the density variable is here introduced as a control variable along with the rest of the dummies. As in the previous specification, we don't want that any possible correlation between the density of technologies in the TS biases any of the coefficients of the technological variables.

It may still be the case that patents records at USPTO may not fully reflect countries' profiles of competences (d), meaning that the US technological market may not be an unbiased sample of the world production of patents and its distribution over countries. We address this issue by including a sample selection test to identify whether the particular selection of countries we use in the analysis may have influenced our results. This test, proposed by Wooldridge (1995), does not require imposing any restriction on distribution of the error term in the regression equation, and allows for arbitrary serial dependence. In order to carry it out we need to calculate the likelihood any given country is part of our sample at any period of time and according to different country-level characteristics, to later obtain the Inverse Mills Ratio (IMR) and include it the regression equation. A significance t-test on the additional variable will then be used to detect the presence of sample selection. As the construction of this test involves defining and describing a whole new set of country-level variables, we relegate a detailed outline of this procedure to the appendix.

The use of secrecy over patenting as a method of protection (a) cannot be measured unless a firm-level survey spanning different technological domains is conducted, however, there are no obvious reasons to believe that this may considerably affect our results, especially after having controlled for specific tech-level effects in the regression.

2.4 Results

In what follows we describe the results the econometric model. Table 2.3 below reports the results of the linear probability models. The first two columns correspond to equation (1) for the two different subsamples defined at the beginning of section 3.2, while the third column corresponds to equation (2). Standard errors were clustered by country, technology, and time according to Cameron et al. (2011).

Results in columns one and two test the likelihood a country will diversify into the production of a particular technology. The former identifies technologies that were not already produced in a country by restricting to cases in which countries show a RTA value less than 0.1 in the previous period, while the later considers RTA values below 0.1 at the starting point of our sample. Both specifications show that having capabilities in related technologies is important when entering into a new technological domain, as reflected by the positive and significant coefficient of the density variable. The likelihood a new technological capability will emerge is higher the closer that technology is with respect to the profile of existing capabilities in that country. However, when the density variable is interacted with countries' GDP per capita, this effect diminishes, showing that having related capabilities is less important as countries develop. Results show that diversification in unrelated technologies is less likely to occur at early stages of development, or putting it in another way, that developing countries tend to diversify incrementally. As expected, it is less likely to find diversification into more complex and less concentrated technologies, as shown by the negative coefficients of the ITC and Herfindahl index variables. There are no significant differences in the likelihood of diversification with respect to the economic value of technologies (no significant coefficient of the Value Added variable).

Column three summarizes the results of the specialization equation, which show that there is a significant and positive reinforcement of having developed related capabilities, as captured by the density coefficient. The latter means that technological production tends to cluster in the technological space. Additionally, it is less likely to find countries specializing in complex and valuable technologies. However, when we interact the tech-level variables with countries' GDP per capita, results show that the likelihood of specialization increases for complex and valuable technologies as countries develop (see positive and significant coefficients of ITC and Value Added variables when interacted with GDP). In sum, developed countries tend to show a distribution of their technological production biased towards the production of less concentrated, more complex, and more valuable technologies. Additionally, note that the coefficient of the sample selection test is not significant in all three specifications. The log size variable is always positive and significant as expected; the straightforward meaning is that in larger technological classes there are higher opportunities of diversification, and this is true at any stage of development (as suggested by the not significant coefficient of the interaction variable log size*GDP).

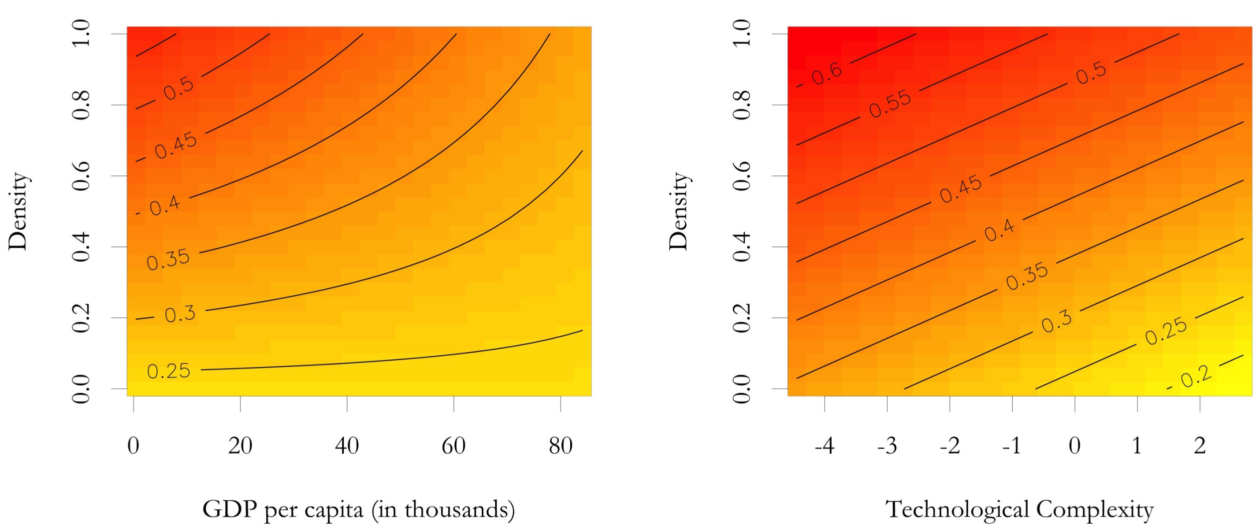
Table 2.3: *Results of the Econometric Model*

	Diversification Eq.	*Diversification Eq.*	*Specialization Eq.*
	(1)	(2)	(3)
Density	0.208**	0.27***	0.87***
	(0.034)	(0.025)	(0.0392)
Density * GDP	-0.0011*	-0.0029***	
	(0.00058)	(0.00050)	
Tech-Level Variables			
Value Added	-0.0005	-0.00021	-0.00061*
	(0.00043)	(0.00030)	(0.00028)
Log Size	0.0185**	0.032***	0.031***
	(0.0069)	(0.0048)	(0.005)
Herfindahl Index	-0.169***	-0.169***	-0.248***
	(0.033)	(0.036)	(0.067)
ITC	-0.025***	-0.0140***	-0.025***
	(0.0034)	(0.0021)	(0.004)
Value Added * GDP			0.000008**
			(0.000003)
Log Size * GDP			-0.000087
			(0.00016)
Herfindahl Index * GDP			-0.0059**
			(0.0022)
ITC * GDP			0.00067*
			(0.00029)
Sample Selection Test	0.011(0.008)	0.0023 (0.005)	0.0107(0.0136)
Adjusted R-Squared	0.079	0.090	0.27
Tech FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
Observations	54,370	58,808	89,440

*p<0.1; **p<0.05; ***p<0.01

Figure 2.2 below summarizes our results regarding the importance of having related capabilities for diversification, and its differential impact along stages of development (darker colors correspond to higher probability values). The left panel shows how the likelihood of diversification changes as proximity to technological capabilities increases (as reflected by the density coefficient on the vertical axis) and along stages of development[6] (i.e. GDP per capita on the horizontal axis). Note that the likelihood of diversification more than doubles (from 0.25 to over 0.5) if we move from related to unrelated technologies for low-income countries (i.e. while going up along the vertical axis from the left-lower corner). When considering high-income countries (rightmost part of the graph along the horizontal axis) the difference in the likelihood of diversification between related and unrelated technologies barely changes[7].

Figure 2.2: *Likelihood of Diversification & Stages of Development*



[6]Technology attributes were set to their average values. Predictions are made over all dummy variables and then averaged together.

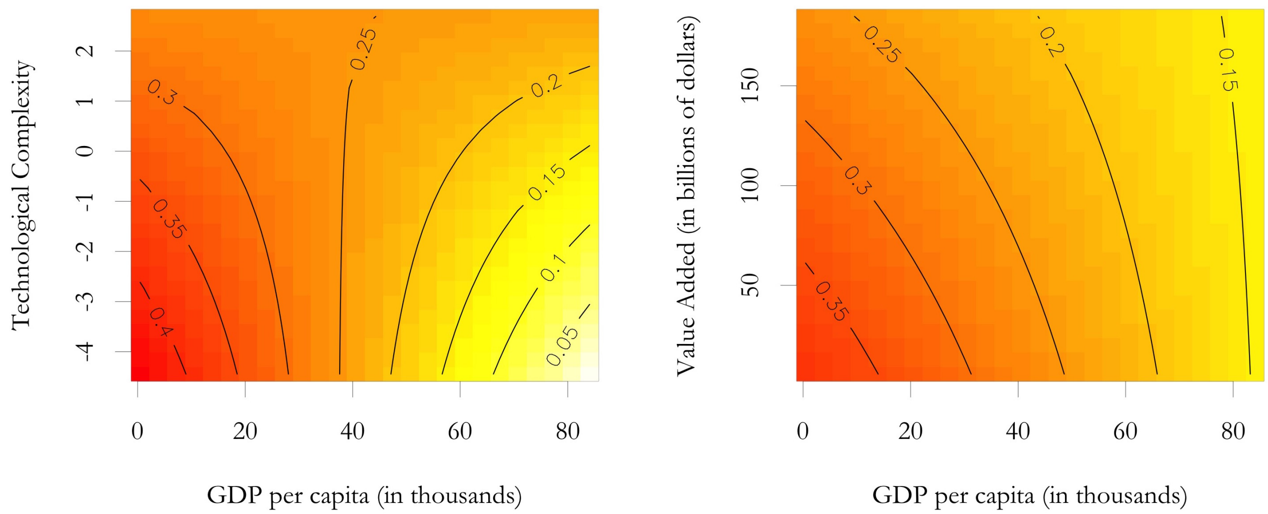
[7]Note that while plotting predicted values we are keeping country and technology fixed effects constant, meaning that these figures do not incorporate shifts in the likelihood of diversification captured by fixed effects.

Figure 2.2 exemplifies another aspect of our results. It shows how different characteristics of technologies such as their complexity and proximity to countries' indigenous competences (density) interact to determine possible paths of technological development. On average, the likelihood of diversification triples (from 0.2 to 0.6) if we move from complex and unrelated technologies (lower-right corner of the figure) to ubiquitous and related technologies (north-upper-left corner of the figure).

Results in Table 2.3 show that different characteristics of technologies, such as their complexity and economic value, can be associated with specialization patterns along the stages of development. It is not so straightforward to see, however, to what extent these statistically significant results can be translated into meaningful changes of specialization profiles. Figure 2.3 below shows specialization profiles (predicted values $S_{c,j,t}$) along different characteristics of the technologies and stages of development (as before, darker colors correspond to higher probability values). The left panel plots predicted probabilities for different values of technological complexity and GDP per capita, showing that it is four times more likely to find low-income countries specializing in less complex technologies than finding high-income countries producing that sort of technologies (it goes from values above 0.4 to less than 0.1). For high-income countries, as opposed to low-income ones, the mode is situated in the upper part of the graph. Similar results can be found for value added (right panel), where production of less valuable technologies is relatively more likely to be found in low-income countries (e.g. 0.35 in low income vs 0.2 for high income countries).

We would like to emphasize that our results do not imply that technological diversification can't happen unless you have certain related level of existing technological competences. Note that independently of their proximity to existing competences, complexity, and value, each technology has different barriers to entry. As shown in Park and Lee (2006) and Lee (2013), latecomer countries

Figure 2.3: *Specialization Patterns & Stages of Development*



can enter technological niches where windows of opportunities open up more frequently, so reducing barriers to entry. These opportunities emerge usually in technologies with short life cycle, where latecomers are less at disadvantage with advanced economies. In these niches, as shown by the experience of Korea, leapfrogging can occur (Lee and Lim, 2001). In our econometric specification these characteristics are absorbed by tech-level fixed effects, as they remain practically invariant over the time span we considered. We show in the appendix that our fixed effects are correlated with the technology-regime variables used in Park and Lee (2006) and Lee (2013).

To conclude, a potential concern with our econometric specification is that results may be biased given that we imposed a linear relationship in the way our variables of interest are affected as GDP per capita increases. An alternative option is to divide the sample according to the stage of development of countries (proxied by the GDP per capita level), and evaluate whether we observe a similar pattern. We carry out this analysis in the appendix and show that results are robust to changes in the specification of the model.

2.5 Concluding Remarks

This study has examined how countries diversify and upgrade their technological production as they develop. We analyzed diversification and specialization patterns of 65 countries using disaggregated patent data by type of technology from the United States Patent and Trademark Office (USPTO) over a period of 15 years. Patenting activity of countries show that the development of new technologies is a highly cumulative and path dependent process, in which technological upgrading emerges out of pre-existing knowledge bases and patterns of specialization. The likelihood of diversification into a new technological activity is higher for those domains that are related to countries' existing profile of competences. Countries climb the ladder of technological development rung by rung, as new capabilities have to be built-up gradually. However, patterns of technological development change in two important ways as countries develop. First, diversification is more heavily constrained by related, indigenous capabilities at early stages of development. At later stages of the development process countries are able to make bigger jumps and develop new technologies that are less and less related to their current knowledge bases. Second, the type of technologies in which developed and developing countries are specialized is different. Along their economic development, countries tend to upgrade their technological structure by specializing into increasingly complex and more valuable technologies.

The key role of technological relatedness in driving changes in the economic structure of countries and regions has been highlighted in several studies, based on different geographical and industrial unit of analysis. Our findings are consistent with Hidalgo et al. (2007) on export diversification of countries, Neffke et al. (2011) on industrial diversification in Swedish regions, and Boschma et al. (2014) on entry of technologies in U.S. cities.

These results have several implications for development strategies of countries. It is important to understand that developing strong capabilities and leadership in technologies out of the blue is nearly impossible. It is tempting to try to become a leader in information technologies or biotech because these technologies are fashionable and profitable, but such a technology-targeted strategy is a lottery. What if countries do not have the specific knowledge and capabilities required to become a successful producer of the targeted technology? In many cases, such a blinded policy will end up in wasted taxes. This issue is particularly important for developing countries, where diversification possibilities are even more heavily constrained by indigenous capabilities. Our results suggest that a more efficient policy strategy consists of carefully assessing the pre-existing knowledge bases of a country, considering indigenous capabilities as a starting point. The specific analysis of the technological strengths and weaknesses of a country can reveal unexploited opportunities, which may lead to alternative development paths. The methods used in this chapter can be applied to assess national knowledge bases and develop strategies for their diversification or upgrading.

Although we suggest that an inside-out strategy of technological development is more desirable than a technology-targeted policy, a word of caution is necessary here. It is important to note that a policy supporting only technologies that are very closely related to existing capabilities is also quite likely to fail. Such a development strategy won't be risky in the short run, but it can lock-in technological development of countries in the long run. This is especially true for developing economies, which tend to make shorter technological jumps (i.e., more related) in less sophisticated activities (i.e. less complex). Therefore, there is a threat for developing countries to become locked in the production of less sophisticated technologies, and innovation policies should not re-enforce this path-dependent process by narrowing down technological opportunities. It is important for developing countries to continuously aim at producing more complex technologies in order to upgrade their technological structure. Sometimes it will require supporting activities that are not necessarily the most related to existing capabilities. In this case, it will require bigger investments (also in complementary education or supporting infrastructure). Another important leg of this strategy is to engage actively in international knowledge networks in order to gradually build-up the necessary competencies.

An alternative strategy for latecomers is to identify technological niches where windows of opportunities can open up more frequently. Park and Lee (2006) show that these opportunities usually emerge in technologies with short life cycle, where barriers of entry are lower and latecomers are less at disadvantage with advanced economies. In these niches, as shown by the experience of Korea, leapfrogging can occur (Lee and Lim, 2001). However, the latter is true only if latecomers have accumulated enough technological capabilities, otherwise rapid technological change becomes an additional barrier for latecomers (Lee, 2013).

An important caveat of this chapter is that we focus exclusively on patents. Patents are praised as the only systematic measure of invention, but also criticized because they only capture some specific types of innovation and technologies. Many generic forms of innovation, especially in developing countries, won't show up in patent data. Similarly, patents don't capture the innovation outcomes that are generated by the imitative activities of firms. These activities can range from pure imitation of existing technologies to creative design of novel processes and products, whose innovative content can be regarded as relevant as that in patents (Bell, 2009). It is important to bear these critiques in mind while interpreting the results and critically assessing the policy implications we are drawing. Also, it is not clear whether the results we report will also hold true for non-manufacturing sectors, or for sectors relying less on patenting activity. This could be addressed in future research combining patents with innovation surveys, both in the manufacturing and service industry. An interesting question here is that service sectors can also serve as catalyzers, enhancing technological diversification possibilities, as they generally span over different types of productive activities.

2.6 Appendix

2.6.1 Test for Sample Selection

As mentioned in the methodological section, one appropriate concern is that the sample we have used to estimate our econometric model may not be an unbiased representation of the world production of patents and its distribution over countries. We address this issue by including a sample selection test to identify whether the particular selection of countries we use in the analysis, those who had activity in all periods, may have influenced our results. This test, proposed by Wooldridge (1995), does not require imposing any restriction on distribution of the error term in the regression equation, and allows for arbitrary serial dependence. In order to carry it out we need to calculate the likelihood any given country was included in our sample at any period of time, and according to different country-level characteristics, to later obtain the Inverse Mills Ratio (IMR) and include it the regression equation. The IMR is computed after estimating a separate probit model at each period. A significance t-test on the additional variable will then be used to detect the presence of sample selection. The purpose of this appendix is to describe the construction of this test, which involves defining and describing a whole new set of country-level variables.

We use the World Bank's World Development Indicators (WDI) database; which provides a diverse collection of development indicators compiled from officially-recognized international sources. Additionally, data on bilateral trade flows and distances were obtained from the BACI database developed by the CEPII.

Table 2.4 below provides a short description of all country-level variables. These variables aim at capturing whether the inclusion in the final sample can be explained by country-specific characteristics, which may have an impact on our estimated coefficients if the country-level fixed effects cannot appropriately control for them in the regression equation.

Table 2.4: *Country-Level Variables*

Variable Name	Description	Source
Distance	Thousands of km to US (using capital cities)	CEPII
Language	Whether or not English is an official language	CEPII
Population	Population in millions	WDI
GDP per capita	GDP per capita in thousands of 2005 US dollars	WDI
Outward Orientation	Share of total exports to GDP	WDI
Trade	Share of exports oriented to the US market	BACI

Distance to the US market, as well as language barriers, directly shift up barriers to entry, therefore affecting the likelihood that firms within a country will regularly patent at the USPTO. Additionally, countries' outward orientation and the importance of the US market as a commercial partner may increase the likelihood of participation in the US 'technological market'. The former by reflecting the need of technological upgrading to compete in international markets, while the later by capturing the effect that strong trade relationships have on the orientation of patenting activity. Lastly, population and GDP per capita aim at capturing size effects and remaining factors that could be related to countries' level of development. Table 2.5 shows descriptive statistics of the variables

Table 2.5 below shows the results of estimating the likelihood any given country was included in our sample at any period of time (a separate probit model for each period, first period is dropped as we use lagged values in the main equation). All country-level variables are highly significant, consistent over time, and in line with what we could expect; the likelihood increases with population, the degree of outward orientation, and the level of development of the country. Additionally,

Table 2.5: *Main Country-Level Descriptive Statistics*

Variable Name	Mean	SD	Min	Max
Distance	8.51	3.57	0.55	16.1
Language	0.31	0.46	0	1
Population	33.54	129.7	0.02	1311
GDP (pc)	8.91	13.30	0.09	84.07
Outward O.	0.89	0.53	0.16	4.79
Trade	0.15	0.19	0	0.91

Number of Countries: 169
Coverage: 1993-2007 (5 intervals of 3 years each)

border and distance effects have the expected signs. Sharing language impacts negatively given the substantial proportion of developing countries, in number, where English is an official language. There is a negative relationship between the proportion of exports directed to the US and the likelihood of a country being included in the sample, this looks somehow unintuitive, as you would expect that strong trade relationships may increase the amount of patenting activity between countries. However, this is mainly due to the fact that most of this US oriented exporters are from developing economies, and mostly exporting low-end goods.

It may be the case that these relationships are non-linear, or that interaction effects are important. We included all possible interactions, and third-degree polynomials of each variable in the version used for estimation; however, as results don't change in any significant way, and because we wanted to provide a parsimonious description, we decided to report this summarized version. Additional results are available upon request.

Table 2.6: *Results (Probit Models)*

	Period 2	Period 3	Period 4	Period 5
	(1)	(2)	(3)	(4)
Intercept	-0.794***	-0.941***	-0.972***	-1.053***
	(0.021)	(0.0201)	(0.019)	(0.019)
Distance	-0.00006***	-0.00006***	-0.00007***	-0.00007***
	(0.000)	(0.000)	(0.000)	(0.000)
Language	-0.818*	-0.901***	-0.938***	-0.917***
	(0.0147)	(0.015)	(0.015)	(0.016)
Border	3.982***	3.709***	3.89***	4.184***
	(0.04)	(0.04)	(0.04)	(0.04)
Trade	-1.552***	-1.235***	-1.38***	-1.65***
	(0.0453)	(0.038)	(0.037)	(0.041)
GDP pc	0.094***	0.094***	0.087***	0.08***
	(0.001)	(0.001)	(0.001)	(0.001)
Population	0.030***	0.028***	0.027***	0.027***
	(0.0005)	(0.0004)	(0.0005)	(0.0004)
Outward Orientation	0.203***	0.307***	0.441***	0.519***
	(0.0096)	(0.011)	(0.011)	(0.012)

*p<0.1; **p<0.05; ***p<0.01

2.6.2 Robustness of the Results: Dividing Countries into Developed and Developing Economies using World Bank's Classification

In our baseline model we interacted the variable GDP per capita with the main variables of interest in order to evaluate whether effects are different as countries develop. This implies imposing a linear relationship in the way our variables of interest are affected as GDP per capita increases. If this assumption does not hold, coefficient estimates could be biased.

An alternative option is to divide countries according to their stage of development using a dummy variable based on the categories created by the World Bank. We can then evaluate whether we observe a similar pattern with respect to the baseline model. Therefore we perform a new estimation dividing countries into High-Income and Low-Middle-Income economies. This classification is taken directly from the World Bank, and is based on GDP per capita[8] . Table 2.7 below provides the list of countries that fall in each category.

Table 2.8 below shows the result of estimating the same econometric model than in the baseline case but interacting the variables with a dummy identifying High Income countries. In columns 1 and 2 we test the likelihood that a country will diversify its technological activities, by patenting in a new technological class in which it had no patenting activity before. As in the baseline model, we identify instances of diversification by restricting to cases in which countries show a RTA value below 0.1 at the starting point of our sample, and in the previous period.

The results reported in Table 2.8 show that having capabilities in related technologies is important when entering into a new technological domain, as reflected by the positive and significant coefficient of the density variable in all the specifications of the model. The above finding confirms what we obtained in the baseline

[8]See https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

model: the likelihood that a new technological capability will emerge in a country is higher the closer a technology is to its existing capabilities. And this latter effect is stronger in developing economies (i.e lower for High-Income countries)[9]. Similarly, we confirm that is less likely to find diversification into more complex and valuable technologies, as shown by the negative coefficients of the ITC and Value Added variables.

Column 3 summarizes the results of the specialization equation. As in the baseline model, we find a significant and positive reinforcement of having developed related capabilities, as captured by the density coefficient. Additionally, these findings confirm that High-Income countries tend to specialize into more complex, more valuable, and less concentrated technologies.

Overall, the above results confirm our previous findings; which point towards the fact that developed countries tend to show a distribution of their technological production biased towards the production of less concentrated, more valuable, and more complex technologies. Additionally, note that the coefficient of the sample selection test is not significant in all three specifications and that the log size variable is always positive and significant as expected.

[9]Note, however, that this is not true for the first column, in which we subset the sample keeping those cases where RTA<0.1 in the previous period. This is partly explained because we end up with a very limited and highly unbalanced panel for the High-Income group.

Table 2.7: *List of Countries by Income Levels*

Low-Middle Income Countries	High-Income Countries
Argentina	Australia
Belarus	Austria
Brazil	Bahamas
Bulgaria	Belgium
China	Canada
Colombia	Chile
Costa Rica	Croatia
Cuba	Czech Republic
Egypt	Denmark
India	Estonia
Jordan	Finland
Lebanon	France
Malaysia	Germany
Mexico	Greece
Morocco	Hong Kong
Pakistan	Hungary
Philippines	Iceland
Russia	Ireland
South Africa	Israel
Sri Lanka	Italy
Thailand	Japan
Turkey	Korea
Ukraine	Latvia
Venezuela	Lithuania
Vietnam	Luxembourg
	Netherlands
	New Zealand
	Norway
	Poland
	Portugal
	Slovakia
	Slovenia
	Spain
	Singapore
	Sweden
	Switzerland
	United Kingdom

Table 2.8: *Results of the Robustness Check*

	Diversification Eq.	*Diversification Eq.*	*Specialization Eq.*
	(1)	(2)	(3)
Density	0.21***	0.338***	0.87***
	(0.024)	(0.037)	(0.037)
Density * GDP	-0.038	-0.095**	
	(0.03)	(0.0412)	
Tech-Level Variables			
Value Added	-0.0008***	-0.0007***	-0.0005*
	(0.00021)	(0.00023)	(0.00027)
Log Size	0.014**	0.026***	0.028***
	(0.0075)	(0.0061)	(0.005)
Herfindahl Index	-0.153***	-0.208***	-0.232***
	(0.041)	(0.036)	(0.046)
ITC	-0.026***	-0.024***	-0.026***
	(0.0048)	(0.0049)	(0.004)
Value Added * GDP			0.00017*
			(0.00009)
Log Size * GDP			-0.00064
			(0.006)
Herfindahl Index * GDP			-0.186**
			(0.0611)
ITC * GDP			0.022**
			(0.00029)
Sample Selection Test	0.008(0.0153)	0.0073 (0.01)	0.014(0.018)
Adjusted R-Squared	0.081	0.081	0.27
Tech FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
Observations	54,370	58,808	89,440

*p<0.1; **p<0.05; ***p<0.01

2.6.3 Technologies Life Cycles and Barriers to Entry

As mentioned in section four, we show here that independently of their proximity to existing competences, complexity, and value, technologies may have different barriers to entry, which are captured in our econometric regressions by technology fixed effects, as illustrated in Figure 2.4 below.

Figure 2.4: *Likelihood of Diversification Across Technologies*

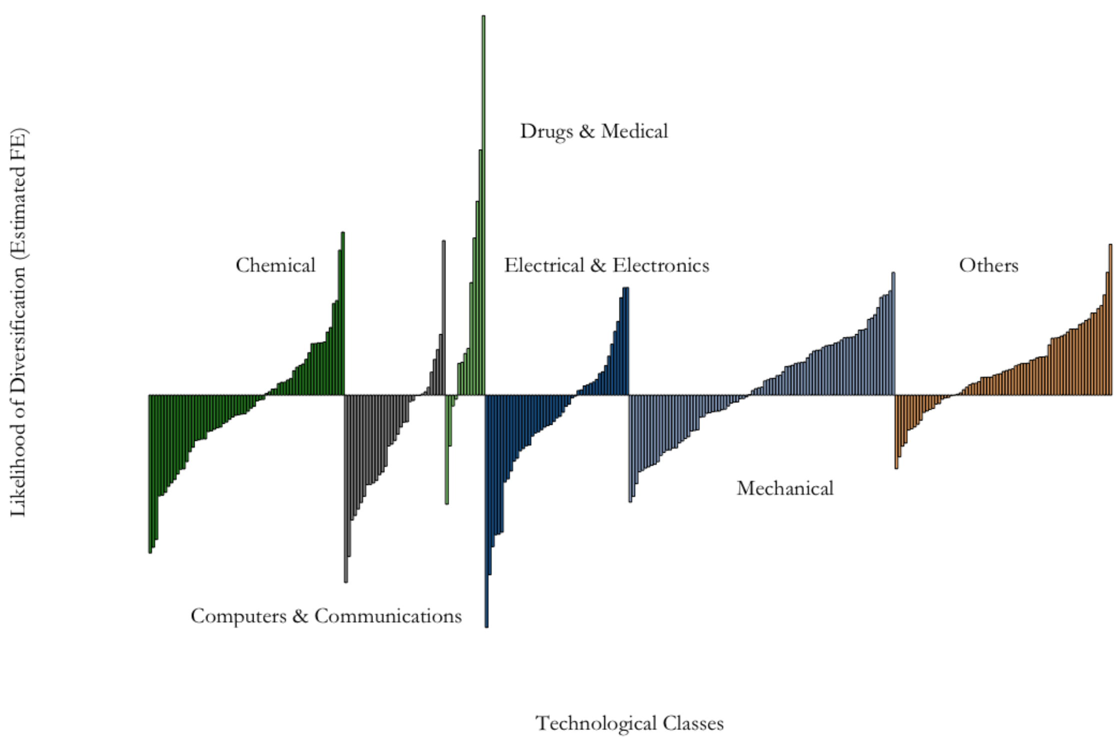


Figure 2.4 plots the coefficient estimates of the technology fixed effects across different technological sectors. All technologies are grouped according to their main economic classification (Chemical, Computers and Communications, Drugs and Medical, Electrical and Electronic, Mechanical, and Others) and are ordered

by their estimated coefficient. Higher values indicate a higher likelihood of diversification in that particular technology (or lower barriers to entry). All values are centered with respect to the average. This graph clearly portrays how disparate technologies are in terms of entry barriers, both within and between categories.

Moreover, we show in Table 2.9 below that these coefficient estimates strongly correlate with the determinants of entry used in Park and Lee (2006)[10]. We replicate the model of Park and Lee (2006) (Table 5, first column), but differently from them, in our econometric specification the dependent variable is given by the fixed-effect estimated coefficients, which proxy the likelihood of technological diversification (i.e. the inverse of barriers to entry). We find that all main variables keep the same direction, including the variable called CYCLETIME, which means that technologies with a shorter-life cycle evidence lower barriers to entry.

[10]Tech-level variables related to different characteristics of the technological regimes can be found at: http://www.keunlee.com

Table 2.9: *Replication of Park and Lee (2006)*

Variables	Likelihood of Entry
OPPORTUN	-0.004
	(0.019)
CUMUL1	-0.257***
	(0.018)
APPRO	0.066**
	(0.032)
INITIAL	0.067***
	(0.009)
CYCLETIME	-0.092***
	(0.007)
Adjusted R-Squared	0.82
Observations	315

*p$<$0.1; **p$<$0.05; ***p$<$0.01

CHAPTER 3

Unveiling the geography of historical patents in the United States from 1836 to 1975

*This article has been produced in collaboration with P.A. Balland and D. Rigby. The PhD candidate is the first author of the article, which has been published in 2016 as "**Unveiling the geography of historical patents in the United States from 1836 to 1975. Scientific data, 3. Nature Publishers.**". We acknowledge financial support from the Urban and Regional Research Centre Utrecht (Department of Human Geography & Planning, Utrecht University) and from Andrea Morrison through his VIDI project number 452-11-013 (Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), Innovational Research Incentives Scheme). We also thank Mathieu Steijn for his help in improving the statistical model to filter out correct locations.*

3.1 Background & Summary

The long-run development of societies depends on the rate at which they innovate. Innovation not only defines opportunities for economic progress but also determines the way that knowledge itself is produced. Invention is increasingly collaborative, generated overwhelmingly within the dense agglomerations of individuals and firms that comprise the world's major urban areas. Innovative cities are at the top of the global value chain, they are characterized by relatively high income per capita and by continuous improvements in average living standards. Clearly, not all residents of the largest cities benefit in the same way from invention, just as not all cities, regions and nations are as inventive as others. At the broadest scales, differences in rates of knowledge production over space and time are linked to geographical factors and to institutions that shape the character of economic and political fortunes. Still, we know relatively little about why particular technologies were developed in some places rather than others, about why specific cities boomed on the backs of some ideas, while other places with competing innovations languished. How do chains of technologies emerge over time building industries and regions in different places while destroying older regimes? And, in a new era of rapid information flow, are the old canons of uneven historical development likely to be discarded or merely revised?

At this time, few options exist for scholars seeking to analyse historical data linking the types of technologies invented to their place of invention. The primary source of information on the geography of knowledge production is the patent document. A patent provides exclusive intellectual property rights on an invention to its inventor (or assignee). In this way patents encourage the development of ideas. More precisely, the USPTO defines a patent as, "... the right to exclude others from making, using, offering for sale, selling or importing the invention". In exchange for such rights, the inventor (or assignee) is requested to provide detailed public disclosure of the patented invention. Public disclosure was designed to spur the diffusion of new ideas. Disclosure has also been key for academic researchers, providing a wealth of information on the business of science. By way of example, Figure 3.1 shows the first page of the Cohen-Boyer rDNA patent that gave birth to the biotechnology industry. Like all patents, this document contains systematic information about the invention, the grant date, the name of the inventor(s) and their home address(es), the name of the assignee and its business address, the date of application, the technological domains to which the patent applies, reference to prior academic publications and other patent documents on which the invention builds, and a brief abstract of the invention. This information is regularly used in economics, geography, and science and technology studies.

Figure 3.1: *The Cohen-Boyer rDNA patent*

United States Patent [19] [11] **4,237,224**

Cohen et al. [45] **Dec. 2, 1980**

[54] **PROCESS FOR PRODUCING BIOLOGICALLY FUNCTIONAL MOLECULAR CHIMERAS**

[75] Inventors: **Stanley N. Cohen,** Portola Valley; **Herbert W. Boyer,** Mill Valley, both of Calif.

[73] Assignee: **Board of Trustees of the Leland Stanford Jr. University,** Stanford, Calif.

[21] Appl. No.: **1,021**

[22] Filed: **Jan. 4, 1979**

Related U.S. Application Data

[63] Continuation-in-part of Ser. No. 959,288, Nov. 9, 1978, which is a continuation-in-part of Ser. No. 687,430, May 17, 1976, abandoned, which is a continuation-in-part of Ser. No. 520,691, Nov. 4, 1974.

[51] **Int. Cl.**^{3} **C12P 21/00**
[52] **U.S. Cl.** .... **435/68;** 435/172; 435/231; 435/183; 435/317; 435/849; 435/820; 435/91; 435/207; 260/112.5 S; 260/27R; 435/212
[58] **Field of Search** .... 195/1, 28 N, 28 R, 112, 195/78, 79; 435/68, 172, 231, 183

[56] **References Cited**

U.S. PATENT DOCUMENTS

3,813,316 5/1974 Chakrabarty .... 195/28 R

OTHER PUBLICATIONS

Morrow et al., Proc. Nat. Acad. Sci. USA, vol. 69, pp. 3365–3369, Nov. 1972.
Morrow et al., Proc. Nat. Acad. Sci. USA, vol. 71, pp. 1743–1747, May 1974.
Hershfield et al., Proc. Nat. Acad. Sci. USA, vol. 71, pp. 3455 et seq. (1974).
Jackson et al., Proc. Nat. Acad. Sci. USA, vol. 69, pp. 2904–2909, Oct. 1972.
Mertz et al., Proc. Nat. Acad. Sci. USA, vol. 69, pp. 3370–3374, Nov. 1972.
Cohen, et al., Proc. Nat. Acad. Sci. USA, vol. 70, pp. 1293–1297, May 1973.
Cohen et al., Proc. Nat. Acad. Sci. USA, vol. 70, pp. 3240–3244, Nov. 1973.
Chang et al., Proc. Nat. Acad. Sci, USA, vol. 71, pp. 1030–1034, Apr. 1974.
Ullrich et al., Science vol. 196, pp. 1313–1319, Jun. 1977.
Singer et al., Science vol. 181, p. 1114 (1973).
Itakura et al., Science vol. 198, pp. 1056–1063 Dec. 1977.
Komaroff et al., Proc. Nat. Acad. Sci. USA, vol. 75, pp. 3727–3731, Aug. 1978.
Chemical and Engineering News, p. 4, May 30, 1977.
Chemical and Engineering News, p. 6, Sep. 11, 1978.

Primary Examiner—Alvin E. Tanenholtz
Attorney, Agent, or Firm—Bertram I. Rowland

[57] **ABSTRACT**

Method and compositions are provided for replication and expression of exogenous genes in microorganisms. Plasmids or virus DNA are cleaved to provide linear DNA having ligatable termini to which is inserted a gene having complementary termini, to provide a biologically functional replicon with a desired phenotypical property. The replicon is inserted into a microorganism cell by transformation. Isolation of the transformants provides cells for replication and expression of the DNA molecules present in the modified plasmid. The method provides a convenient and efficient way to introduce genetic capability into microorganisms for the production of nucleic acids and proteins, such as medically or commercially useful enzymes, which may have direct usefulness, or may find expression in the production of drugs, such as hormones, antibiotics, or the like, fixation of nitrogen, fermentation, utilization of specific feedstocks, or the like.

14 Claims, No Drawings

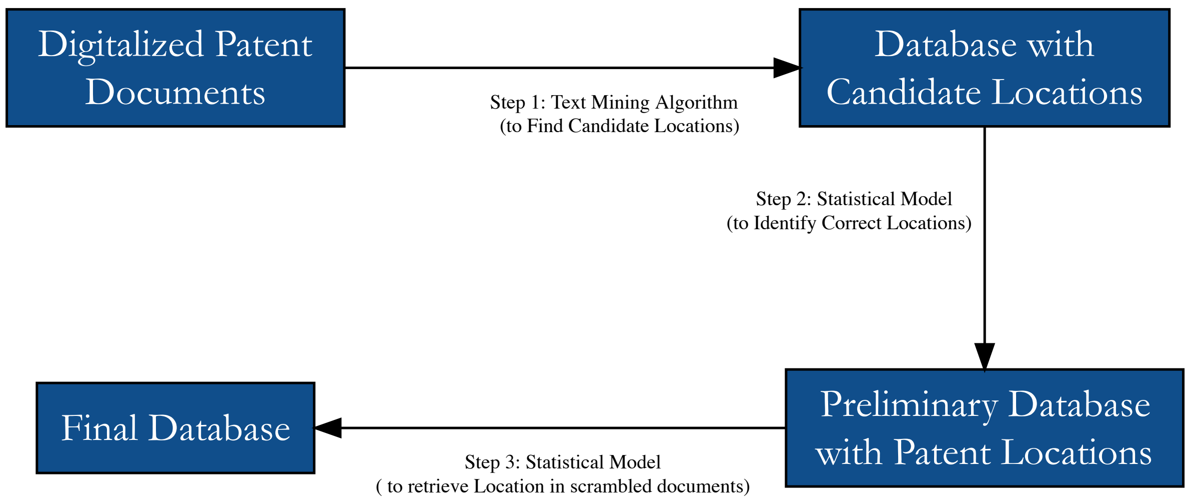
Original image of the front page of the Cohen-Boyer rDNA patent granted by the USPTO in 1980. The front page shows the different type of systematic information that a patent document contains, such as the inventors home addresses, the technological fields, and the references to prior art.

Although patent data are freely available from the USPTO Patent Full-Text and Image Database, they are not always available in a format that can be directly used for applied research. For some research questions, the raw data first have to be cleaned and processed (location disambiguation, or inventor/assignee name disambiguation for instance). A few structured, geo-referenced datasets have been developed over the past couple of years. One of the most commonly used is the patent dataset of the National Bureau of Economic Research (NBER), providing information on the state of first inventor for patents from 1975 to 1999. Another widely used database for US patents is the Patent Network Dataverse, providing longitude and latitude coordinates of inventor addresses for patents granted by the USPTO from 1975 to 2010. In a similar fashion, the REGPAT dataset of the Organisation for Economic Co-operation and Development (OECD) provides inventor locations (NUTS3 level for Europe, TL3 for other OECD countries) for patents filed to the European Patent Office (EPO) or to the World Intellectual Property Organization (WIPO) from 1978 to 2011 (OECD, 2015).

However, these datasets only provide detailed geographical information on patents granted since 1975, the year when the USPTO began to record patents electronically. The main objective of this chapter is to present HistPat, a well-structured, ready-to-use, comprehensive, and geo-referenced dataset of historical patents in the United States covering the years 1836 to 1975. HistPat contains geographical information (at the county level) of approximately 2.8 million patent documents (around 83% of all patents granted to US residents).

HistPat is built using optically recognized patents documents made available by Reed Tech and Google. We develop a methodological procedure to retrieve geographical information from those patent documents that can be divided into three steps (see Figure 3.2). First, we use a standard text-mining algorithm to find potential locations within these patent documents. Second, we propose and calibrate a statistical model to identify correct locations from all possible candidate locations. Third, we exploit data from related patents for scrambled documents. HistPat is a valuable database that should be of interest to researchers in disciplines such as geography, economics, history, network science, and science and technology studies.

Figure 3.2: *Workflow Summary*



Data collection workflow in three main steps: (1) find potential locations within patent documents, (2) identify correct locations from all possible candidate locations, and (3) retrieve the geographical location in scrambled patent documents.

3.2 Methods

This section describes the methodological procedure used to obtain the location of inventors and/or assignees from optically recognized patents documents (plain text documents). It is divided into three steps, as described in Figure 3.2 above.

3.2.1 Step 1: Finding candidate locations within patent documents

The final database on the geography of historical U.S. patents - HistPat - was built using bulk data from the United States Patent and Trademark Office. In 2006, the USPTO entered into a series of agreements with Reed Tech and Google to digitalize all available patent documents, making historical patent data available in bulk form. This bulk data contains ZIP or TAR files with TIFF or PDF images, concatenated XML or structured ASCII files, and can be accessed at: http://www.uspto.gov/learning-and-resources/electronic-bulk-data-products.

The dataset presented in this chapter has been constructed using this data, covering a period ranging from 1836 to 1975. Even though the first patent dates back to 1790, coverage between 1790 and 1836 is scattered and not entirely reliable. This is because a fire at the USPTO destroyed the records of thousands of granted patents and pending applications in 1836. Individual patents can also be accessed without using bulk data through the "Google Patents" search engine: https://patents.google.com. In this subsection we outline a procedure to create a database of 'candidate' geographical locations from the digitalized patent documents. This database will later be evaluated to assess the likelihood that a 'candidate' location is the actual location of an inventor or an assignee. This procedure is divided into two stages. We first identify all possible candidate locations within patent documents. Second, we generate a set of variables providing information about those locations such as their proximity to inventors' names, their position within the patent document, and other features. Identification of candidate locations depends upon access to a comprehensive list of town, city, and county names within the United States. We use two sources for this task. The first is provided by the U.S. Census Bureau at https://www.census.gov/geo/reference/codes/place.html. The second is the online gazetteer provided by Falling Rain Genomics, Inc and available at http://www.fallingrain.com/world/US/. The gazetteer is used to supplement neighbourhood names that are sometimes missing in census data. Historical US patent documents reference the addresses of inventors and assignees by naming the town, county and state where individuals and/or firms were located. Armed with a list of place names within the U.S., standard text detection algorithms can be used to detect the presence of these names within patent documents. Fast and reliable packages for text mining algorithms can be found in R software. We use the stringr (version 1.0.0), stringi (version 1.0.1), and tm (version 0.6.1) packages. Once we have a list of candidate locations we evaluate them in the context of the patent document, generating a set of covariates for each of them. Figure 3.3 describes a typical input in this first step. Note that for any location name, an entire set of potential candidate locations may be generated, as there are many places with the same name in different states.

Table 3.1 shows the typical output of this first step. The length of the document (in number of characters) is captured by the variable named 'Length'. This variable remains constant within the document and it is best used in combination with other variables to standardize values and ease comparison across documents. The variable 'Location' identifies where in the document the name of the location was first found (as there may be multiple mentions of the same name). Additionally, other terms may be used (in combination with the location name) to create variables providing valuable information. For instance, the variable 'State' gives a value of 1 if the state name of the candidate location was found in the document. An additional variable measuring proximity between the name of the candidate location and the state name (if found), or inventor(s) name, also proved useful.

Figure 3.3: *Input to step 1*

G. C. **FRENCH**.
PLUE CLEANER.
APPLIOATION FILED OCT. 18, 1911.
1,051,323, Patented Jan.21, 1913.
. INVENTOR 62076 af'rencfi ATTORNEYS **GEORGE** C. **FRENCH**, OF **ELDORA**, **IOWA**.
FLUE-CLEANER.
Specification of Letters Patent.
Patented Jan. 21, 1913.
Application filed October 18, 1911. Serial No. 655,255.
To all whom it may concern:
Be it known that I, **GEORGE** C. **FRENCH**, a citizen of the United States, and a resident of **Eldora**, in the county of **Hardin** and State of **Iowa**, have invented new and useful Improvements in Flue-Cleaners, of which the following is a full, clear, and exact description.
This invention relates to flue cleaners for freeing the flues of boilers and other apparatus from accumulations of soot, scale and the like, and has reference more particularly to a flue cleaner which comprises an elongated flexible body having an adjustable scraper mounted thereon.

.
.
.

Having thus described my invention, I claim as new, and desire to secure by Letters Patent:
1. c In a fine cleaner in combination, a body formed of two parts having a loose joint permitting the angular relative movement of said parts, a flexible, helical scraper, and adjustable means on said respective parts of said body, to secure said scraper to said body.
2. A flue cleaner comprising a body consisting of two parts, the first of said parts having a reduced extension and an openended socket member mounted upon said extension, a follower mounted upon said re duced extension and held in place by said socket member, the second of said parts having a head received in said socket member, and being movably held in position relative to said other part thereby, a helical scraper having one end secured to said follower and having the other end mounted upon the second of said parts, and means for adjustably securing said scraper relative to said parts.
3. A flue cleaner comprising a body consisting of a member having a reduced threaded part, a socket nut screwed upon said threaded part, and a bolt rod having a head located within said'socket nut, said member being provided with a shoulder, a follower positioned upon said member between said shoulder and said socket nut, and a helical scraper having one end adjustably secured to said bolt rod and having the other end secured to said follower.
4. A flue cleaner comprising a body consisting of a member having a reduced, threaded part, a socket nut screwed upon said threaded part, and a bolt rod having a head located within said socket nut, said member being provided with a shoulder, a follower positioned upon said member, between said shoulder and said socket nut, and consisting of a rounded member, a conical washer clamped between said follower and said socket nut, said bolt rod having a threaded portion, a helical scraper having an eye positioned upon said threaded part of said bolt, and having an end clamped between said follower and said washer, and nuts upon said threaded part. of said bolt, for securing said eye of said scraper adjustably in position.
In testimony whereof I have signed my name to this specification in the presence of two subscribing witnesses.
GEORGE C. **FRENCH**.
Witnesses:
HOWARD E. Morrn'r'r, **MATTHEW** D. **KENNEDY**.
Copies of this patent may be obtained for five cents each, by addressing the Commissioner of Patents,
' **Washington**, D. 0.

Example of an input in step 1. In this case, OCR text related to a patent on flue-cleaners by George C. French, a resident of Eldora, in the county of Hardin (Iowa).

Table 3.1: *Output from step 1*

Variable	Type	Comment
Min. Location	Integer	Candidate Location (CL) first appearance in the document.
Street	Dummy	1 if CL is located to the right of the words 'ST', 'AVENUE', 'ROAD', 'RD', 'BLVD', or 'AVE'. 0 otherwise.
City	Dummy	1 if 'CITY' is part of the CL name.
Frequency	Integer	Number of times the CL was found.
State	Dummy	1 if the name of the corresponding state was found.
Min. Location State	Integer	State first appearance in the document.
State Distance	Set of dummy variables	Dummies corresponding to intervals of character distances between CL and the state name.
Countries	Dummy	1 if the name of a western european country appears in the document.
Country Distance	Dummy	1 if a country name as specified above is found close to the CL. 0 otherwise.
Cutoff	Dummy	1 if the CL or the state has been found after the 50% of the document length. 0 otherwise.
Substring	Dummy	1 if the CL is a substring of another CL within the same patent. 0 otherwise. (i.e. York for New York)
Nchar	Integer	Number of characters of the CL.
Detected Name	Dummy	1 if the CL matches any part of the inventor or assignee name. 0 otherwise.
W State	Dummy	1 if at least 1 state name has been found for other CL within the same patent document. 0 otherwise.
Rel. Min. Location	Countinuous	Min Location over the length of the document. Varies between 0 and 1.
City	Dummy	1 if the CL was found next to the word 'CITY'. 0 otherwise.
County	Dummy	1 if the CL was found next to the word 'COUNTY'. 0 otherwise.
COC	Dummy	1 if more than one CL of the same county ID co-occur within the same patent document. 0 otherwise.
WX	Continuous	Index constructed with all the aforementioned variables for 'competing' CL within the same patent.

Constructing a set of variables for each potential location is crucial, as we will use them to evaluate the likelihood that each candidate location is the true location of the patent. Step 2 outlines a statistical procedure to filter out the correct locations from all available possibilities. Table 3.1 provides a detailed description of the variables we constructed.

3.2.2 Step 2: Filtering correct locations

The objective of this subsection is to discuss the design of a statistical model that allocates probabilities to candidate locations that signify their likelihood of being the real location of a patent. These probabilities are generated by using the observed attributes of each location (see Table 3.2).

We do this by training and evaluating the predictive performance of three popular and well-studied statistical procedures (Neural Networks (NN), K-th Nearest Neighbours (KNN), and a Probit model). For training purposes we use a manually collected sample that identifies the correct locations for a randomly selected subset of patent documents. We inputted manually the correct location for approximately 7000 patent documents, which were selected randomly from all available patents covering the period 1836 to 1975. More specifically, let the output or response variable of the statistical model take two possible values from the finite set $Y = 0, 1$; where the category '1' identifies correct locations within patents. Let $X = X_{ij} = (X_{ij}^1, X_{ij}^2, ..., X_{ij}^p)$ be a vector of p predictors (or attributes) for location i within patent document j. If we treat Y as a quantitative output we allow predictions of Y (denoted $\widehat{Y}$) to fall within the interval $[0, 1]$. Additionally, assume there exists a set of measurements (x_{ij}, y_{ij}) for a randomly selected subset of patents $j = 1, ..., N$ that we will call training set. Statistical decision theory provides a framework to evaluate problems of this sort. Within this framework we aim at finding a function $f(.)$ to predict Y_{ij} given X_{ij}. This framework requires specifying a loss function $L(Y, f(X))$ that penalizes errors in prediction. We seek to find an

Table 3.2: *Attributes of each candidate location*

Publication Number	City	State	Length	Location	State Located
US1051323	ELDORA	CO	7482	150	0
US1051323	ELDORA	FL	7482	150	0
US1051323	ELDORA	IA	7482	150	1
US1051323	ELDORA	NJ	7482	150	0
US1051323	ELDORA	PA	7482	150	0
US1051323	ELDORA	WV	7482	150	0
US1051323	FEBRUARY	TN	7482	2211	0
US1051323	FRENCH	AR	7482	7	0
US1051323	FRENCH	ID	7482	7	0
US1051323	FRENCH	MN	7482	7	0
US1051323	FRENCH	NM	7482	7	0
US1051323	FRENCH	VA	7482	7	0
US1051323	GEORGE	IA	7482	129	1
US1051323	GEORGE	KS	7482	129	0
US1051323	GEORGE	MO	7482	129	0
US1051323	GEORGE	MS	7482	129	0
US1051323	GEORGE	MT	7482	129	0
US1051323	GEORGE	NC	7482	129	0
US1051323	GEORGE	OR	7482	129	0
US1051323	GEORGE	TX	7482	129	0
US1051323	GEORGE	WA	7482	129	1
US1051323	GEORGE	AR	7482	129	0
US1051323	HARDIN	AR	7482	426	0
US1051323	HARDIN	CA	7482	426	0
US1051323	HARDIN	CO	7482	426	0
US1051323	HARDIN	GA	7482	426	0
US1051323	HARDIN	IA	7482	426	1
US1051323	HARDIN	IL	7482	426	0
US1051323	HARDIN	KY	7482	426	0
US1051323	HARDIN	MO	7482	426	0
US1051323	HARDIN	MT	7482	426	0
US1051323	HARDIN	NC	7482	426	0
US1051323	HARDIN	OH	7482	426	0
US1051323	HARDIN	OR	7482	426	0
US1051323	HARDIN	TN	7482	426	0
US1051323	HARDIN	TX	7482	426	0
US1051323	IOWA	IA	7482	158	1
US1051323	IOWA	LA	7482	158	0
US1051323	IOWA	PA	7482	158	0
US1051323	IOWA	WI	7482	158	0
US1051323	KENNEDY	AL	7482	7355	0
US1051323	KENNEDY	CA	7482	7355	0
US1051323	KENNEDY	GA	7482	7355	0
US1051323	KENNEDY	IA	7482	7355	1
US1051323	KENNEDY	IL	7482	7355	0
US1051323	KENNEDY	IN	7482	7355	0
US1051323	KENNEDY	KY	7482	7355	0
US1051323	KENNEDY	MN	7482	7355	0
US1051323	KENNEDY	MO	7482	7355	0
US1051323	KENNEDY	NE	7482	7355	0
US1051323	KENNEDY	NM	7482	7355	0
US1051323	KENNEDY	NY	7482	7355	0
US1051323	KENNEDY	PA	7482	7355	0
US1051323	KENNEDY	SD	7482	7355	0
US1051323	KENNEDY	WA	7482	7355	1
US1051323	KENNEDY	WI	7482	7355	0
US1051323	MATTHEW	KY	7482	7344	0
US1051323	MATTHEW	NC	7482	7344	0
US1051323	MATTHEW	TN	7482	7344	0

approximation $\widehat{f(X)}$ to the relationship between the predictors and the output Y. Probit and NN models can be grouped within the class of Projection Pursuit Regressions where $f(X)$ can be defined as:

$$f(X) = \sum_{m=1}^{M} g_m(\omega_m^T X) \tag{3.1}$$

with a loss function of the form

$$L(Y, g_m(.), \omega_m^T, X) = \sum_{i=1}^{n} [y_i - \sum_{m=1}^{M} g_m(\omega_m^T X)]^2 \tag{3.2}$$

Our aim is to approximate the parameters of this model by minimizing the loss function. What differentiates Probit and NN models are the assumptions over the parameters. If we let $M = 1$ and assume $g_m = g$ to be the Cumulative Distribution Function (CDF) of the standard normal distribution we get a Probit model. What differentiates NN models is that they use linear combinations of the predictors to construct a set of indexes Z_m that are combined in linear form to estimate $\widehat{Y}$. Thus,

$$\begin{aligned} &Z_m = \sigma(\alpha_{0m} + \alpha_m^T X_m), m = 1, ..., M \\ &T = \beta_0 + \sum_{m=1}^{M} \beta_m^T Z_m \\ &f(X) = g(T) \end{aligned} \tag{3.3}$$

where the activation function σ and the output function g could be chosen to be the logistic function. Note that the NN model proposed here can be understood as logistic regression using Z_m as covariates. The intermediate inputs Z_m are called hidden units because their values are not observed directly. The last statistical model we implement does not require any statistical fitting. The KNN model consists of finding for any given point x_0, the K-th nearest neighbours within a set of training points (X, Y); to later classify x_0 using a decision rule based on the information provided by the K-th nearest neighbours. We use the Minkowski distance metric to find the nearest neighbours. Predictors are standardized beforehand, we set $\theta = 2$.

$$Distance(x_0, X_{ij}) = (\sum_{r=1}^{p} |x_0 - x_{ij}|^{\theta})^{\frac{1}{p}}, \quad with \quad \theta > 1 \tag{3.4}$$

After the K-th nearest neighbours are found, x_0 is classified implementing a decision rule over all output values within the neighbourhood of x_0 (i.e. all $y_{ij} \in N_{x_0}$). We use the Epanechnikov kernel function to weight neighbours according to their distances and predict the value of y_0 as a weighted average of all $y_{ij} \in N_{x_0}$. Predictions of these procedures ($\widehat{Y}$) will lie in in the interval $[0, 1]$. We can then classify each location within the groups $G = [Correct, Incorrect]$ according to the following rule:

$$\begin{aligned} \widehat{G} &= Correct \quad if \quad \widehat{Y} > \mu \\ \widehat{G} &= Incorrect \quad if \quad \widehat{Y} \leqq \mu \end{aligned} \tag{3.5}$$

Where μ is a threshold parameter that falls in the interval $[0, 1]$, used to discriminate correct from incorrect locations. This might be interpreted as a threshold likelihood that potential locations should pass to be considered as real locations in our database. This classification rule will typically be subject to misclassification error. However, as μ increases, the probability of misclassifying an incorrect location should decrease. Table 3.3 shows a typical output. Note that

Table 3.3: *Output of the filtering process*

Publication Number	City	State	Length	Location	State Located	$\widehat{Y}$
US1051323	ELDORA	IA	7482	150	1	1
US1051323	HARDIN	IA	7482	426	1	0.99

the table includes a new variable with the value of $\widehat{Y}$ for each location. In this example we only keep the locations predicted as true with likelihood above 50%.

Note that the three statistical models proposed in this chapter can be clearly ordered in terms of their parametric constraints. The Probit model, being the most restrictive of all, has the advantage of speed as the number of parameters to be learned from the data is lower. Commonly used searching algorithms, such as iteratively reweighted least squares (IRLS), can be used to choose the parameters that minimize the loss function. An additional feature is that we are able to provide an interpretation of the effect of our predictors on the output. It is often the case, however, that NN and KNN models outperform Parametric Single Index Models (PSIM) in terms of predictive power (Friedman et al., 2001; Jeffrey, 2002). As prediction is the main objective, the simplicity and interpretability of PSIM may impose constraints we don't want or need. NN models are more flexible and have been proven to approximate nonlinear relationships relatively well. They tend to outperform PSIM in most empirical applications (Friedman et al., 2001; Jeffrey, 2002). There are, however, some shortcomings. First of all, it is usually difficult to interpret the effect of predictors as they are masked within the hidden units. Additionally, NN models tend to have a considerable number of weights to be estimated, often leading to the risk of overfitting the data if parameters and optimization procedures are not chosen appropriately. We use the so-called resilient back-propagation algorithm to minimize the loss function (Anastasiadis et al., 2005; Günther and Fritsch, 2010; Riedmiller and Braun, 1993) This modifies weights after calculating the gradient of the error function until a local minimum is reached. An appealing feature of this procedure is that different learning rates can be assigned to different weights that make the procedure more robust when compared to traditional back-propagation algorithms. Being completely non parametric, KNN models tend to impose an even higher computational burden. Note that they usually require finding the neighbours and storing the entire training set to be matched against query points. In our case it requires $N * p$ operations per x_0. However, KNN models have proven successful in a variety of classification problems, especially when decision boundaries are very irregular (Friedman et al., 2001; LeCun et al., 1990; Tibshirani et al., 2003) An appealing feature is that KNN models are unstructured and don't impose any particular parametric restriction, nor do they require any model to be specified. As in the case of NN models, they are not useful for understanding the relationship between the predictors and the outcome and may be unstable under some circumstances. As our final goal is to correctly predict as many locations as possible while minimizing errors, the final decision over competing alternatives will be entirely based on predictive performance. The inclusion of these three particular competing alternatives is based on the wide variety of scenarios they could accommodate. The idea is that other researchers wanting to expand or improve this database could have a set of flexible tools at their disposal. Note that for the particular problem at hand we have an important advantage over the usual predictive endeavours for we can see the Data Generating Process (DGP) via the patent documents themselves. Moreover, this DGP barely changes over time. This means that we can create attributes of locations knowing beforehand whether and how they will work. By way of an example, let us say one is interested in tracking down the emergence of new technologies or chemical components by searching for references to those technologies (i.e. internal combustion, polyethylene, etc.). In principle, the same exact procedure could be applied to the set of available documents, replacing location names by these keywords. If these keywords appear in any part of the document, evaluation of the appropriateness of located terms may be more difficult. If decision boundaries are more irregular, less restrictive approaches may be preferable.

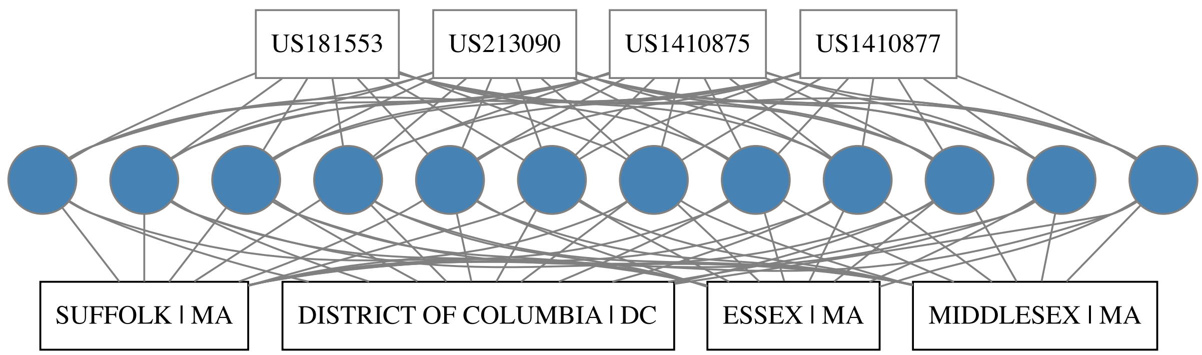
3.2.3 Step 3: Including location for 'unreadable' patent documents

At this stage we have a preliminary database with the correct locations (or predicted correct) for around 2.65 million patents documents. Even though the Optical Character Recognition (OCR) software succeeded in providing an accurate and detailed digitalized description of most patent documents, some of them still remained 'unreadable' (or 'machine-unreadable' to be fair). This means the OCR software was unable to recognize scrambled, broken and unconnected characters or symbols for some documents. As a result, locations could not be retrieved for those patents. It is possible, however, to make use of the bibliographic information on patents to infer a location for those 'unreadable' documents. This can be done by evaluating other patents of the same inventors and/or assignees. The idea is to evaluate, for all 'unreadable' patent documents, a set of potential locations using the predicted locations in step 2 and the fact that we often have bibliographical information related to the patent[1].

In this step, we create a set of 'potential' locations for every 'unreadable' patent whenever the same inventor or assignee has another patent with an identified location (retrieved in step 2). As in step 2, we create a set of attributes for 'potential' locations that will be related to the number of times that location was found in other patents under the same inventor/assignee name, the ubiquity of the inventor/assignee name, etc. Figure 3.4 below summarizes this procedure. It describes how to construct a database of possible locations for 'unreadable' patent

[1] See for instance the following example of a scrambled document with an 'unreadable' location: https://www.google.com/patents/US6469.
Available bibliographical information can be found here:
http://worldwide.espacenet.com/publicationDetails/biblio?CC=US&NR=6469.

Figure 3.4: *Retrieving possible locations of 'unreadable' patent documents*



This figure describes how to find location for unreadable patent documents. We use here a set of patents invented by Alexander Graham Bell as an example. Patent numbers at the top of the figure correspond to documents for which a location couldn't be found, while blue coloured dots represent patent documents with an identified location from step 2. The division between unreadable and readable patents is only for illustrative purposes. All of these patents contain an assigned location from step 2. Lines connecting unreadable patents with blue dots mean those patents share the same inventor name (i.e. Alexander Graham Bell). We can connect those blue dots with found locations to create a geo-referenced database of unreadable OCR text.

documents using the bibliographical information about the inventor/assignee name.

In this example there are multiple locations to consider because the inventor lived in different places. Table 3.4 shows how we structure the information displayed in Figure 3.4. The variable named 'Frequency' counts how many times that location appeared. Note that this variable does not vary across patents, as it is a characteristic associated with the inventor name rather than the specific patent document. However, we can make use of the bibliographical information to include variables that capture the heterogeneity across 'unreadable' patents by identifying, for instance, how many of the blue-coloured patents were filled in the same year. The variable 'Year Coinc.' counts how many of the positive matches in 'Frequency' correspond to the same year of the 'unreadable' patent. This variable provides valuable information to disambiguate among locations when the inventor has moved during the period, as in this case (the correct location for patents US1410877, US1410875, and US213090 is Washington DC, and Boston MA for

Table 3.4: *Input in step 3*

Inventor	PN	County	State	Year	Class	Frequency	Class Coinc.	Year Coinc.
A. G. Bell	US1410875	DC	DC	1922	114	7	1	1
A. G. Bell	US1410875	Essex	MA	1922	114	2	0	0
A. G. Bell	US1410875	Middelsex	MA	1922	114	1	0	0
A. G. Bell	US1410875	Suffolk	MA	1922	114	3	0	0
A. G. Bell	US1410877	DC	DC	1922	114	7	1	1
A. G. Bell	US1410877	Essex	MA	1922	114	2	0	0
A. G. Bell	US1410877	Middelsex	MA	1922	114	1	0	0
A. G. Bell	US1410877	Suffolk	MA	1922	114	3	0	0
A. G. Bell	US181553	DC	DC	1876	310	7	0	0
A. G. Bell	US181553	Essex	MA	1876	310	2	0	4
A. G. Bell	US181553	Middelsex	MA	1876	310	1	0	0
A. G. Bell	US181553	Suffolk	MA	1876	310	3	0	0
A. G. Bell	US213090	DC	DC	1879	381	7	3	1
A. G. Bell	US213090	DC	DC	1879	379	7	5	1
A. G. Bell	US213090	Essex	MA	1879	379	2	1	0
A. G. Bell	US213090	Essex	MA	1879	381	2	0	0
A. G. Bell	US213090	Middelsex	MA	1879	381	1	0	0
A. G. Bell	US213090	Middelsex	MA	1879	379	1	0	0
A. G. Bell	US213090	Suffolk	MA	1879	379	3	2	0
A. G. Bell	US213090	Suffolk	MA	1879	381	3	4	0

US181553). Information about technological classes may also help disambiguating among locations when the inventor name is very common, by considering also the area of expertise of the inventor.

Note that the information provided by the coincidence of class and year can be incorporated directly in the network of Figure 3.4. It is possible to create additional networks that link 'unreadable' patent documents only with patents of the same inventor and within the same technological class. Frequencies can be calculated for this sub-network. This procedure has the advantage of reducing the dimensionality of the network, which may be handy when the number of pairs to evaluate is very high, as in this case. We filter correct locations at this time in exactly the same way as in step 2 (above). We divide the manually collected sample into training and test sets to train the same three econometric procedures and evaluate them according to their predictive performance. Note that any errors already present in the preliminary database coming from step 2 may be carried to this stage. However, the threshold coefficient μ can be set arbitrarily to determine the desired error tolerance level. Locations predicted as 'correct' will be appended to the database of step 2. Table 3.5 lists all variables we constructed for this step.

Table 3.5: *Variables in step 3*

Variable	Type	Comment
State	Dummy	1 if the name of the corresponding state was found.
Name Match	Dummy	1 if any place name within potential counties from stage 2 is found.
Nmatch	Integer	Number of name matches in 'Name Match'.
Frequency PAL	Integer	Frequencies for locations obtained from the Patent to Assignee to Location Network (PAL).
Frequency PIL	Integer	Frequencies for locations obtained from the Patent to Inventor to Location Network (PIL).
Proportion PAL	Continuous	Proportion for locations obtained from the Patent to Assignee to Location Network (PAL).
Proportion PIL	Continuous	Proportion for locations obtained from the Patent to Inventor to Location Network (PIL).
Frequency PACL	Integer	Frequencies for locations obtained from the Patent to Assignee to Location to Class Network (PACL).
Frequency PICL	Integer	Frequencies for locations obtained from the Patent to Inventor to Location to Class Network (PICL).
Proportion PACL	Continuous	Proportion for locations obtained from the Patent to Assignee to Location to Class Network (PACL).
Proportion PICL	Continuous	Proportion for locations obtained from the Patent to Inventor to Location to Class Network (PICL).
Frequency PAYL	Integer	Frequencies for locations obtained from the Patent to Assignee to Location to Year Network (PAYL).
Frequency PIYL	Integer	Frequencies for locations obtained from the Patent to Inventor to Location to Year Network (PIYL).
Proportion PAYL	Continuous	Proportion for locations obtained from the Patent to Assignee to Location to Year Network (PAYL).
Proportion PIYL	Continuous	Proportion for locations obtained from the Patent to Inventor to Location to Year Network (PIYL).
Ubiquity	Integer	Ubiquity of the name of the inventor/assignee.
WX	Continuous	Index constructed with all the aforementioned variables for 'competing' CL within the same patent.

3.3 Code Availability

All procedures implemented in this project were written in R software (Version 3.3.1). We used text mining algorithms from the following packages: stringr (version 1.0.0), stringi (version 1.0.1), and tm (version 0.6.1). We provide a simplified example of the original code to facilitate the reproduction of the procedures described in this chapter, with access details provided in the Data Citation 1: Harvard Dataverse http://dx.doi.org/10.7910/DVN/BPC15W, under the name 'Replication Example'. The entire code is available upon request.

3.4 Data Records

The result of this procedure is a database that we will refer to as HistPat. HistPat and supporting data are archived at the Harvard Dataverse, Harvard University, with access details provided in the Data Citation 1: Harvard Dataverse http://dx.doi.org/10.7910/DVN/BPC15W. The 'HistPat.csv' file, within the folder named "HistPat Dataset", contains seven columns and 3,496,301 rows. Each row corresponds to a location in a patent document while the columns provide the following information:

Table 3.6: *Data Records*

Variable Name	Description
PN	Patent Document Publication Number as shown in patent documents
FIPS	County subdivision FIPS code as specified by the US Census Bureau (https://www.census.gov/geo/reference/codes/place.html)
State	State postal code as specified by the US Census Bureau (https://www.census.gov/geo/reference/codes/place.html)
County	County Name
Source	Identifies how the patent location was obtained. One of the following types:
	- MCS: Manually Collected Sample
	- MCU: Manually Contributed by Users
	- Step 2: Automatically inputted, corresponds to the second step described in this document
	- Step 3: Automatically inputted, corresponds to the third step described in this document
Alpha	Expected accuracy for automatically derived locations. A value of 5, 2.5, or 1 means that you should expect 5, 2.5, or 1 wrongly assigned locations every 100 patent documents, respectively.
Year	Year of publication (grant year)

The variable "PN" gives the patent publication number, as shown in patent documents. Users can search individual patents listed in HistPat by copying and pasting this patent document publication number in the Google Patent search engine: https://patents.google.com/patent/. For instance, the patent for the phonograph (PN = US200521), invented by Thomas Edison in 1878 can be found at this address: https://patents.google.com/patent/US200521.Users can also use this number to append HistPat to other existing datasets such as the NBER patent data (Hall et al., 2001a). In this case, the corresponding variable name is "patent", and only includes numeric values, i.e. "US200521" would be "200521".

3.5 Technical Validation

All three procedures have tuning parameters to be learned from the data or to be imposed exogenously. For instance, in the case of NN models, weights are learned from the data while the number of hidden units is usually set by the researcher. In Probit models coefficients are estimated, while the number of neighbours in KNN models is usually chosen exogenously. Results of this section are obtained using 25 neighbours for the KNN model, and allowing only one layer and 30 hidden units in the NN model. Results are robust to departures from these values.

In this section we test the performance of these three different alternatives. The main goal is to choose the best procedure in terms of predictive accuracy and coverage. A performance assessment over an independent test set is crucial for this sort of procedure as there is a risk that models will over-fit the training set. Over-fitting the training set occurs when parameters of the model are tuned in such a way that they become suitable only for that particular training set, without being able to generalize and correctly predict new data. In a first sub-section (3.5.1) we provide a comparison across procedures using training and test samples of equal size. We use these results to choose the 'best' performing procedure.

After having chosen the 'best' performing model we also test whether predicted locations evidence any sort of bias. Note that we are identifying locations by name matching, based on an imperfect OCR procedure. It may be the case that some particular locations are either more difficult to recognize or to evaluate properly. For instance, as the length of the location name increases, the likelihood of misspelling increases too. However, if a location with a long name is detected, the likelihood the model considers it correct increases. This may generate a bias towards correctly predicting some locations more often than others. We test this in the second subsection (3.5.2) by comparing the distribution of locations in our final database to the one collected manually, both across time and technological domains.

In both subsections (3.5.1) and (3.5.2) we only show the result of evaluating all three procedures for what we called the second step. This means that we only include the comparison across procedures for the case where we aim at predicting which candidate locations are correct. Remember that we apply a similar procedure also to identify the locations in scrambled documents (i.e. step three). We do not show the comparison for this later case because predicting performance of models is almost identical to the one obtained in step 2.

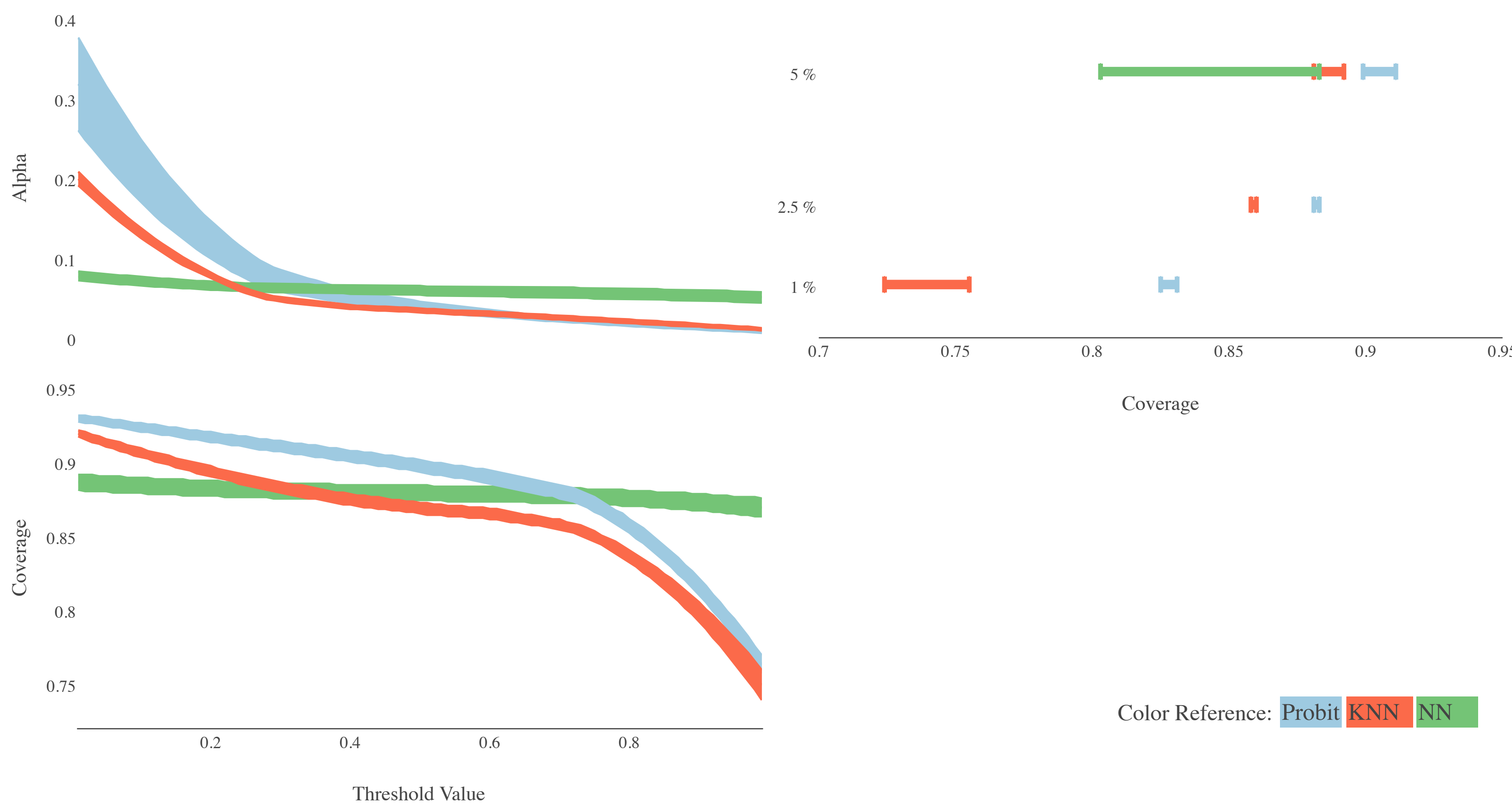
3.5.1 Predicting model performance

Figure 3.5 compares all three alternatives in terms of their predictive performance. We use the value μ to filter locations that are predicted to be true above the specified threshold. We evaluate, for any value of μ, how many errors are contained in the final sample (we call this parameter alpha) and what percentage of all patents within the sample contain at least one location (we call it coverage). Specifically, we compute alpha (for a particular value of μ) as the percentage of locations that were wrongly codified as 'correct' (i.e. they passed the filter even though they are not the correct location of the patent). The final coverage is computed as the proportion of patents that have at least one location.

Figure 3.5 should be read in a clock-wise direction from the bottom-left. Different colours were assigned to alternative procedures (blue to Probit, green to NN, and red to KNN). The first graph shows how coverage decreases as we increase the value of our threshold (μ). As expected, requiring higher predicted probabilities decreases the coverage of the final sample. However, as can be seen in the second (top-left) graph, this also decreases the probability that we will commit mistakes and include 'incorrect' locations in the final sample. These two graphs share the horizontal axis and show how the coverage and alpha change as we move the threshold value μ. The thickness of the line represents the 95% confidence interval (calculated repeating 100 times the procedure for different randomly selected training and test sets). Note that the Probit model (blue) always has a higher rate of coverage than the KNN model (red). When it comes to avoiding mistakes, however, the Probit model is inaccurate for lower values of μ but quickly recovers and performs better than the KNN model for higher values of μ.

The NN model (green) seems to be very insensitive to changes in μ. In fact it starts being sensitive for values that are very close to 1, which cannot be captured by this figure unless scales are changed. This fact exemplifies the inappropriateness of using the threshold value μ as a reference to compare across models. Instead, we use μ to set up a level of alpha and then evaluate procedures by comparing their coverages. In this way we are able to fix the number of mistakes we are willing to commit, and then choose the preferred procedure as the one that maximizes the coverage.

Figure 3.5: *Comparison across models*



Predictive performance of the different models (Probit = blue; NN = green; KNN = red) in terms of coverage (share of geo-referenced patents) and reliability (probability that the location is correct).

The top-right graph ranks procedures (in terms of coverage) after having fixed a desired level of alpha (vertical axis). This graph shares the same vertical axis as the one on its left but it is scaled differently, only for three values of alpha (5%, 2.5%, and 1%). Note that the Probit model outperforms all other procedures as it obtains the highest coverage (horizontal axis) for any given value of alpha. This result also holds also when we compare procedures in step 3. As a result, we are going to predict whether candidate locations are correct relying on the probabilities we obtain after evaluating covariates of each location based on our calibrated Probit model.

3.5.2 Geographical Distribution of the Final Sample

One concern is whether the set of final locations in our sample (those that passed the Probit filtering) are representative of the true geographical distribution of patent locations. It may be the case that the ubiquity of some city names triggered false positives beyond what can be considered statistically acceptable. We propose to perform a Pearson's chi-squared test to evaluate whether observed differences between the geographical distribution of the manually collected sample and our final database can be considered statistically insignificant. We use a final sample targeting an error rate (α) of 5%, meaning that we set the filtering parameter μ so as to admit, at most, a misclassification rate of 5%.

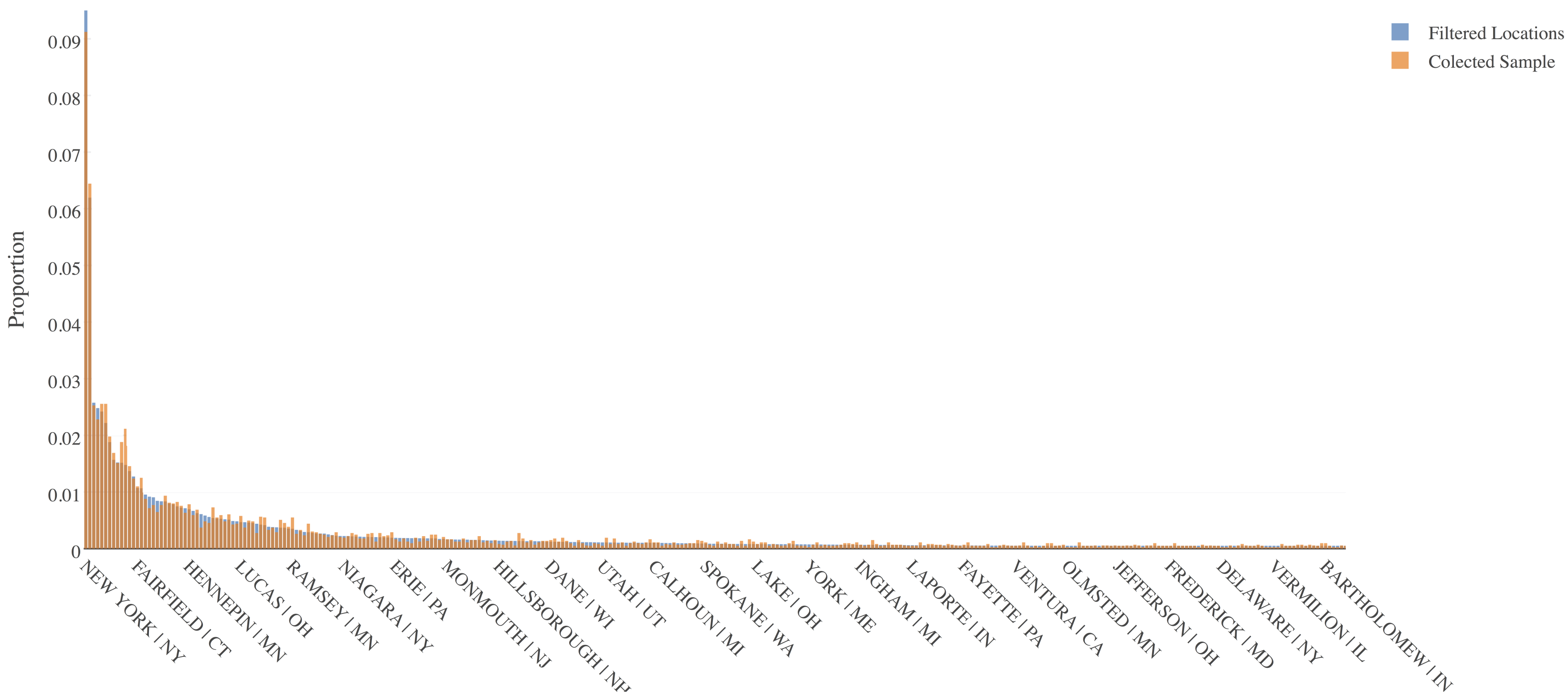
Let $p_1^{FD}, p_2^{FD}, p_3^{FD}...p_k^{FD}$ be the proportion of patents coming from locations $1, 2, 3...k$ in the Final Database (FD) and let $p_1^{MCS}, p_2^{MCS}, p_3^{MCS}...p_k^{MCS}$ be the proportion of patents in the Manually Collected Sample (MCS), for those same locations. Pearson's chi squared test evaluates the following hypothesis:

$$\begin{aligned} H_0: &\quad p_1^{FD} = p_1^{MCS}, p_2^{FD} = p_2^{MCS}, ..., p_k^{FD} = p_k^{MCS} \\ H_1: &\quad p_i^{FD} \neq p_i^{MCS} \quad for any i \end{aligned} \tag{3.6}$$

Note that the test will reject the null hypothesis if a significant difference is found for any particular location. The statistic is calculated as the sum of the standardized counts of all k locations, which is asymptotically chi-square distributed with $k-1$ degrees of freedom. Specifically, $X^2 = N \sum_{i=1}^{k} p_i^{FD} (\frac{p_i^{MCS} - p_i^{FD}}{p_i^{FD}})^2$, Where N represents the total number of observations.

Figure 3.6 shows graphically how similar both samples are, in terms of their geographical distribution. In fact, the statistical test over these two distributions gives a value of $X^2 = 2705.5$ (with a p-value of 0.975) leading us not to reject the null hypothesis that both distributions are statistically equivalent.

Figure 3.6: *Comparing distributions*



The graph compares the geographical distribution of the sample of patents collected by hand with the geographical distribution of the patents geo-referenced with our algorithm.

We also test whether this result holds along the most relevant dimensions in our sample, by type of technological domain and over time. A genuine concern may be that the statistical procedure performs poorly for particular technologies; this may happen if some technologies use a vocabulary that makes detection harder. For instance, mining or extractive technologies may reference locations to describe soil characteristics increasing the likelihood the procedure will include a false positive. In addition, changes in the way documents have been constructed may also have an effect on the likelihood that correct locations are included.

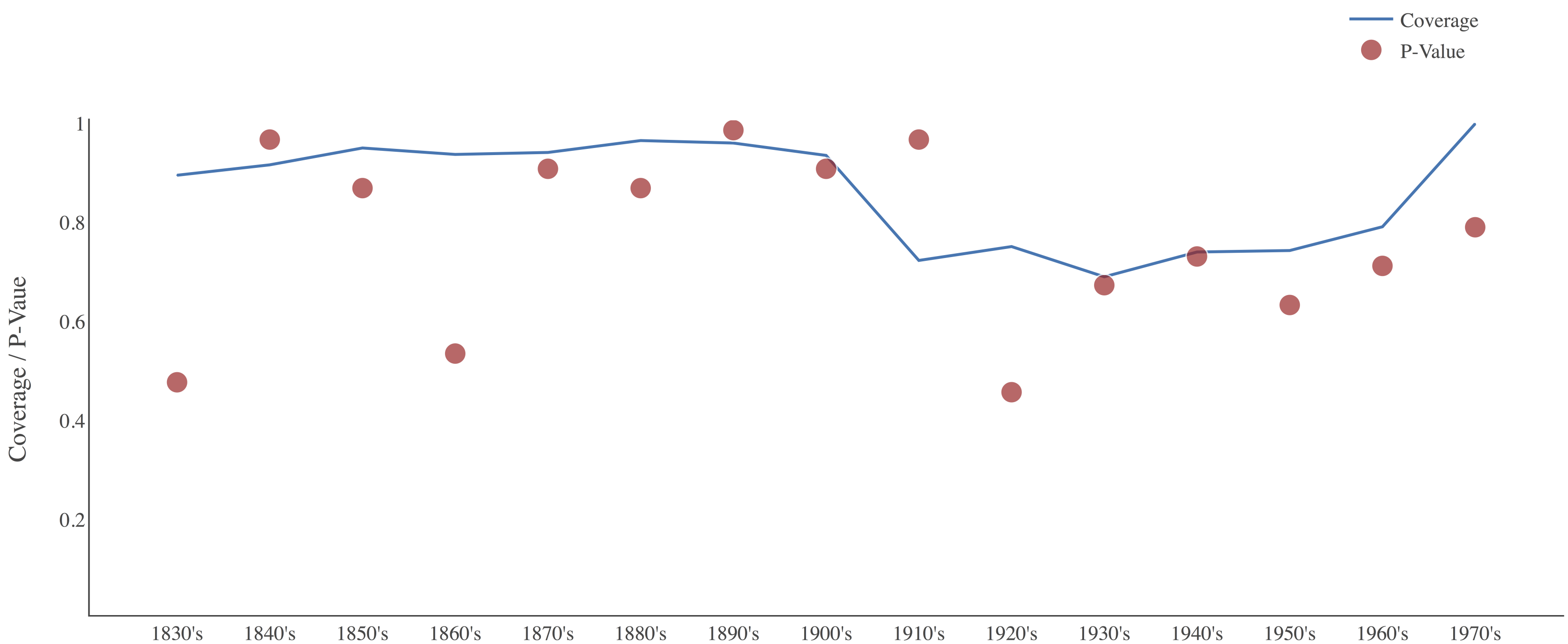
Figure 3.7 evaluates the similarity of the geographical distribution of the final database and the manually collected sample over different periods of time. We perform the same test as before but divide the sample into 15 periods of 10 years each.

Figure 3.7 shows that we maintain a relatively high rate of retrieval over all periods with the maximum value above 99% for the 1970s, and the lowest rates between 1910 and 1950 with a coverage that does not fall below 69%. Note that we never reject the null hypothesis that the final database is geographically unbiased for any period (the lowest p-value is 0.42).

Figure 3.8 below is analogous to the previous figure but considers different technological domains instead of time periods. As before, results show that we have a relatively high and homogeneous rate of coverage across technological domains. Also, we never reject the null hypothesis that our final sample is unbiased (the lowest p-value is 0.15 for Heating technologies).

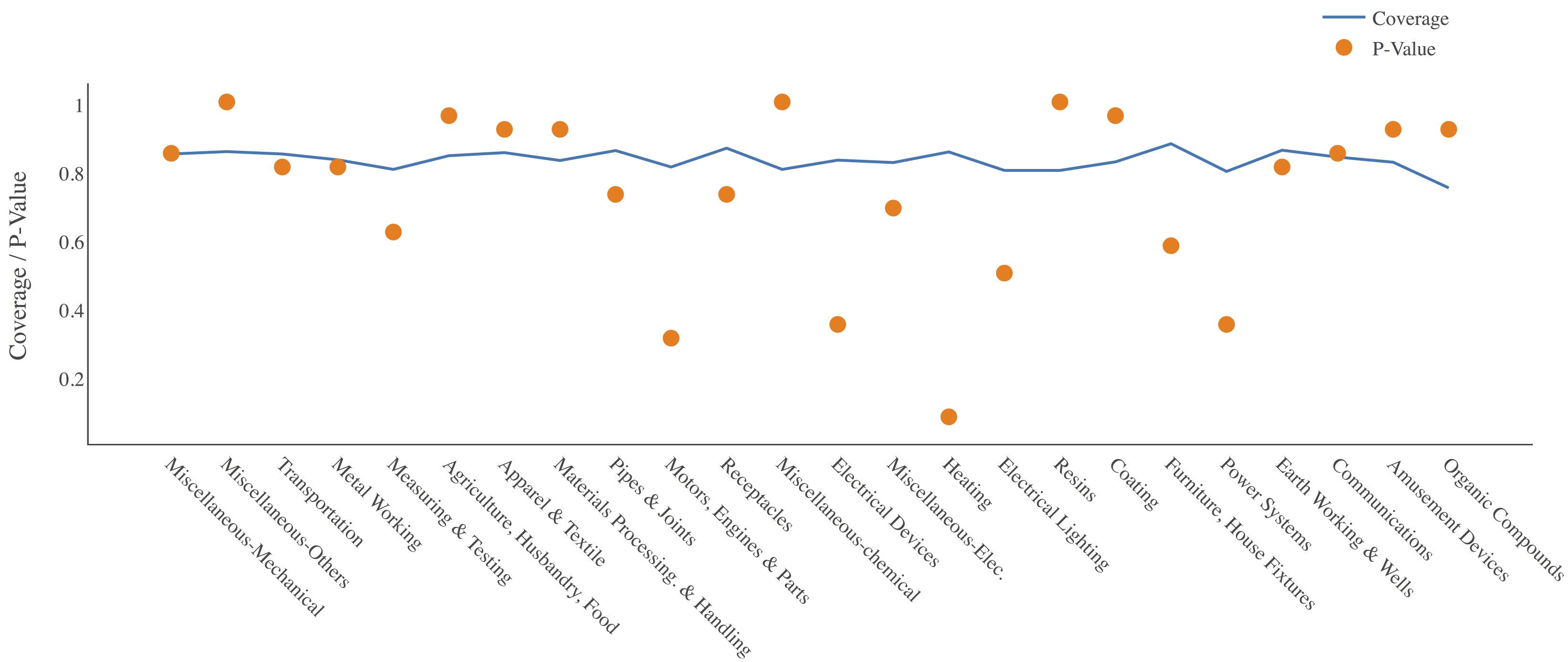
Recall that all these tests were done by setting a threshold level that corresponds to an error rate (alpha) of 5%. Even higher p-values are obtained at alpha values of 2.5% and 1%.

Figure 3.7: *Coverage over time*



The graph shows the share of geo-referenced patents (coverage) for each period and the associated p-values. Red coloured dots represent the p-values of the statistical chi-squared test for each period, while the blue line shows the coverage of the sample over that period. As before, coverage is calculated as the proportion of patents with at least one location in the final database.

Figure 3.8: *Coverage across technologies*



The graph shows the share of geo-referenced patents (coverage) for different technological fields. The size of the dots represents the share of patents in each category.

3.6 Usage Notes

A more detailed visualization of the database (including maps) can be found at https://histpat.shinyapps.io/histpat/. We plan to include new updates of this database to include manually collected data for those patents we could not retrieve automatically. We recommend checking for the latest version as we continuously update the database to include manually collected locations for those patents that couldn't be inputted by one of our procedures.

CHAPTER 4

From ideas to trade

This chapter has been produced in collaboration with C. Pan and F. Yu.

4.1 Introduction

One of the oldest and most well-known theories of international trade, the Ricardian theory, highlights the role of technological dispersion as the key driver of bilateral trade. Differences in technological capabilities across sectors and across countries determine who exports what good. Countries will benefit by specializing in those goods in which they have a comparative advantage, exchanging them for other goods. Eaton and Kortum (2002), henceforth EK, develop a general equilibrium model that extends the Ricardian framework to many countries and many goods, capturing how the opposing forces of technological change and geographic barriers affect bilateral trade. This seminal model, however, assumes that technological dispersion is the same for every country. In this chapter we show that technological dispersion varies substantially across countries, and that this variation plays an important role in bilateral trade; both theoretically and empirically.

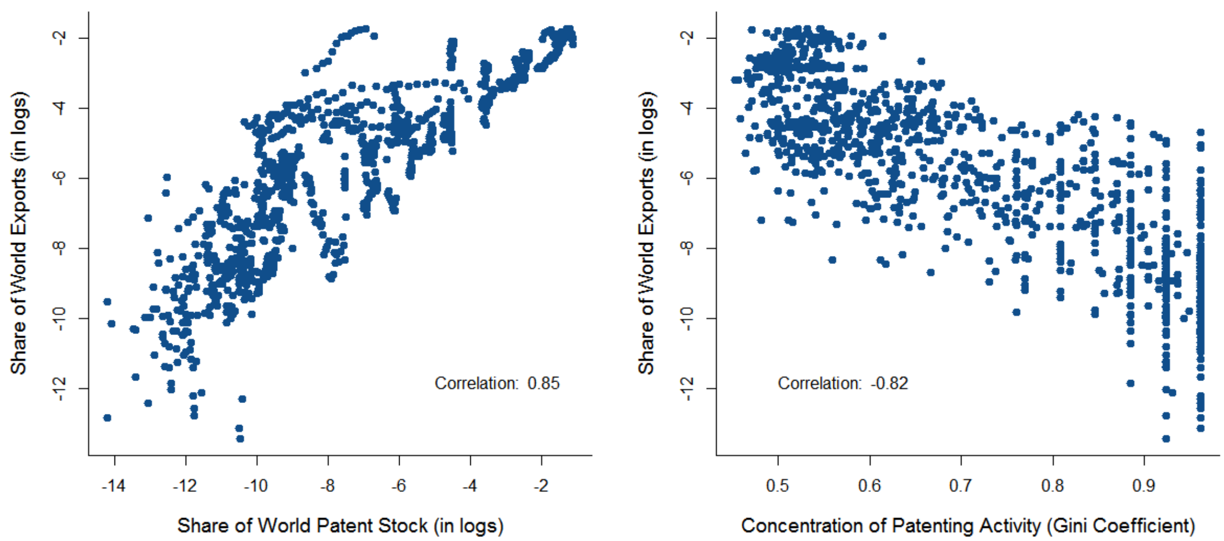
Figure 4.1 showcases our main argument by providing evidence that technology and trade are related in both the first and second moments. The left panel plots the the share of world technology, measured by patents, against the share of world exports for many countries and years. Each dot represents a country-year, from a pool of 84 countries over the period 1985-2000. This positive association between the stock of technology and trade is not surprising and is embedded in most models of trade and innovation[1]. The right panel of Figure 4.1, however, reveals two stylized facts that motivate our work: First, dispersion in patenting activities (technology) varies across countries and years, and second, it is highly correlated with bilateral trade. In particular, using the Gini coefficient to measure the dispersion of patents across sectors in the economy, we show that countries with more concentrated patenting activity also export less[2]. Even though the graphs provide simple correlations, these strong relationships survive the inclusion of country-specific dummy variables. They suggest that cross-country differences in the dispersion of technological know-how (which are not accounted for in EK's multi-good and multi-country Ricardian model) might matter for exports.

The aim of this chapter is to study how cross-country differences in the allocation of technology affect bilateral trade. In particular, as suggested by Figure 4.1, we intend to better understand to what extent the stock and dispersion of technology affect exports and how they interact to do so. To answer these questions, several challenges need to be overcome. The main one is to reconcile theory with

[1]Models that explain the role of technology in trade, such as EK, predict that countries with a higher stock of technology will export more. In addition, studies on the role of international trade in innovation show that both higher exports (through larger markets; see Lileeva and Trefler (2010) and Bustos (2011)) and imports (through increased competition; see Bloom et al. (2015) and Steinwender (2015)) lead to higher innovation. Therefore, both channels predict a positive correlation between the stock of technology and trade.

[2]A low Gini indicates high dispersion: Patents are evenly distributed across the economy (and vice versa)

Figure 4.1: *Trade and Patents*



empirics. Most empirical studies on the effect of innovation on bilateral exports are either specific to a country (and/or sector) or purely empirical. While it has been shown that the gravity equation can be derived from various models, it is common practice to apply transformations to the canonical gravity equation as an ex post exercise. A second challenge is finding a theoretical model that is able to explain the empirical regularities observed in the data while still being simple enough to be tested. To date, theory has failed to predict trade in a world where both the stock and dispersion of technology vary from one country to another[3]. Finally, most studies that use theoretical models lack data on technology to test their predictions. For instance, in EK and many subsequent papers that use their Ricardian model, the effect of technology on trade is trapped inside a country fixed-effect dummy rather than estimated in isolation from other effects. In other studies, technological levels are derived as a residual of the model, mean-

[3]An exception to this is Costinot et al. (2012). Note that Eaton and Kortum (2002) include a technological dispersion parameter in their model, but it is assumed to be the same for all countries. Thus, the only technological difference between them arises from the overall stock of ideas.

ing that the theory cannot be properly tested. Therefore, we need measures of technology that are independent of the model and allow us to test its predictions.

We build on Eaton and Kortum (2002) and develop a Ricardian model in which the process of innovation determines both the stock *and* dispersion of technology in each country. As in the EK model, a higher stock of technology and lower relative input costs foster exports. Unlike in the EK model, a country's overall comparative advantage (which determines overall exports) depends on technological dispersion. Intuitively, our model contains an additional term: the interaction between technological dispersion and input costs. In particular, technological dispersion governs the advantage of having lower input costs. Llower dispersion benefits countries with lower input costs since their exports are determined by these and not technological differences. The opposite is the case when technological dispersion is high, as exports by these countries are determined by differences in technology and not costs. In other words, what matters is the covariance between relative input costs and technological dispersion. A country exports more when this covariance is negative, meaning that competitors with higher costs have lower technological dispersion. In addition, an interesting feature of our framework is that when we impose a common dispersion of technology across countries the model simplifies to the EK model. This allows for a direct comparison and assessment of the gains of our more general framework. We derive a gravity equation and study how changes in a country's costs and technological dispersion affect trade flows.

On the empirical side, the main difficulty lies in finding measures of the key technological variables that allow us to test the predictions of our model. These should accurately represent absolute and comparative advantage in ideas in each country, be independent of exports, and be consistent with theory. There exist several revealed comparative advantage measures in the literature, dating back to the famous 1989 Balassa Index. However, most of these lack theoretical foundations and cannot truly represent the drivers of exports (in a Ricardian spirit) since they are unable to separate the causes of exports from consequences.[4] In contrast to previous studies (mostly based on EK) that estimate technological dispersion as a coefficient on trade costs, we wish to measure it and test it in our model[5].

To test effects of technological innovation on bilateral exports we create a novel dataset of historical patents (the longest to date) and use these to construct measures of the stock and dispersion of technology by country and year. Specifically, we take patent grants at the United States Patent and Trademark Office (USPTO) from 1836 to 2000 and add geographic location (country of origin) based on the inventor's residence.[6] We use patent counts by country and year as our measure of the stock of knowledge and create a measure of technological dispersion across each economy by estimating the Kortum (1997) idea-generating model that serves as the microfoundation of EK. Our estimated dispersion parameters are in the range of EK and other previous estimates in the literature (obtained using different samples and techniques), such as Costinot et al. (2012) and Simonovska and Waugh (2011).

This chapter contributes to an extensive literature concerned with the role of technological advance in international trade that dates back to David Ricardo's famous 1817 model. Recent extensions of the classical theory include EK's general equilibrium multi-country setting and the multi-sector extensions of Caliendo and Parro (2014), Chor (2010), Costinot et al. (2012), and Shikher (2011). The latter develops a model that introduces factor endowments and leads to an HO-Ricardian hybrid. It departs from this literature in three main respects.

[4]An exception to this is provided by Costinot et al. (2012) and Leromain and Orefice (2014), who develop measures of comparative advantage that isolate exporter-specific characteristics that might drive trade flows.

[5]In those models technological dispersion can only be estimated and not tested.

[6]For patents previous to 1975 we went through the digitalised patents available in Google via Reed Tech and collected all the necessary information. The procedure we followed is described in detail under the Data section.

First, to construct our technology measures, we use data on patents that better reflect (technological) productivity than other indicators (such as wholesale prices)[7]. Second, rather than simply estimating the dispersion parameter, our model is able to capture the effect of technological dispersion on bilateral trade. Finally, since our model was constructed to embed the benchmark EK model, the two can be easily compared and the gains of incorporating a country-varying technological dispersion parameter can be easily assessed.

This chapter is also related to empirical studies concerned with both testing the Ricardian model and constructing comparative advantage measures and studying their evolution. Examples of the these include Kerr (2013), Simonovska and Waugh (2011), Levchenko and Zhang (2016), Leromain and Orefice (2014), and Bolatto (2013). Our technological dispersion (comparative advantage) country-specific measures differ from those in the literature in that they are derived from an idea-generating process and thus consistent with the theory that microfounds our Ricardian model.

The remainder of the chapter is organized as follows. Section 4.2 describes our theoretical framework. Section 4.3 presents our data and discusses the empirical specification. We derive a gravity equation from our theoretical model and use it as the estimating equation. Section 4.4 tests our model using panel data. Our empirical results match our theoretical predictions, suggesting that both the stock and distribution of knowledge play a fundamental role in bilateral trade. Several robustness tests are performed to confirm our results. Finally, Section 4.5 concludes and suggests implications for policy.

[7]Some studies have used R&D data to measure patent stocks, but not dispersion.

4.2 Theoretical Framework

We develop a simple Ricardian model of innovation and trade that builds on Eaton and Kortum (2002) and incorporates country-level technological heterogeneities. As in previous studies, the model accounts for differences in the technological stock across countries. Unlike previous studies, it also accounts for differences in how countries distribute their technological stock across industries.

4.2.1 Model Setup

The world economy consists of N countries indexed by $i = 1, ..., N$ and a continuum of goods indexed by $j \in [0, 1]$. Under constant returns to scale, the cost of producing one unit of good j is $c_i/z_i(j)$, where z_i denotes the number of units of the good produced by one unit of input (efficiency), and c_i is the input cost in country i. Geographic barriers are introduced by means of an iceberg cost $d_{ni} > 1$, the cost of delivering one unit from i to n. Perfect competition makes the price that country i charges in country n for one unit of good j equal to the cost of delivering one unit to n.

$$p_{ni}(j) = \left(\frac{c_i}{z_i(j)}\right) d_{ni}$$

The actual price that buyers in country n will pay for good j is the lowest across all sources: $p_{nj} = \min_i\{p_{ni}(j)\}$. Country i's efficiency in producing good j is the realization of a random variable z_i (drawn independently for each j) from its country-specific Frechet probability distribution $F_i(z) = e^{-Tz^{-\theta}}$. Buyers in country n buy from the cheapest source, meaning that the probability that country i

provides a good at the lowest price in country n is

$$\begin{aligned}
\pi_{ni} &= Pr(P_{ni}(j) \leq min_{k\neq i}P_{nk}(j)) \\
&= Pr\left(\frac{c_i d_{ni}}{Z_i} \leq \min_{k\neq i}\frac{c_i d_{nk}}{Z_k}\right) \\
&= \Pi_{k\neq i}^{N} E\left(Pr\left(Z_k \leq z_i \frac{c_k d_{nk}}{c_i d_{ni}}\middle| z_i\right)\right) \qquad (4.1)
\end{aligned}$$

4.2.2 Technology: the role of T and θ

The distribution of efficiencies provides the key to understanding the role of technology in trade. In particular, the Frechet distribution is governed by two parameters, T and θ, that depict two aspects of the countries' technological capabilities. T represents the overall stock of technology, or *absolute advantage*. A higher T increases the likelihood that goods produced by country i are more efficient (require less labor per unit). Statistically, T governs the location of the distribution. Figure 4.2 shows that increases in T shift the Frechet distribution to the right, making higher-efficiency productivity draws for all goods more likely. The parameter θ represents the dispersion of technology, or in EK's terms *the force of comparative advantage*. It measures dispersion in the labor requirement (efficiency) across goods. As shown in Figure 4.3, θ determines the shape of the distribution. A high θ means that all of the input requirements (or efficiencies) drawn from the country-specific Frechet distribution are close to the mean: The country is similarly productive in all of its sectors.[8] In other words, the force of comparative advantage is weak. In this model a country will sell a good only if it is the lowest-cost supplier. The position of the Frechet curve for each country, determined by the country's T and θ, will determine the efficiency draws and thus the export probability. The more "to the right" the Frechet curve is, the more pro-

[8]Thus, it is neither exceptionally bad nor exceptionally good at anything.

ductive the country will be in all the goods that it produces and the more likely it will be to export these.

However what does it mean, in practice, for a country to increase T or θ? Although these are exogenously assigned to countries in this model, in reality, countries can choose how to allocate their knowledge. As time passes countries accumulate more technology, i.e. by means of R&D expenditures, therefore raising T.[9] The evolution of θ depends on how this new technology is allocated across different industries. If it goes to industries that were already technology abundant relative to the rest, then θ will decrease and the difference in efficiencies across industries will become even more pronounced. A low θ thus refers to a very uneven distribution of technology across sectors. Conversely, allocating the new technology toward industries with technological scarcity will bring all efficiencies closer. However, helping the most inefficient sectors, which will raise θ, comes at the expense of the most productive industries.

Eaton and Kortum (2002) assume that the distribution of country i's efficiency Z_i is $F_i(z) = e^{-T_i z^{-\theta}}$. Since θ is fixed, countries only differ in their stock of technology T (absolute advantage) and the world can be perfectly described by Figure 4.2. Countries draw their efficiencies from similar distributions (in shape), and thus, their differences arise from some distributions being shifted to the right due to a larger stock of technology. Our contribution to the literature is to allow countries to differ in how they distribute their technology across their industries. By introducing a country-specific technological dispersion parameter θ_i, the world now looks like Figure 4.4. The probability of exporting depends on *both* the country-specific stock and dispersion of technology, meaning that it is not so obvious what countries ought to do to “move to the right" and become more productive than the rest of the world.

[9] T can never decrease in this model since it refers to the stock of ideas rather than physical capital.

Figure 4.2: *Changes in Absolute Advantage (T) Holding Comparative Advantage (θ) Fixed*

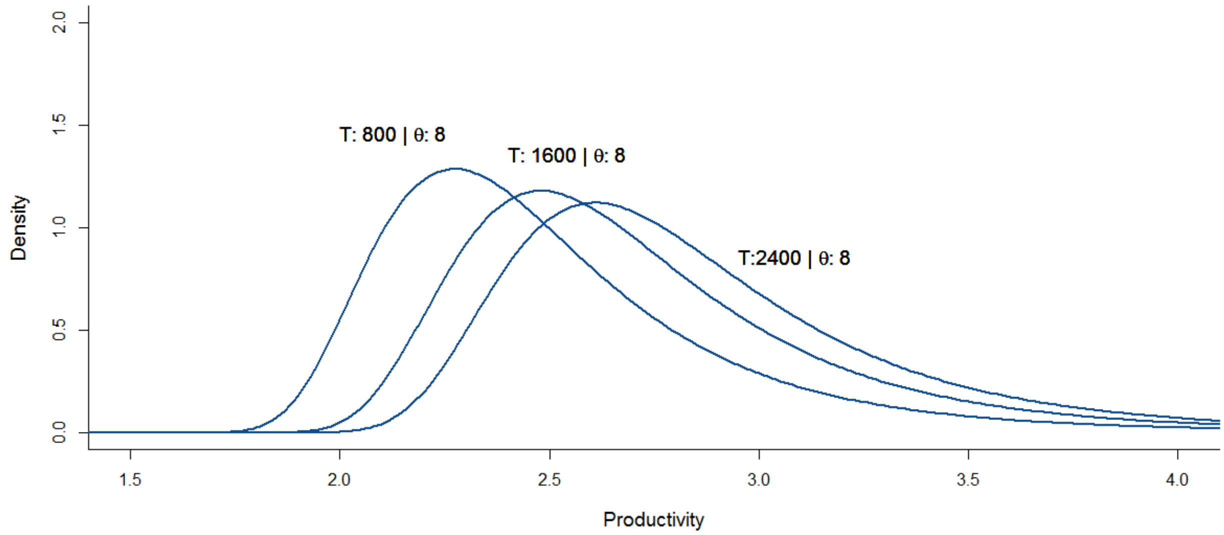


Figure 4.3: *Changes in Comparative Advantage (θ), Holding Absolute Advantage (T) Fixed.*

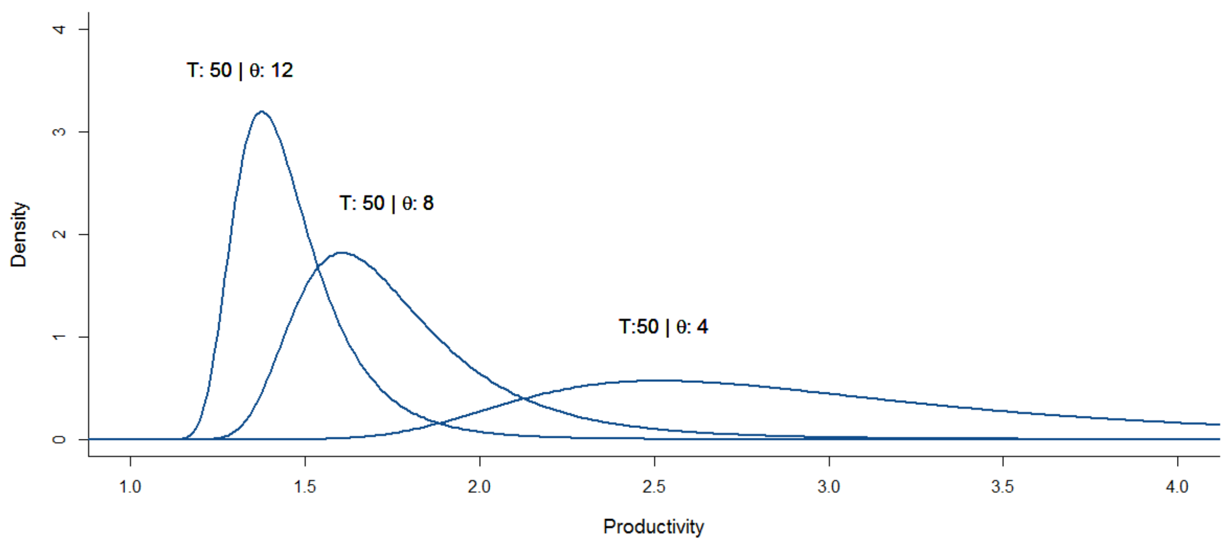
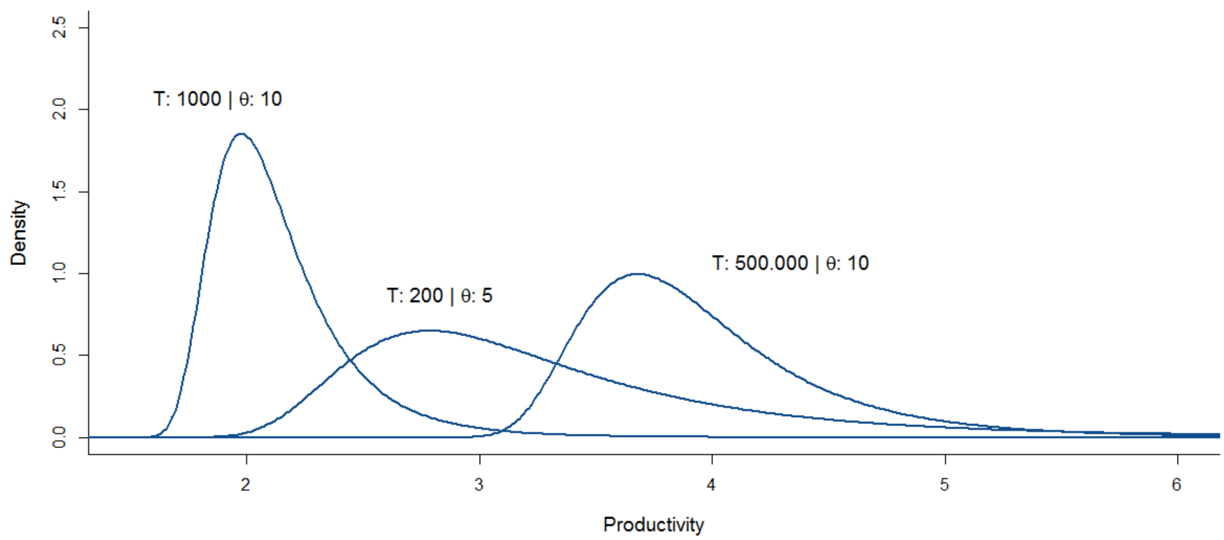


Figure 4.4: *Simultaneous Changes in Absolute and Comparative Advantage*



4.2.3 A small model with country-specific comparative advantage

To develop some intuition on the role of the country-specific comparative advantage, we will focus on a model with only two countries, Home and Foreign, and assume that z_i is log-normally distributed:

$$z_i \sim LN(\mu_i, \sigma_i^2)$$

where $i \in \{\text{Home}, \text{Foreign}\}$. Here μ_i captures the level of absolute advantage of country i and σ_i^2 is a measure of dispersion. In the spirit of Dornbusch et al. (1977), we can order sectors based on decreasing comparative advantage from the perspective of Home. This defines the downward-sloping curve $A(j)$

$$\begin{aligned} j &= P\left(\frac{z_H}{z_F} > A(j)\right) \\ &= 1 - \Phi\left(\frac{\ln A(j) - (\mu_H - \mu_F)}{\sqrt{\sigma_H^2 + \sigma_F^2}}\right) \end{aligned}$$

In equilibrium there is a cut-off sector j^* such that Home produces goods $j \in [0, j^*]$ defined by

$$\frac{w_H}{w_F} = A(j^*)$$

such that

$$j^* = 1 - \Phi\left(\frac{\ln w_H - \ln w_F - (\mu_H - \mu_F)}{\sqrt{\sigma_H^2 + \sigma_F^2}}\right) \tag{4.2}$$

What happens to the range of goods produced at Home when the dispersion parameter σ_H^2 increases? It depends on the equilibrium relative wage that appears in the numerator of equation (4.2).

If we impose CES preferences then each country spends a fraction j^* of its income on goods produced at Home. Total spending in equilibrium has to equal total wages at Home

$$w_H L_H = j^* w_H L_H + j^* w_F L_F \tag{4.3}$$

Starting from a symmetric case in which $L_H = L_F$, it can be shown that (4.2) and (4.3) imply

$$\frac{\partial j^*}{\partial \sigma_H^2} > 0 \iff \ln w_H - \ln w_F > \mu_H - \mu_F \iff \mu_H < \mu_F$$

meaning that an increase in technological dispersion increases the range of goods produced at Home if and only if Home is a technological follower (has less absolute advantage than Foreign). The intuition behind this result is as follows. Home is on average more expensive than Foreign and thus produces in equilibrium a narrower range of goods. As a result, an increase in dispersion dampens the effect of absolute advantage and therefore increases the range of goods produced in the country with less absolute advantage (technological follower). Notice however that the effect depends on the initial values of dispersion by country σ_H^2 and σ_F^2.

4.2.4 Full model with country-specific comparative advantage

When θ_k is country-specific, the probability that country i provides a good at the lowest price in country n is

$$\begin{aligned}
\pi_{ni} &= P(p_{ni}(j) \leq \min_{k\neq i} p_{nk}(j)) \\
&= P\left(\frac{c_i d_{ni}}{z_i(j)} \leq \min_{k\neq i} \frac{c_k d_{nk}}{z_k(j)}\right) \\
&= \int_0^\infty \Pi_{k\neq i}^N e^{-T_k\left(\frac{c_k d_{nk} z_i}{c_i d_{ni}}\right)^{-\theta_k}} \theta_i T_i z_i^{-\theta_i-1} e^{-T_i z_i^{-\theta_i}} dz_i \\
&= \int_0^\infty e^{-\sum_{k=1}^N T_k\left(\frac{c_k d_{nk} z_i}{c_i d_{ni}}\right)^{-\theta_k}} \theta_i T_i z_i^{-\theta_i-1} dz_i
\end{aligned} \tag{4.4}$$

In the appendix we show that the term inside the integral can be replaced by

$$\sum_{k=1}^N T_k \left(\frac{c_k d_{nk} z_i}{c_i d_{ni}}\right)^{-\theta_k} = \sum_{k=1}^N T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right)^{-\theta_k} z_i^{-\theta_i} + \varepsilon$$

where the expectation of the approximation error $E(\varepsilon)$ is second order. This approximation treats the sum with heterogeneous θ_k as if it were a sum with a fixed θ_i, as differences in the exponents will tend to cancel out.[10] With this approximation, we can simplify expression (4.4) to obtain

$$\pi_{ni} = \frac{T_i}{\sum_{k=1}^N T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right)^{-\theta_k}} \tag{4.5}$$

The term in the denominator is a measure of world competitiveness relative to country i. It reflects how much cheaper (in terms of input and transport costs)

[10] See the appendix for an analysis of the accuracy of this approximation.

the rest of the world is relative to i. Intuitively, exporter i will be more successful if its technological level is higher than world competitiveness. Since in this model π_{ni} is also the fraction that country n spends on goods from i, equation (4.5) becomes

$$\frac{X_{ni}}{X_n} = \frac{T_i}{\sum_{k=1}^{N} T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right)^{-\theta_k}} \tag{4.6}$$

The left-hand side variable is normalized bilateral imports: i's imports from n adjusted by Home purchases. We can think of (4.6) as the model's gravity equation, as it relates normalized bilateral trade to the stock of technology and relative input and transport costs (such as wages and geographic distance). Taking logs and expanding with respect to the θ_k parameters up to a first order we obtain

$$\ln \frac{X_{ni}}{X_n} = \ln T_i - \ln \left(\sum_{k=1}^{N} T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right)^{-\theta} \right) + \sum_{k=1}^{N} \alpha_k \ln \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right) (\theta_k - \theta) \tag{4.7}$$

where α_k is the relative standing of country k in world competitiveness

$$\alpha_k = \frac{T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right)^{-\theta_k}}{\sum_{j=1}^{N} T_j \left(\frac{c_j d_{nj}}{c_i d_{ni}}\right)^{-\theta_j}}$$

Equation (4.7) summarizes our model. Bilateral trade is related to the exporter's technological stock or absolute advantage (first term), a world competitiveness index (relative to i, second term), and a comparative advantage term. Note that, if all countries have the same dispersion parameter, $\theta_k = \theta \; \forall k$, the last term equals zero and our model simplifies to the benchmark Eaton and Kortum

(2002).[11] This is an advantage of our setup, compared to others, as it allows for easy comparison between the models.

The EK model, contained in the first two terms, provides a gravity equation to study the effect of trade costs, represented by geography and technology, on trade patterns. In particular, it predicts that exports from country i are higher when it has more absolute advantage relative to world competitiveness. That is, country i will export more to country n the more technology that it has accumulated and the higher the trade costs of the rest of the countries are (relative to i).[12] The only role of the world technological dispersion parameter θ is in shaping the elasticity of imports with respect to input costs and geographic barriers.

Note that an increase in the relative cost of any given country k will have a similar effect (henceforth, the EK effect) on i, as the model imposes a common dispersion θ for every country. In other words, in the EK model, from the perspective of i's exports to n, it is irrelevant which competitor k experiences an increase in relative costs. Similarly, any increase in world dispersion (θ) will equally benefit all exporters. However, as we already anticipated, technological dispersion (measured by patenting data) varies significantly across countries, and these differences are important determinants of bilateral trade.

The key feature of our model is, compared to the standard EK gravity, the additional (third) term of equation (4.7). The comparative advantage term helps us to better understand the effect of both a change in a country's relative costs and a change in the dispersion parameter on exports by introducing an effect that has thus far been neglected in the literature: A country's exports also depend on its relative world standing regarding costs and technology. Thus, any changes in a competitor's k input costs or technological dispersion (i.e. it becomes more technologically specialized or diversified) will affect i's exports to n.

[11] $\ln \frac{X_{ni}}{X_n} = \ln T_i - \ln \left(\sum_{k=1}^{N} T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}} \right)^{-\theta} \right)$

[12] The trade costs have an exporter-specific component (input costs) and a bilateral pair component (i.e. geographic distance).

With regard to relative costs, the baseline EK effect still holds and is captured by our world competitiveness (second) term. An increase in k's costs relative to i will benefit i's exports to n. However, there is an additional effect coming from the comparative advantage (third) term, and thus, the overall effect can be augmented or dampened depending on the magnitude of the productivity dispersion parameter θ_k. The effect will be higher when θ_k is large, that is, when productivity dispersion is lower. Since country k's force of comparative advantage is weaker, any given change in k's costs has a larger effect on i's exports.

Our model captures, through the comparative advantage term, how the country-specific productivity dispersion co-varies with relative costs. We can see from equation (4.7) that the effect of an increase in θ_k depends on the sign of the log of relative costs. If country k is relatively less competitive (has higher costs relative to i), then the log term is positive and an increase in θ_k increases exports from i to n. Intuitively, a larger θ_k dampens the force of comparative advantage and increases the effect of a difference in relative costs. This effect was already present in the two-country model of Section 4.2.3. A reduction in productivity dispersion increases the exports of the country that is relatively less expensive. In the multi-country case, we also observe that this effect is larger when α_k is large: An increase in θ_k favors country i's exports, especially if country k is highly competitive relative to the world's average competitiveness index.

When is country k relatively more expensive than country i? The answer was already provided in the two-country model of Section 4.2.3, which suggested that in a two-country world with log-normal productivity, the country with a higher absolute advantage was relatively less expensive. If we assume that labor is the only input in production, the equilibrium in the multi-country model is given by

$$w_i L_i = \sum_{n=1}^{N} \pi_{ni} w_n L_n$$

We can solve the model in closed form if we assume no trade barriers, meaning that $d_{ni} = 1$ for all country pairs, and a common dispersion parameter θ for all countries. In that case, we obtain

$$\frac{w_n/T_n^{1/\theta}}{w_i/T_i^{1/\theta}} = \left(\frac{T_n^{1/\theta}/L_n}{T_i^{1/\theta}/L_i}\right)^{-\frac{1}{1+\theta}}$$

meaning that if country n has higher absolute advantage measured by $T_n^{1/\theta}$ then it will be more competitive in terms of productivity-adjusted labor costs.

4.3 Data Description

We construct a unique panel dataset containing measures of the stock of technology T, the productivity dispersion θ, bilateral trade, input costs, trade costs, and other bilateral characteristics (such as shared language or border) for 84 developed and developing countries, from 1980 to 2000.[13] We follow Eaton and Kortum (2010) in understanding technology as the outcome of a process that begins with an idea and therefore use patent data to measure the technological stock and dispersion. These two are then used to construct the absolute and comparative advantage variables T and θ in a theory-consistent way, which constitutes one of the main contributions of this chapter. For the rest of the variables, we follow the literature in choosing widely used measures and databases. To measure trade, we use UN COMTRADE bilateral imports. Data on GDP per capita by country and year come from the World Bank's World Development Indicators (WDI). Our measures of trade costs (geographic distance, common language, border, common currency, common colonizer, etc.) by bilateral pair are from the CEPII gravdata dataset. Finally, we use data on wages by country and year from

[13]A list of all countries can be found in the Data Appendix.

the International Labour Organization (ILO). Below, we describe the sources of the technological data and the construction of the key variables.

4.3.1 Patent Data

Patent grants at the USPTO are our indicator of the technological capabilities of countries. Data on patenting activity covering the period 1975-2000 were obtained from the "Patent Network Dataverse" developed by the Institute for Quantitative Social Science at Harvard University (Lai et al., 2011) using original data from the USPTO. This database contains all patents granted by the USPTO to resident and non-resident inventors along with their address information, which we used to determine and assign the origin of the patent.[14] To identify older patent grants (pre-1975) at the USPTO, we developed an algorithm that retrieves the location information of optically recognized (OCR) historical patent documents. In 2006, the USPTO began a series of no-cost agreements with Reed Tech and Google to digitalize all available patent documents dating back to 1790, making OCR patent documents available to anyone free of charge.[15] This algorithm finds references to geographic locations (country names) within patent documents to later evaluate the likelihood that a reference is indeed the location of an inventor/assignee in a specific patent. Our algorithm is analogous to that used in Petralia et al. (2016) but for international patents.[16]

In summary, for each patent at the USPTO since 1836, we were able to retrieve information on the country of origin, the year it was granted, the patent class, and the number of citations. The latter is used, in line with the innovation literature,

[14]This assumes that the knowledge is where the inventor. If a patent has several inventors in different locations, we assigned an entire patent count to each country. The results do not change if instead a proportional fraction is assigned to each country.

[15]Even though the earliest patent available dates back to 1790, coverage between 1790 and 1836 is scattered and not reliable. This is because a fire at the USPTO destroyed file histories of thousands of patents and pending applications in 1836. For more information see https://www.google.com/googlebooks/uspto.html, and http://www.uspto.gov/learning-and-resources/electronic-bulk-data-products.

[16]See our Data Appendix for further detail.

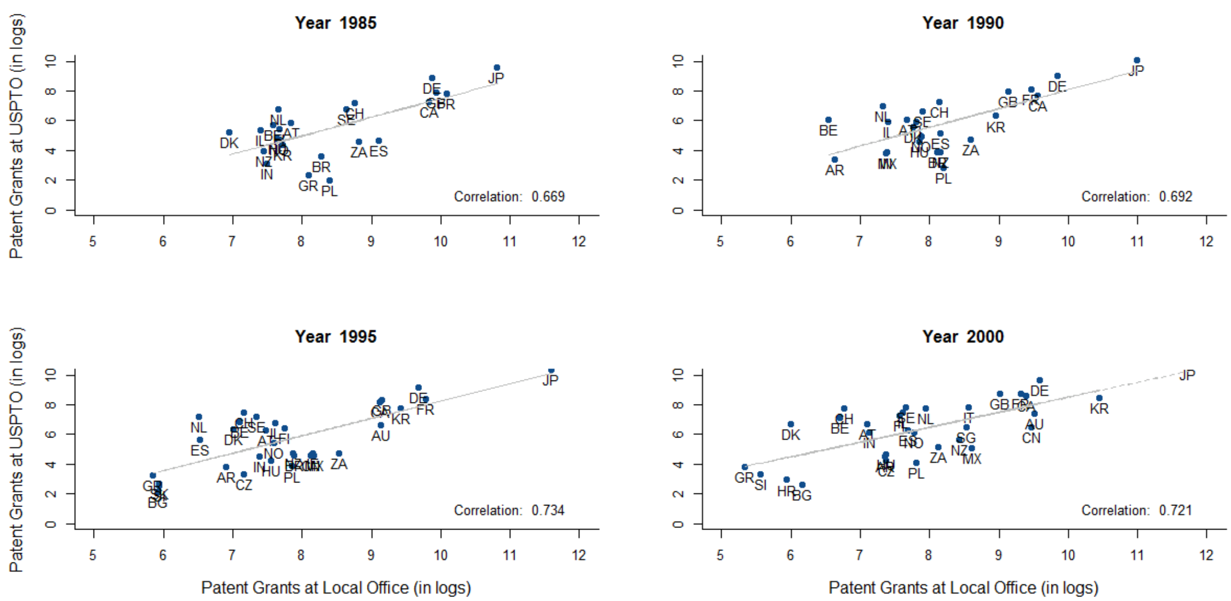
as a measure of patent quality. Quality dispersion will be crucial for our empirics since they will identify technological dispersion, as we show below. Note that, to avoid a home bias effect, we exclude US patents from the analysis and focus on foreign patents in the US. We chose to use data on patents at USPTO rather than individual patent offices for easier comparison between countries (the same criteria for all countries) and to ensure greater reliability and the availability of scanned historic patents. In the next section, we show that patents at the USPTO are a good measure of countries' technological innovation and describe the evolution of foreign patents over time.

Patent Descriptives

The usefulness of patent grants at the USPTO as a measure of countries' technological innovation relies on two assumptions: that patents indeed capture technological progress and that the patenting behavior of countries is similar at home and abroad (USPTO). The first assumption has been widely debated on the innovation literature, which concludes that patents and R&D expenditures are the best available proxies for technological innovation. Since patents are an output rather than an input measure, they best represent countries' technological stock. The second claim deserves closer scrutiny. Figure 4.5 compares patent grants at the USPTO with grants at the local (home) office since 1980. The comparison for earlier dates can be seen in Figure 4.10 in the Appendix. There is a strong correlation between the two patenting activities: Countries that patent more at home also patent more abroad, in particular at the USPTO.

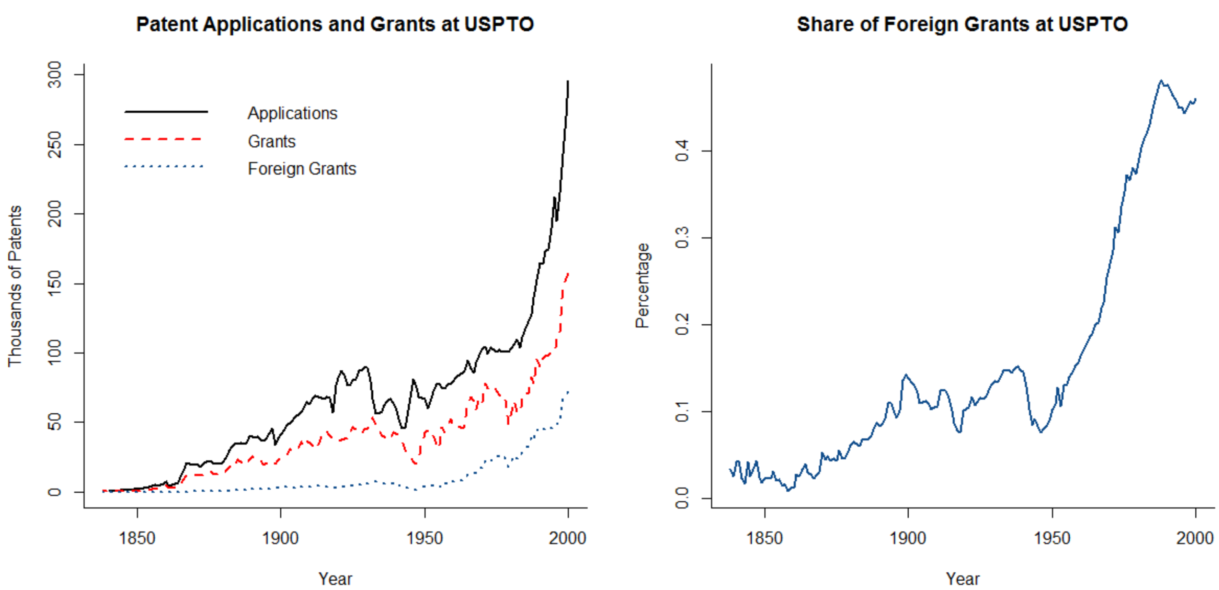
The historical evolution of patent grants is depicted by Figures 4.6 and 4.7. Foreign patent grants at the USPTO have increased drastically since the 1830s, as seen in Figure 4.6, with a clear change of pace circa the 1950s. They went from a negligible 3% to a modest 15% of total patent grants between the middle and end of the 19th century and later jumped to represent more than 40% of overall patents at the USPTO by the end of the 20th century. The exponential increase in foreign grants at the USPTO can be partially attributed to the contribution of Japan as a key player in the world production of technological knowledge. In fact, as shown in Figure 4.7 below, Japan climbed from producing less than 3% of all patent grants in the 1950s to producing nearly 50% by the turn of the century. The countries that consistently patent the most over the entire time period are Japan, Germany, Great Britain, France, and Canada.

Figure 4.5: *Grants at local office vs. USPTO*



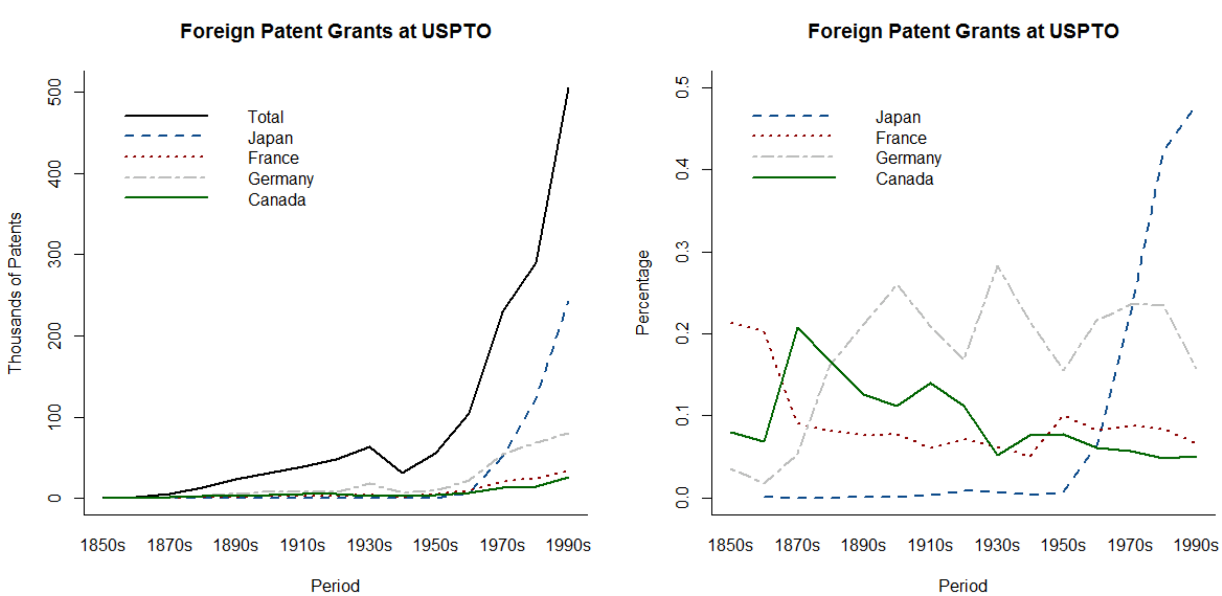
Source: Own elaboration based on USPTO and WIPO statistics.

Figure 4.6: *Patent Applications and Grants at the USPTO*



Source: Own elaboration based on USPTO and WIPO statistics.

Figure 4.7: *Composition of Foreign Patent Grants at the USPTO*



Source: Own elaboration based on HistPat and Harvard Patent Dataverse.

Measures of T and θ

We construct our measures of technological stock (T) and dispersion (θ) by combining the patent data with a general equilibrium model of technological change from Kortum (1997). This idea-generating model provides the micro-foundation for the Frechet distributions that we assume in the main model, yielding a theory-consistent approach to estimating the distribution parameters T and θ that will allow us to construct our technological variables of interest.

According to Eaton and Kortum (2010), an idea is the core of technology and can be described as "a recipe to produce some good j with some efficiency q (quality of the idea) at some location i". In this model, ideas arrive to researchers as a Poisson process with an intensity (arrival rate) that depends on both current research effort and the history of the arrival of ideas $T(t) = \int_{-\infty}^{t} R(\tau)d\tau$, where $R(\tau)$ is past research effort.[17] Ideas have a quality Q with a Pareto probability distribution:

$$P(Q > q) = \begin{cases} \left(\frac{q}{\underline{q}}\right)^{-\theta} & q \geq q \\ 1 & q < q \end{cases}$$

It follows that ideas with quality $Q \geq q$ arrive to researchers with intensity $T(t)q^{-\theta}$. It can be shown that if the distribution of ideas is Pareto, then the distribution of the best ideas is Frechet with parameters T and θ.

The probability that y_i ideas (patents) with quality q_i arrive in a given year is

$$P(Y = y) = \prod_i e^{-T(t)q_i^{-\theta}} \frac{(T(t)q_i^{-\theta})^{y_i}}{y_i!}$$

[17]Eaton and Kortum (2010) describe this as a "no-forgetting" feature of the model.

We obtain values for idea qualities based on citations and proxy for T with the stock of patents at time t.

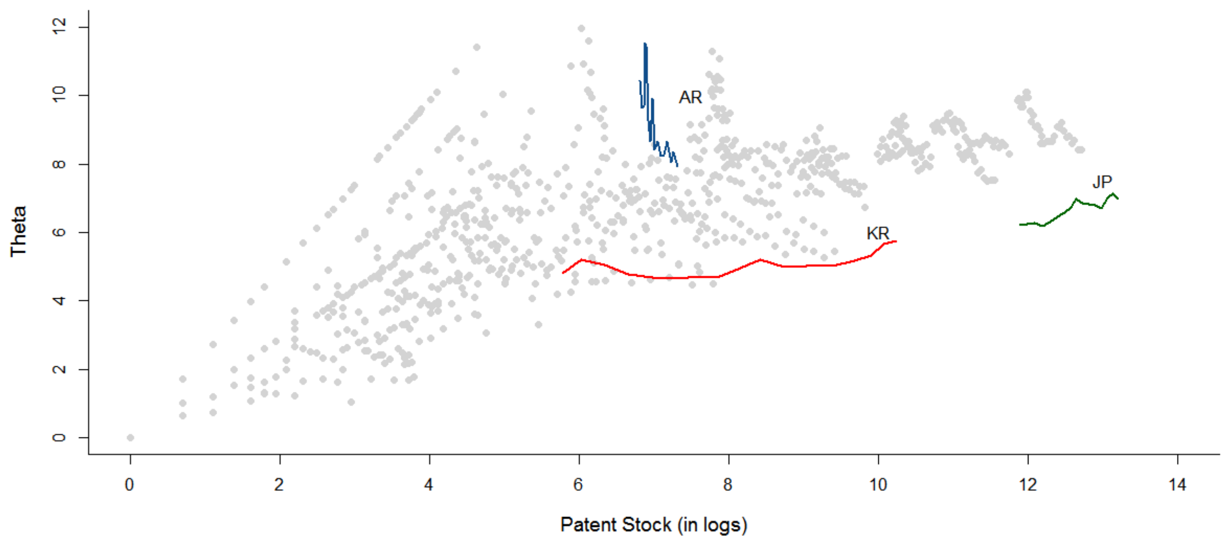
$$T(t) = \sum_{k=1836}^{t} \text{Patents}_k$$

We estimate one Poisson model per country and year, and retrieve $\hat{\theta}_i$. Figure 4.8 plots the estimated θ_i against the technological stock T_i for all countries and all years in our sample. Each dot represents a country-year for a pool of 84 countries during the period 1980-2000[18]. We see that all of our estimates for θ_i lie between 1 and 12, which is consistent with other estimates in the literature[19]. We follow three countries, Argentina, Korea, and Japan, over time and plot their implied country-specific Frechet distributions in Figure 4.9. Throughout the time period, we see an impressive accumulation of technological stock in Korea, which rapidly catches up with the technological leaders on the right-hand side. Japan has the highest technological capabilities and has experienced a reduction in dispersion, which means that differences in efficiency across sectors are shrinking. Finally, Argentina has experienced a mild increase in the stock of technology but a sharp reduction of θ (an increase in technological dispersion).

[18]Not all years are available for all countries

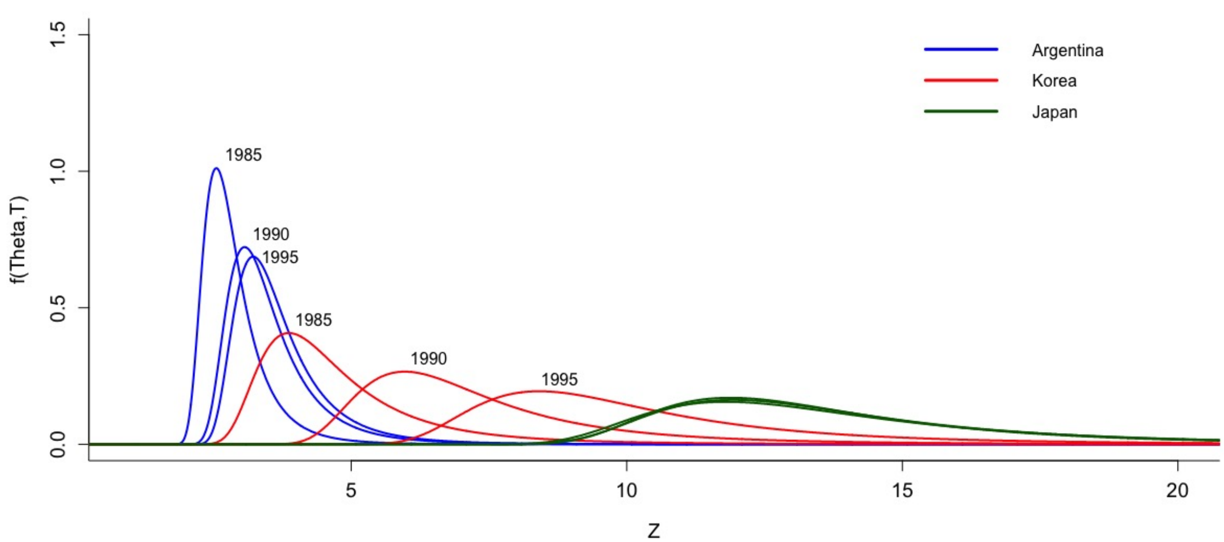
[19]See Eaton and Kortum (2002), Simonovska and Waugh (2011), and Leromain and Orefice (2014) for some examples.

Figure 4.8: T *and estimated* $\hat{\theta}$



Source: Own elaboration for 84 countries in the period 1980-2000.

Figure 4.9: T *and estimated* $\hat{\theta}$



Source: Own elaboration for based on the T_i and θ_i of Argentina, Korea, and Japan.

4.4 Estimation Results

To understand the role that the stock and distribution of knowledge play in bilateral trade, we estimate the gravity equation (4.7) derived from our model as follows:

$$\ln Ish_{ni,t} = \beta_0 + \beta_1 AA_{i,t} + \beta_2 WC_{ni,t} + \beta_3 CA_{ni,t} + \epsilon_{ni,t} \tag{4.8}$$

where Ish_{ni} denotes the bilateral imports of country i from country n as a share of i's total spending, AA_i is the absolute advantage (or technological stock) of the exporter, WC_{ni} is a world competitiveness index relative to i, and CA_{ni} is country i's comparative advantage when selling to n. Finally, t represents time (in years). All variables are as defined in equation (4.7) and have been calculated using the data described above.

Our model is a generalized version of Eaton and Kortum (2002), represented by the terms $AA_{i,t}$ and $WC_{ni,t}$. One advantage of this setup is that it allows for a straightforward comparison between the models, and thus, we can assess the relevance of the distribution of knowledge that enters through our term $CA_{ni,t}$. Moreover, this term represents our main contribution, as it introduces the two main aspects of classical Ricardian theory into the multi-country setup: Technological dispersion is country-specific, and it matters for bilateral exports. In our formulation, the latter takes the form of an augmenting (or dampening) trade effect that decreases the exporter's relative costs. Our theory predicts that β_2 is negative and β_1 and β_3 are positive. Bilateral exports from i to n decrease with i's trading costs (relative to the rest of the world's) and increase with i's technological stock and relative force of comparative advantage. In particular, as discussed above, country i will benefit more from a decrease in a competitor's force of comparative advantage (reduced technological dispersion) the cheaper i is relative to its competitors.

Table 4.1 reports the results of estimating equation (4.8) for over 2200 country pairs and 18 years, comparing the baseline EK model (first two terms) to our generalized (full) model across different specifications. To keep our econometric model closer to the empirical literature, we include the usual controls. Note that the economic interpretation of the comparative advantage coefficient is not straightforward. The CA term represents how relative costs covary with technological dispersion, and thus, we interpret its coefficient as the elasticity of exports with respect to the cost-dispersion covariance.

The results in columns (1) and (2) of Table 4.1 reveal that, after accounting for exporter fixed effects, all absolute advantage, world competitiveness, and comparative advantage terms matter for bilateral trade. All coefficients have the expected sign and are highly significant, which supports our model and suggests that the standard EK model was omitting a relevant determinant of exports. Columns (3) and (4) of Table 4.1 report the results of estimating both models after adding time fixed effects. The coefficient on absolute advantage decreases while the others remain virtually unchanged. Finally, columns (5) and (6) of Table 4.1 report the results of estimating both models after introducing the typical bilateral trade cost determinants from the gravity literature. These allow one to control for common shared characteristics that are constant over time, such as a common language, a shared border, colonial ties and a shared past. The full table of results (including these gravity coefficients) can be seen in the Appendix. The results are very similar to the previous specification and all the gravity variables are significant and have the expected signs. This is our preferred specification. Overall, our results suggest that, in line with the existing Ricardian literature, an increase in the overall stock of technology or a decrease in relative costs of exporter i increases its exports to n. In addition, we find that the comparative advantage term matters, supporting our augmented (full) model. This evidence suggests that changes in country-specific technological dispersion are important determinants of bilateral trade.

In Table 4.2, we show that our main finding is robust to alternative specifications and measures of costs. Thus far, we have followed Eaton and Kortum (2002) in using wages to measure costs. One concern with this measure is that wages only capture the labor component of production costs; therefore, we now turn to using the wholesale price index (WPI) as an alternative measure. Columns (1) and (2) of Table 4.2 report the results of estimating our preferred specification when using wages and WPI, respectively. All coefficients are similar in magnitude, have the expected signs, and are significant. Columns (3) and (4) of Table 4.2 report the estimation results after adding exporter-time fixed effects with both measures of costs. The comparative advantage term is absorbed by the fixed effects, but both the world competitiveness term and the comparative advantage term are statistically significant and have the right signs. Finally, we drop the bilateral controls and exporter-time effects and replace them with bilateral-pair fixed effects. Columns (5) and (6) report the results of estimating equation (4.8) using both measures of costs. As expected, the significance of the comparative advantage effect is diminished, but all results still hold when using wages as the cost measure.

Table 4.1: *Baseline Results*

	Dependent variable is bilateral imports (as a share of total spending)					
	(1)	(2)	(3)	(4)	(5)	(6)
	EK	full	EK	full	EK	full
Absolute Advantage	0.554***	0.594***	0.245***	0.258***	0.227***	0.238***
	(20.99)	(22.46)	(6.51)	(6.94)	(6.32)	(6.71)
World Competitiveness	-0.105***	-0.0960***	-0.109***	-0.0980***	-0.0784***	-0.0689***
	(-52.19)	(-42.01)	(-52.47)	(-41.45)	(-38.97)	(-30.77)
Comparative Advantage		0.0431***		0.0585***		0.0510***
		(10.06)		(11.64)		(10.94)
Exporter FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	No	No	Yes	Yes	Yes	Yes
Bilateral controls	No	No	No	No	Yes	Yes
Observations	19230	19230	19230	19230	19230	19230

t statistics in parentheses

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

Table 4.2: *Robustness*

	Dependent variable is bilateral imports (as a share of total spending)					
	(1)	(2)	(3)	(4)	(5)	(6)
Absolute Advantage	0.238***	0.136***	-	-	0.737***	0.602***
	(6.71)	(3.63)			(28.85)	(24.23)
World Competitiveness	-0.0689***	-0.0128***	-0.0709***	-0.0169***	-0.0154***	-0.00278**
	(-30.77)	(-8.83)	(-30.93)	(-9.29)	(-3.02)	(-2.04)
Comparative Advantage	0.0510***	0.0208***	0.0548***	0.0328***	0.0142**	0.00156
	(10.94)	(5.76)	(11.35)	(7.39)	(2.53)	(0.55)
Bilateral controls	Yes	Yes	Yes	Yes	No	No
Exporter-time FE	No	No	Yes	Yes	No	No
Bilateral pair FE	No	No	No	No	Yes	Yes
Cost measure	wages	WPI	wages	WPI	wages	WPI
Observations	19230	15171	19230	15171	19230	15171

t statistics in parentheses

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

4.5 Concluding Remarks

This chapter documents that there is a wide range of country-level technological dispersion as measured by patent data. In addition, this measure appears to be highly correlated with international trade flows. The literature, however, has managed to explain only the relationship between the technological stock and trade flows, thus ignoring the role of dispersion.

Building on Eaton and Kortum (2002), we have developed a Ricardian model that takes into account the role of absolute advantage and technological dispersion, both of which are country-specific. We have shown that one is not independent of the other: Theory predicts that their interaction is also important for explaining bilateral trade. In particular, the effect of increasing technological dispersion depends on the relative world standing of a given country, as measured by its technological stock. In the model, technological dispersion governs the advantage of having lower input costs. A lower dispersion benefits countries with lower input costs since their exports are determined by these and not technological differences. The effect is reversed for the countries with which they are competing. This observation results in the main prediction of the model: A given country's exports increase when competitors with low costs have high technological dispersion and competitors with high costs have low technological dispersion.

To test our theory, we rely on measures of technological variables that allow us to test the predictions of the model. These should represent absolute and comparative advantage, originate from outside of the model and be consistent with theory. To this end, we create a novel dataset of historical patents (the longest to date) using patent grants at the USPTO. We use the dataset to construct measures of the stock and dispersion of technology that are consistent with the theoretical model. We combine the patent data with data on exports, patents, input costs, income, expenditures, and bilateral pair characteristics for 84 developed and developing exporters in the period 1983-2000.

Our empirical results confirm our theoretical predictions: Both the stock and the dispersion of technology are important determinants of exports, and countries are affected by their relative world standing in terms of costs and technological stock and dispersion. This finding has important policy implications. The optimal strategy for a country that could change its technological profile depends on where it stands in the world scenario. Low-cost countries are better off competing on costs, while high-cost countries are better off competing on comparative advantage. Future work should focus on using this model to explain the diversification patterns in export behavior.

4.6 Appendix

4.6.1 Approximation

We use the approximation

$$\sum_k \delta_k z_i^{\theta_i - \theta_k} \approx 1$$

where δ_k is the share of country k in the index of world competitiveness relative to country i

$$\delta_k = \frac{T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right)^{-\theta_k}}{\sum_{j=1}^N T_j \left(\frac{c_j d_{nj}}{c_i d_{ni}}\right)^{-\theta_j}}$$

to integrate

$$\int_0^\infty e^{-\sum_{k=1}^N T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right)^{-\theta_k} z_i^{-\theta_i}} \theta_i T_i z_i^{-\theta_i - 1} dz_i$$

instead of

$$\int_0^\infty e^{-\sum_{k=1}^N T_k \left(\frac{c_k d_{nk} z_i}{c_i d_{ni}}\right)^{-\theta_k}} \theta_i T_i z_i^{-\theta_i - 1} dz_i$$

We want to bound the difference

$$\left| \int_0^\infty e^{-\sum_{k=1}^N T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right)^{-\theta_k} z_i^{-\theta_i}} \theta_i T_i z_i^{-\theta_i - 1} dz_i - \int_0^\infty e^{-\sum_{k=1}^N T_k \left(\frac{c_k d_{nk} z_i}{c_i d_{ni}}\right)^{-\theta_k}} \theta_i T_i z_i^{-\theta_i - 1} dz_i \right| < \varepsilon$$

Instead we will work with

$$\left| \int_{\underline{z}}^{\overline{z}} e^{-\sum_{k=1}^N T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right)^{-\theta_k} z_i^{-\theta_i}} \theta_i T_i z_i^{-\theta_i - 1} dz_i - \int_{\underline{z}}^{\overline{z}} e^{-\sum_{k=1}^N T_k \left(\frac{c_k d_{nk} z_i}{c_i d_{ni}}\right)^{-\theta_k}} \theta_i T_i z_i^{-\theta_i - 1} dz_i \right| < \varepsilon$$

which holds if

$$\int_{\underline{z}}^{\overline{z}} \left| e^{-\sum_{k=1}^N T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right)^{-\theta_k} z_i^{-\theta_i}} - e^{-\sum_{k=1}^N T_k \left(\frac{c_k d_{nk} z_i}{c_i d_{ni}}\right)^{-\theta_k}} \right| \theta_i T_i z_i^{-\theta_i - 1} dz_i < \varepsilon$$

wer impose the integration limits so that the integral does not blow up, but $\underline{z}$ can be made arbitrarily small. This holds if

$$\left| e^{-\sum_{k=1}^{N} T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right)^{-\theta_k} z_i^{-\theta_i}} - e^{-\sum_{k=1}^{N} T_k \left(\frac{c_k d_{nk} z_i}{c_i d_{ni}}\right)^{-\theta_k}} \right| < M_\varepsilon$$

where

$$M_\varepsilon \equiv \frac{\varepsilon}{T(\underline{z}^{-\theta} - \overline{z}^{-\theta})}$$

which holds if

$$\left| \frac{e^{-\sum_{k=1}^{N} T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right)^{-\theta_k} z_i^{-\theta_i}} - e^{-\sum_{k=1}^{N} T_k \left(\frac{c_k d_{nk} z_i}{c_i d_{ni}}\right)^{-\theta_k}}}{e^{-\sum_{k=1}^{N} T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right)^{-\theta_k} z_i^{-\theta_i}}} \right| < M_\varepsilon$$

or

$$\left| 1 - e^{\sum_{k=1}^{N} T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right)^{-\theta_k} z_i^{-\theta_i} - \sum_{k=1}^{N} T_k \left(\frac{c_k d_{nk} z_i}{c_i d_{ni}}\right)^{-\theta_k}} \right| < M_\varepsilon$$

This holds if

$$\left| 1 - e^{-\left|\sum_{k=1}^{N} T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right)^{-\theta_k} z_i^{-\theta_i} - \sum_{k=1}^{N} T_k \left(\frac{c_k d_{nk} z_i}{c_i d_{ni}}\right)^{-\theta_k}\right|} \right| < M_\varepsilon$$

or alternatively

$$-\ln(1 + M_\varepsilon) < \left| \sum_{k=1}^{N} T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right)^{-\theta_k} z_i^{-\theta_i} - \sum_{k=1}^{N} T_k \left(\frac{c_k d_{nk} z_i}{c_i d_{ni}}\right)^{-\theta_k} \right| < -\ln(1 - M_\varepsilon)$$

And a first order Taylor expansion gives

$$-\frac{\ln(1 + M_\varepsilon)}{\ln \underline{z}/\underline{z}^{\theta}} < \left| \sum_{k=1}^{N} T_k \left(\frac{c_k d_{nk}}{c_i d_{ni}}\right)^{-\theta_k} |\theta_k - \theta_i| \right| < -\frac{\ln(1 - M_\varepsilon)}{\ln \overline{z}/\overline{z}^{\theta}}$$

So the approximation will be better when θ_k is close to θ_i.

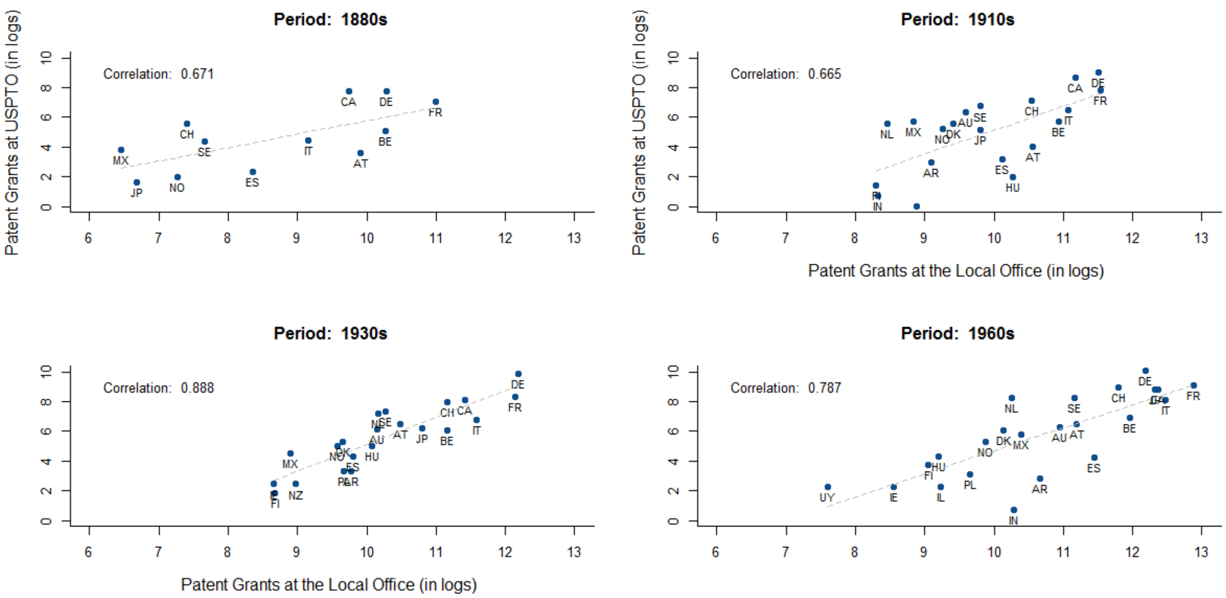
4.6.2 Data

We consider all countries with patent and trade data availability, with a few exceptions. We exclude communist and former USSR countries due to unreliable data and large historical gaps, tiny countries (population less than 500 thousand), and US territories due to the home bias effect. This yields a total of 84 countries.

Countries			
Albania	Egypt	Rep. of Korea	Portugal
Argentina	Spain	Lebanon	Paraguay
Australia	Finland	Sri Lanka	Senegal
Austria	France	Morocco	Singapore
Belgium	Gabon	Madagascar	El Salvador
Bangladesh	United Kingdom	Mexico	Suriname
Bulgaria	Ghana	Mali	Slovakia
Bahrain	Guinea	Mauritius	Slovenia
Bolivia	Greece	Malaysia	Sweden
Brazil	Guatemala	Niger	Syria
Canada	Guyana	Nicaragua	Thailand
Switzerland	Honduras	Netherlands	Trinidad and Tobago
Chile	Croatia	Norway	Tunisia
China	Hungary	Nepal	Turkey
Cameroon	India	New Zealand	Tanzania
Colombia	Ireland	Oman	Uganda
Costa Rica	Israel	Pakistan	Uruguay
Cyprus	Italy	Panama	Viet Nam
Czech Rep.	Jordan	Peru	Yemen
Germany	Japan	Philippines	South Africa
Denmark	Kenya	Poland	Zimbabwe

4.6.3 Additional Figures

Figure 4.10: *Grants at local office vs. USPTO (pre 1980)*



Source: Own elaboration based on USPTO and WIPO statistics.

4.6.4 Regression Results (Full Table)

Table 4.3: *Baseline Results - Full Table*

	(1)	(2)	(3)	(4)	(5)	(6)
Absolute Advantage	0.554***	0.594***	0.245***	0.258***	0.227***	0.238***
	(20.99)	(22.46)	(6.51)	(6.94)	(6.32)	(6.71)
World Competitiveness	-0.105***	-0.0960***	-0.109***	-0.0980***	-0.0784***	-0.0689***
	(-52.19)	(-42.01)	(-52.47)	(-41.45)	(-38.97)	(-30.77)
Comparative Advantage		0.0431***		0.0585***		0.0510***
		(10.06)		(11.64)		(10.94)
Shared border					1.673***	1.691***
					(21.33)	(21.57)
Common language					0.720***	0.728***
					(7.73)	(7.92)
Have had colonial links					0.266**	0.267**
					(2.89)	(2.87)
Common colonizer					0.794***	0.721***
					(8.40)	(7.74)
Are/were the same country					2.173***	2.172***
					(15.41)	(15.40)
Constant	-9.583***	-9.991***	-7.553***	-7.722***	-8.761***	-8.897***
	(-43.58)	(-44.56)	(-25.19)	(-25.98)	(-30.77)	(-31.46)
Observations	19230	19230	19230	19230	19230	19230
Adjusted R^2	0.572	0.575	0.576	0.580	0.627	0.630

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

CHAPTER 5

Unravelling the Trail of a GPT: The Case of Electrical & Electronic Technologies from 1860 to 1930

This chapter is single authored. I would like to thank all the participants at the Harvard Kennedy School CID Growth Lab seminar, U.S. Census Bureau seminar, University of Barcelona AQR seminar, the RSA 2017 conference, the Technical University of Eindhoven seminar, and the 2016 RIDGE Forum at the University of Buenos Aires for their useful comments. Additionally, I profited greatly from the comments of and interactions with P.A. Balland, R. Boschma, C. Esposito, A. Morrison, P. Moser, S. Perez, and M. Svarc.

5.1 Introduction

Technological change has marked the pace of socio-economic progress in recent Western history, growing at unprecedented and continuously increasing rates. It has been argued that episodes of acceleration in technological progress were driven by particular technologies. These technologies, given their revolutionary nature, have had the power to change the pace and direction of economic progress, as well as to transform the social and political structures surrounding them. Widely known examples are the steam engine and the electricity. Information and communication technologies (ICTs) are often mentioned as a contemporaneous example.

In economics, these revolutionary technologies are referred to as "General Purpose Technologies" (GPTs). GPTs are characterized by possessing a wide scope for continuous improvement and elaboration, on the one hand, and high complementarity on the other. The latter means that a GPT should be able to diffuse across a wide range of sectors, not only because it is used an input in many different products and processes, but also because it is a technological complement of existing and new technologies. These characteristics are what make GPTs "engines of growth".

While theoretical models have advanced greatly, providing a precise and coherent characterization of the economic implications of the diffussion of a GPT (Bresnahan and Trajtenberg (1995), Helpman and Trajtenberg (1998b) ,Helpman and Trajtenberg (1998a), and Aghion and Howitt (2000)), a lack of convincing and comprehensive empirical evidence has called into question the relevance and usefulness of the notion of GPTs (Field, 2008).

The main empirical challenges are, on the one hand, to find evidence of a GPT having a real impact on the economy, as theory provides clear predictions on how the diffusion of a GPT should affect growth and wages. Thus far, most of the evidence has been collected at an aggregated level (national), placing an immediate limitation on the possibility to univocally relate economic changes to the diffusion of a GPT. See, for instance, David (1990), Greenwood (1997), David and Wright (1999), Crafts and Mills (2004), Crafts (2004), and Jovanovic and Rousseau (2005). Additional evidence has been provided in the form of detailed historical accounts on the economic and societal changes generated by several GPT candidates throughout history (Lipsey et al., 2005). A valuable contribution toward providing evidence at a finer level of disaggregation is made by Rosenberg and Trajtenberg (2004). They use county-level information on the adoption of the Corliss steam engine (an alleged GPT) during the late nineteenth century and show that it had a positive effect on population growth.

A second empirical challenge consist of providing a measurable way of characterizing GPTs. As Lipsey et al. (2005) note, "if the concept of a GPT is to be useful, then GPTs must be identifiable". Although work has been done on this subject, no clear consensus has been reached. Hall et al. (2006) propose a series of indicators based on a group of selected patents granted by the United States Patent and Trademark Office (USPTO); however, these indicators are not able to fully portray ICTs in a way that is consistent with theory. Additionally, Moser and Nicholas (2004) use a similar set of measures to evaluate whether electricity matches the GPT criteria based on a sample of historical patents assigned to publicly traded companies in the 1920s. They reject the hypothesis that electricity was a GPT. Also using patent data, Feldman and Yoon (2012) argue that technologies of genetic material recombination[1] exhibit some characteristics of a GPT. Therefore, the empirical evidence remains inconclusive: either there is no particular technology that is capable of fulfilling the criteria established by theory, or current measures are not an appropriate means of identifying GPTs.

Therefore, there is a need for comprehensive empirical evidence on the effect of GPTs over the economy on the one hand, and a lack of consensus on how to empirically identify them on the other. This chapter addresses these two issues by combining economic and demographic data provided by the U.S Census Bureau and IPUMS (the Integrated Public Use Microdata Series) with a novel database containing detailed information on the geographical location, as well as the type of technological domain, of patents granted by the USPTO dating back to 1836.

This chapter begins by addressing the issue of whether or not Electrical & Electronic (E&E) technologies can be considered a GPT by creating a set of patent-based measures intended to capture the main characteristics of a GPT in the data. We use these measures to show that E&E technologies evinced GPT-consistent behavior between 1860 and 1930. This establishes the foundation

[1]https://www.google.com/patents/US4237224

for the main point of the chapter: finding evidence of a GPT (E&E technologies) having a real impact on the economy.

Using economic and demographic data in combination with patenting activity per county, we show that the adoption of E&E technologies had a positive effect on wages and the per capita income growth after the 1900s. The empirical strategy adopted here relies on using measures of the adoption of E&E technologies prior to the 1870s as an instrument to predict the adoption of E&E technologies between 1900 and 1930. It assumes that the early adoption of E&E technologies (prior to the invention of the electric bulb or the establishment of the first power plant) was determinant to the inventive structure of places 50 years later while not being correlated with the events that will determine growth and wages between 1900 and 1930. Jointly, these two results provide a comprehensive picture of the emergence, evolution, and diffusion of E&E technologies during this period and their effect on the economy.

The chapter is organized as follows: The next section describes the different sources of data used in this study, while Section 5.3 proposes and discusses the criteria to identify the main characteristics of a GPT using patent data. After showing that E&E technologies fulfill these criteria, Section 5.4 develops and evaluates an empirical strategy to show that the adoption of E&E technologies had a positive impact on output and wages after the 1900s. Section 5.5 concludes the chapter.

5.2 Data Sources

This section describes the different sources of data used in this study, which when combined allow for a comprehensive overview of the emergence, evolution, development, and diffusion of E&E technologies in their historical context. Empirical studies on the evolution of technologies are often limited in quantity and scope by the availability of data and the nature of the object of study. After all, the diffusion of technologies takes decades, and the most interesting cases occurred long before data began to be collected systematically. For instance, the Corliss steam engine discussed in Rosenberg and Trajtenberg (2004) was patented in 1849[2], while Edison's carbon filament incandescent lamp mentioned in David (1990) dates back to 1880[3]. This is also true for the invention of integrated circuits, allegedly the engine of the ICT revolution, which can be traced back in the U.S. to 1959[4]. In this study, we focus on the period from 1860 to 1930, which is considered to cover the emergence, development, and diffusion of E&E (David, 1990; Goldfarb, 2005; Greenwood, 1997; Lipsey et al., 2005). we overcome data limitations by merging several independently developed datasets, which contain information about the technological and geographical attributes of patented inventions in the U.S., as well as county-level economic and demographic data.

First, we use information about the technological class along with the full description of patent documents made available by the USPTO[5]. In 2006, the USPTO entered into a series of agreements with Reed Tech and Google to digitalize all available patent documents, thereby making historical patent data available in bulk form. These bulk data contain ZIP or TAR files with TIFF or PDF images, concatenated XML or structured ASCII files with all available information in patent documents dating back to 1836. This means that for every patent document ever granted since 1836 it is possible to identify its technological class and access the full description of the invention.

Patents are classified into technological classes according to the type of invention to which they claim rights. There are currently more than 400 different technological classes in use, and whenever a new class is created, or an existing

[2]See https://www.google.com/patents/US6162
[3]See https://www.google.com/patents/US223898
[4]See https://www.google.com/patents/US3138743. The first integrated circuit is attributed to Werner Jacobi (Siemens AG) in 1949 (https://www.google.com/patents/DE833366).
[5]http://www.uspto.gov/learning-and-resources/electronic-bulk-data-products.

one re-defined, all available patents are re-classified to maintain temporal consitency. Furthermore, patents can be grouped into broad economically relevant categories (Chemical, Computer and Communications (C&C), Drugs and Medical (D&M), Electrical and Electronics (E&E), Mechanical, and Others)[6]. During the period considered C&C and D&M technologies represented approximately 1% of total patenting activity, meaning that they were not treated as separate categories but placed within the category Others[7]. Table 5.1 below shows the distribution of patenting activity over time across these broad categories[8].

Table 5.1: *Distribution of Patenting Activity Over Time*

	Chemical	E&E	Mechanical	Others
1850	10.17	1.070	39.630	49.130
1860	9.560	1.860	35.480	53.100
1870	9.490	1.640	33.620	55.250
1880	8.160	3.290	34.480	54.070
1890	6.280	6.490	38.460	48.770
1900	8.920	5.330	37.940	47.820
1920	8.000	7.780	40.340	43.880
1930	12.09	9.640	36.140	42.130

There are considerable differences among categories. Mechanical and Chemical technologies are amongst the most abundant types, which is a consequence of them being at a more advanced stage of maturity than E&E technologies. Note that while Mechanical, Chemical, and Others exhibited stagnant or decreasing participation over this period, E&E showed a marked and steady increase in their share. In fact, E&E technologies exhibited a nearly 10-fold increase in activity

[6] See Hall et al. (2001b) for details. The concordance is available at http://www.nber.org/patents/.

[7] Our results do not differ if they are considered as separate categories.

[8] Patents can be assigned to more than one technological class or broad category depending on the scope of the claims made. In this table only the main technological classification is considered.

during this period, going from representing 1% of all inventive activity to nearly 10%.

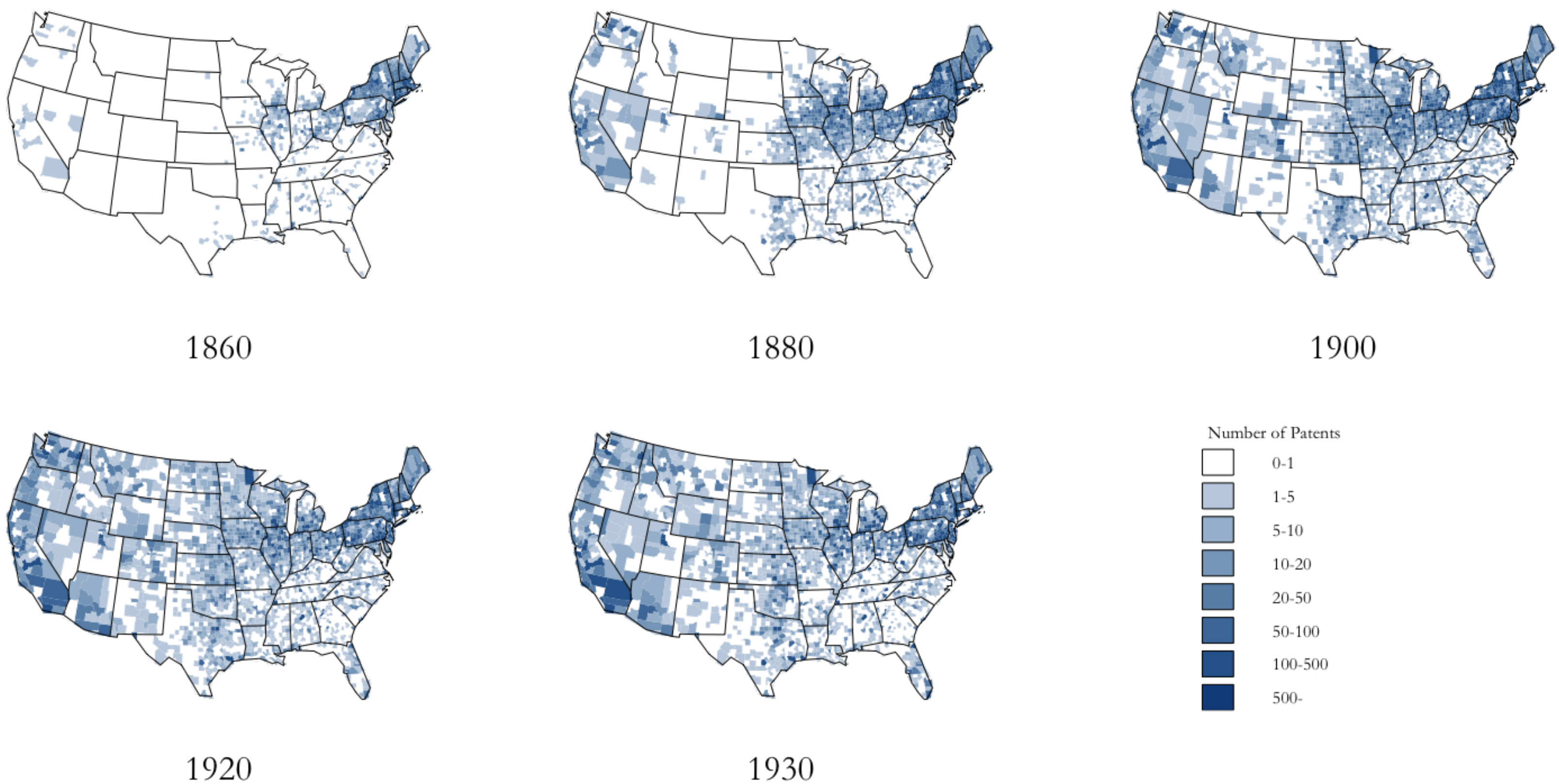
These broad categories differ also in terms of the number of technological classes of which composed. There are more than 80 different technological classes within Chemical, approximately 120 within Mechanical, and 54 in E&E technologies. Others account for approximately 180 classes. Information about technological classes has been widely used in empirical studies, usually to create measures of the technological diversity of places (Boschma et al., 2014) and the generality of particular inventions (Feldman and Yoon, 2012; Hall et al., 2001b, 2006; Moser and Nicholas, 2004).

The second source of data comes from the novel data initiative in Petralia et al. (2016), which provides county-level information on the location of the inventor(s) and/or assignee(s) for most patents granted since 1836[9]. This permits researchers to track the diffusion of inventions in space. This database, HistPat, was built using optically recognized and publicly available patent documents at USPTO, combining text-mining algorithms with a statistical model to identify locations[10]

Figure 5.1 below describes the diffusion of patenting activity over space and time. It shows how it first spread from the north-eastern states, where most inventive activity was concentrated before the 1860s, toward the south-east to later reach the west coast by the 1880s. Patent activity closely followed the development of urban centers and migration (Ager and Brückner, 2013; Burchardi et al., 2016). The central states joined the technological race by the turn of the century.

[9]The data can be downloaded at https://dataverse.harvard.edu/dataverse/HistPat
[10]The entire procedure is documented in Petralia et al. (2016).

Figure 5.1: *The Geography of Patenting Activity over the Years*



Source: Own elaboration based on HistPat

The geographical diffusion of inventive activities was quite heterogeneous across technologies, as shown in Table 5.2. This table displays the correlation of patenting activity across different categories throughout the entire period, where each vector takes value one whenever patenting activity in a county was found and zero otherwise[11]. Note that this heterogeneity in the geographical pattern of diffusion provides an excellent opportunity to exploit time-place differences in the adoption of technologies. This can be used to evaluate the relationship between the evolution of economic outcomes and the diffusion of different technologies.

Table 5.2: *Geographical Correlation in Patenting Activity*

	E&E	Mechanic	Chemical	Others
E and E	1	0.359	0.436	0.335
Mechanic	0.359	1	0.454	0.596
Chemical	0.436	0.454	1	0.432
Others	0.335	0.596	0.432	1

This brings us to the next data source used in this chapter. The county-level technological development described above can be combined with economic and demographic data provided by the U.S. Census Bureau and IPUMS. In particular, we use the database provided by the Inter-University Consortium for Political and Social Research (ICPSR), which contains detailed decennial county-level data on demographic, economic, and social variables that were collected by the U.S. Census Bureau[12]. In Addition, we use data on occupations provided by IPUMS[13].

Table 5.3 below shows the evolution of the main economic and demographic variables in the sample. Note that this period exhibited rapid and continuous

[11]All counties are included in the sample. If only counties with positive patenting activity are considered, the correlations are considerably lower.
[12]Downloadable at https://www.icpsr.umich.edu/icpsrweb/
[13]See https://usa.ipums.org/usa/intro.shtml

growth of population, which had quadrupled since 1860. Immigration was certainly one of the main explanatory factors; by 1880, more than 10% of the population was foreign-born, with this number rising to 20% if second-generation immigrants are included (Burchardi et al., 2016). Migration was mostly unregulated until World War I, after which a series of restrictions were incorporated, ultimately resulting in the establishment of a quota system in 1921[14].

Table 5.3: *Evolution of Main Census Variables*

	1860	1870	1880	1890	1900	1920	1930
Population (in millions)	31.409	38.542	50.150	62.610	75.726	105.967	123.143
Foreign Population (in millions)	4.130	5.562	6.677	9.246	10.428	13.713	13.366
Labor Force (in millions)	1.276	2.054	2.733	4.710	5.316	9.056	8.751
Output (in billions)	1.753	2.534	4.104	8.080	12.126	24.261	32.764
Output per Capita	55.827	65.740	81.825	129.049	160.130	228.950	266.061
Wages (share)	0.206	0.306	0.231	0.282	0.192	0.432	0.352

Notes: Output corresponds to ouput in manufactures at 1850's constant prices. Wages are expressed as a share of the total output, while Labor Force counts hands employed in manufacture.

The remarkable increase in manufacturing production during this era was crucial for the positioning of the U.S. as a world leader, which was realized after the 1900s when the U.S. surpassed Great Britain in terms of world share and per capita levels of total manufacturing output (David, 1990). Additionally, Field (2006) estimates that in the 1920s, manufacturing alone explained more than 80% of total factor productivity (TFP) growth.

There were many factors contributing to this remarkable success. For instance, Wright (1990) argues that it involved the greater exploitation of U.S. geological potential after examining the factor content of trade in manufactured

[14]Daniels (1990) provides a detailed description of the nature and composition of immigration flows during this period.

goods, while Ager and Brückner (2013) relates higher growth with immigration and the cultural diversity of places. Additionally, (Acemoglu et al., 2016) suggests that part of the exceptional technological dynamism in this period can be explained by an immensely capable and effective state. The existence of these explanatory factors lead to the inclusion of two more variables that were collected independently of previous sources. The first is the number of post offices per county, to proxy for state presence as in Acemoglu et al. (2016)[15]. The second counts the number of working mineral deposits within counties using geo-located information provided by the United States Geological Survey (USGS)[16].

Because these variables only play a marginal role in the econometric specification (acting as controls), and their construction did not involve any methodological challenges, we do not provide any further description of them. Additional details can be found in the appendix.

If, as Field (2003) notes, this was one of the most innovative periods in U.S. history, is it possible that E&E technologies were the engine of it? The next Section begins by discussing a way to identify the main characteristics of a GPT in the data to later conclude that E&E technologies exhibited unusual dynamism, which is consistent with the expected behaviour of a GPT. This forms the foundation for Section 5.4, which tests the impact of E&E adoption on economic outcomes.

5.3 Toward a characterization of GPTs

The aim of this section is to provide a means of identifying the main characteristics of a GPT in patent data. Given the amount of information and the level of detail contained in patent documents, it is natural to begin looking for ways of characterizing GPTs using patent data. Every patent provides information on

[15]Original records can be found at https://catalog.hathitrust.org/Record/002137107
[16]Available here: https://mrdata.usgs.gov/mrds/about.php

the technological nature of the invention, the geographical location of the inventor, and the prior art, among other things. This implies that one could identify whether a patent has claimed, for instance, rights to the invention of a new electrical device, a new function for a chemical compound, or both. This means that a patent can claim rights to different types of components or technologies that have been created and combined to produce a product.

Even though there is no agreement on how to measure GPTs, there is a clear understanding on what defines them. According to Helpman and Trajtenberg (1998b), Helpman and Trajtenberg (1998a), Moser and Nicholas (2004), Lipsey et al. (2005), and Jovanovic and Rousseau (2005), GPTs must have the following characteristics:

1. **Wide scope for improvement and elaboration.** They should be able to go through a process of technical advance after they are first introduced, a continuous process of technological improvement.

2. **Potential for use in a wide variety of products and processes.** They should spread across and be used in most sectors.

3. **Strong complementaries with existing or potential new technologies.** They should have an impact on existing and new technologies, not only by creating the need to alter and combine many of the existing technologies but also by increasing the opportunities to develop new technologies in combination with them.

Previous studies have not been able to identify the alleged GPT-characteristics for electricity and ICTs using patent data (Moser and Nicholas (2004) and Hall et al. (2006), respectively). If the wealth of information contained in patent documents does not allow one to portray a theory-consistent picture for two of the most commonly mentioned examples of GPTs, then we might well question the criteria used to identify GPTs can ever be fulfilled or to what extent patent data represent a useful instrument.

One of the main limitations imposed by patent data is the inability to go far enough back in time to cover periods that include the emergence, development, and diffusion of technologies. This is because patent documents began to be digitalized systematically in 1975. However, the use of citations has been a common approach to overcome this limitation, as they date farther back in time and link present and previous inventions (in some instances, for more than one hundred years). Moser and Nicholas (2004) and Hall et al. (2006) use citation-based measures to address the question of weather electricity and ICTs, respectively, behaved as GPTs.

The possibility of tracing back in time the knowledge embodied in patents through citations relies on two assumptions: first, that direct citations provide a comprehensive picture of the type of knowledge contained in a patent and, second, that the dynamics of patent citations are sufficiently invariant that the knowledge composition of patents survives over time.

In this chapter we use contemporaneous patent-related measures to identify the alleged characteristics of GPTs. The scope of the data allows us to have a comprehensive view of the emergence, evolution, development, and diffusion of technologies in their historical context. It focuses on the most iconic example of a GPT, electricity. In what follows, we propose three patent-related indicators that can be used to identify these characteristics[17].

[17]Lipsey et al. (2005) argues that there is a fourth condition that any GPT must satisfy, i.e. having a variety of uses. This refers to the number of distinct uses that are made of a single technology. He emphasizes that having a variety of uses is not synonymous with being widely used (see Section 5.3.2). In the particular case of electricity, he makes the convincing argument that it fulfills this requirement. Note that electricity can be used, for example, as a power source, for illumination, and as a means of communication. For this reason, this issue is not explored further.

5.3.1 Wide Scope for Improvement and Elaboration

As the literature suggests, GPTs should be able to go through a continuous process of technological improvement. This notion is based on the fact that most technologies are originally introduced as unrefined versions of their best self. What distinguishes a GPT is the distance to this most efficient, mature version of itself, which entails developing and perfecting it for its many uses, as well as adapting it to a wide arrange of complementary and yet potentially unrelated technologies.

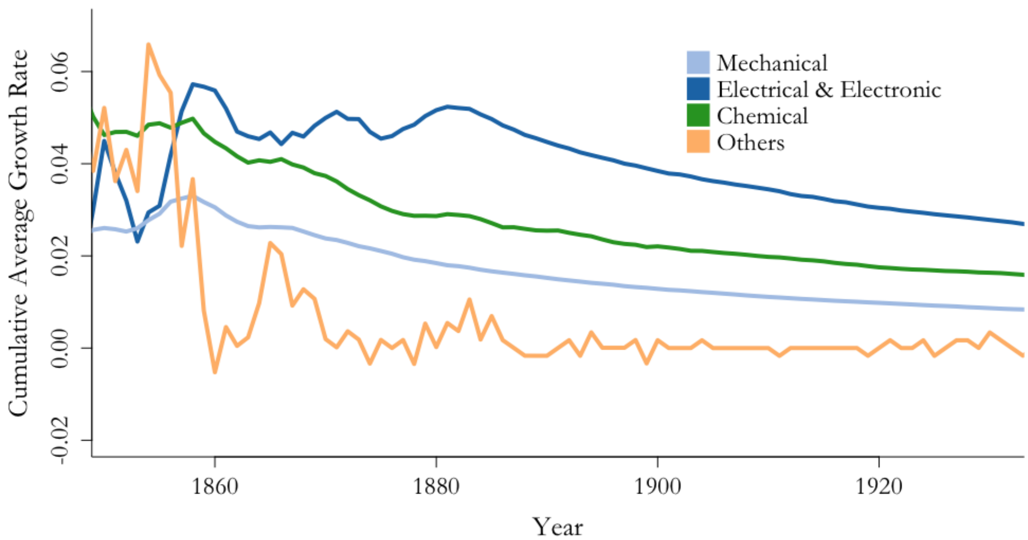
This is probably the least challenging characteristic to relate to data. Previous empirical approaches have often used patent growth to measure the extent to and pace at which technologies have been advancing. For instance, Jovanovic and Rousseau (2005) examine the growth rate of total patenting activity in the U.S and relate changes in its pace to the electric and ICT era. Moser and Nicholas (2004) demonstrate that patent activity in the category of E&E technologies grew the fastest in the 1920s. Similarly, Hall et al. (2006) find that classes related to C&C technologies to grew faster than others after the 1980s.

Here, we also consider growth in patenting activity: in particular, we use the Cumulative Average Growth Rate (CAGR) of patent production per technological category:

$$CAGR_{(t,j)} = \left(\frac{P_{(t,j)}}{P_{(t_0,j)}} \right)^{\frac{1}{t-t_0}} - 1$$

where $P_{(t_0,j)}$ and $P_{(t,j)}$ denote the number of patents produced in the initial year and at time t, respectively, for category $j \, \epsilon \, \{E\&E, Mechanical, Chemical, Others\}$. This indicator measures the geometric progression ratio of patenting activity, providing a constant rate of growth over the period.

Figure 5.2: *CAGR of Patent Production per Category*



Source: Own elaboration based on USPTO Patent Data

Figure 5.2 below displays the CAGR of patenting for every broad category since 1850.[18] It shows that the number of E&E patents granted at the USPTO grew the fastest during the second half of the nineteenth century and up to the end of our period (1930). The results are in line with previous findings, regarding the prominent role of E&E technologies after the 1850s.

The above-average growth in patenting activity for E&E technologies is also evident when looking at Table 5.1 in Section 5.2. This table shows that E&E technologies increased their share by an order of magnitude since 1850, from representing 1% of all patenting activity to nearly 10% in 1930.

[18]The initial year is 1840 because this is when every broad category started to show positive patenting activity.

5.3.2 Potential for Use in a Wide Variety of Products and Processes

It is argued that as GPTs evolve and develop they should spread throughout the economy, given their potential to be used as an input in many different applications. For example, electricity as a power source diffused through a wide range of sectors. It is used for household appliances and in transportation technologies, as well as to power a varied number of industries. Additionally, the ability of electricity to drive chemical reactions, as well as to make digital information processing possible, drastically expanded its range of uses.

Several approaches have been used to evaluate the pace and span of the diffusion of electricity throughout the economy. One is to consider the overall, nationwide, electrification of factories and households. David (1990) documents that the electric power used for mechanical drive capacity in the U.S. reached more than 50% by 1920, while Goldfarb (2005) and Duboff (1979) find that by 1929, the ratio of electric motor power to total motor power reached 82%, on average. Jovanovic and Rousseau (2005) show that by 1929, nearly 70% of households had electrical connections.

Regarding patent data, Moser and Nicholas (2004) and Hall et al. (2006) rely on patent citations to measure how wide the range of applicability of E&E and C&C technologies were, respectively. They use the technological diversity of citing patents to evaluate the generality of any cited patent. Therefore, the generality of a patent depends on how technologically heterogeneous its citing patents were.

We consider a different approach, which exploits the wealth of information contained in patent documents and provides a characterization of E&E technologies in their historical context. Note that all patents provide a detailed description of the invention, which can be scanned to identify keywords related the use of electronic and electrical components, notions, or principles (see the appendix to find the list of words). Therefore, patents that match these criteria but do not belong to the category of E&E can be considered 'users of electricity'. This set of patents will typically include inventions that use electricity-related terms because they rely on electricity-related components, notions, or principles but do not produce any particular technological improvement in that area.

Figure 5.3 can be used as an example. It shows the first page of the patent number 2,956,114 assigned to Ampex Corporation in 1960 for a broad band magnetic tape system (tape recorder). This particular patent falls under the category of C&C technologies, and cites patents only in C&C and Mechanical. This implies that by using a citation-based method, or the standard technological classification, we are not able to identify that the invention uses electrical components.

Examining the patent description, however, provides enough clues for a text mining algorithm to detect the electrical nature of it. Figure 5.3 highlights electricity-related words that are used to identify that this invention uses E&E technologies as inputs.

After identifying which inventions are "users" of E&E technologies, it is therefore possible to evaluate how wide the variety of technological products and processes using E&E technologies is. The most straightforward way to do this consist of counting the number of different technological categories that are users of E&E technologies at any point in time, which provides a contemporaneous measure of the pervasiveness of E&E technologies as inventive inputs.

Figure 5.4 below shows the share of technological classes in which E&E-related vocabulary has been found. This share is calculated with respect to the total number of different technological classes available at any point in time, which are not E&E. For instance, Figure 5.4 shows that by 1850, fewer than 5% of all non-E&E technological classes were using E&E-related vocabulary to describe inventions. By the end of the 1930s, this share had increased to more than 80%.

Figure 5.3: *Ampex Broad Band Magnetic Tape System (1960)*

United States Patent Office

2,956,114

Patented Oct. 11, 1960

1

2,956,114

BROAD BAND MAGNETIC TAPE SYSTEM AND METHOD

Charles P. Ginsburg, Los Altos, Shelby F. Henderson, Jr., Woodside, Ray M. Dolby, Cupertino, and Charles E. Anderson, Belmont, Calif., assignors to Ampex Corporation, Redwood City, Calif., a corporation of California

Filed July 25, 1955, Ser. No. 524,004

8 Claims. (Cl. 178—6.6)

This invention relates generally to electromagnetic tape systems, methods and apparatus, particularly to systems and methods of this character capable of recording and reproducing signal intelligence over a wide frequency spectrum, including for example, video frequencies.

Various problems are involved when it is attempted to record and reproduce frequencies over a wide spectrum, as for example frequencies ranging higher than one megacycle, on magnetic tape. Assuming the use of reasonable tape speeds, conventional equipment is limited with respect to its usable frequency range. The recordable range can be increased by increasing the speed of the tape, but the speeds required for the recording of such high frequencies are such that the system becomes impractical because of the large amount of tape employed for a given recording period. It is possible to reduce the linear tape speed by recording successive tracks extending laterally across the tape. Equipment with this purpose involves the use of magnetic record units which are mounted to sweep successively across the coated surface of the tape while the tape is being advanced in the direction of its length. While this arrangement makes it theoretically possible to provide relative speeds such that frequencies up to four megacycles or higher can be recorded, its application necessarily involves a number of problems. For example the outputs of the several heads are subject to amplitude variations, due to various causes such as lack of exact registration on the recorded track, amplitude variations in the record because of slight variations in pressure between the several heads, and slight variations in the electrical characteristics of the heads. The conventional magnetic tape recording system, using currents varying in amplitude for application to the recording head, is particularly susceptible to undesired amplitude variations. The undesired signal variations cause distortion of the reproduced signal, and make it difficult if not impossible to reproduce the original frequency spectrum with reasonable fidelity, and particularly with sufficient fidelity to permit the recording and reproduction of television or like visual images.

The present invention is predicated upon certain discoveries which we have made, and which we utilize to advantage in the present invention. Particularly we have found that a wide frequency spectrum can be successfully recorded and reproduced by the use of a frequency modulation system in which the deviation of the carrier is small relative to the highest frequency components to be transmitted. In other words we have found that it is practical to use what can be referred to as narrow band F.-M. Narrow band F.-M. means that the ratio of $\Delta f/f_m$ is relatively small, and in actual practice can be of the order of 0.2, where Δf represents deviation corresponding to maximum signal amplitude and f_m represents the highest modulating frequency. Likewise we have found that the limit of f_m can be made reasonably close to the carrier frequency. We have also discovered that the center or carrier frequency can be so selected that it is near the upper recordable frequency limit of the apparatus, which

2

as previously explained is generally determined by the relative speed between the heads and the tape and the characteristics of the head. When the carrier frequency is so selected the recording system depends upon single sideband or vestigial sideband transmission. In other words the upper band of frequencies containing the frequency modulation components is not recorded or reproduced to any substantial extent. We have found that such a magnetic record can be reproduced to provide, after demodulation, the original modulating frequencies with a good degree of fidelity.

In addition to the foregoing, a practical system for the recording and reproduction of frequencies over a wide spectrum requires highly accurate speed control means for both recording and reproduction.

It is an object of the present invention to provide a system and method for the recording and reproduction of a wide frequency band, which will be relatively immune to spurious variations in signal strength.

Another object of the invention is to provide a system and method of the above character which, when used for the recording and reproduction of video frequencies, makes possible the reproduction of visual images with good fidelity.

Another object of the invention is to provide a system and method of making use of narrow band frequency modulation for recording over a wide frequency band.

Another object of the invention is to provide improved means for controlling the speed of operation of various parts during recording and reproduction.

Another object of the invention is to provide a system and apparatus for the recording of frequency components over a wide spectrum, such as video frequencies, which utilizes a plurality of record heads sweeping laterally across a magnetic tape, but without causing troublesome distortion or disturbances of the reproduced signal due to amplitude variations.

Additional objects and features of the invention will appear from the following description in detail in conjunction with the accompanying drawings.

Referring to the drawings:

Figure 1 is a circuit diagram illustrating a complete recording and reproducing system incorporating the present invention.

Figure 2 is a circuit diagram illustrating a modification of Figure 1.

Figure 3 is a plan view schematically illustrating mechanism for mounting the magnetic heads and for transporting the tape.

Figure 4 is a cross sectional view taken along the line 4—4 of Figure 3.

Figure 5 is a cross sectional detail taken along the line 5—5 of Figure 3.

Figure 6 is an enlarged cross sectional detail illustrating the guide means for the tape and the manner in which the tape is contacted by the magnetic heads.

Figure 7 is an enlarged detail illustrating means for engaging the lower edge of the tape, while it is passing through the guide means.

Figure 8 is a cross sectional detail taken along the line 8—8 of Figure 3, and showing suitable pulse generating means.

Figure 9 is a schematic view illustrating the pulse generating means and the cathode follower which may connect to the same.

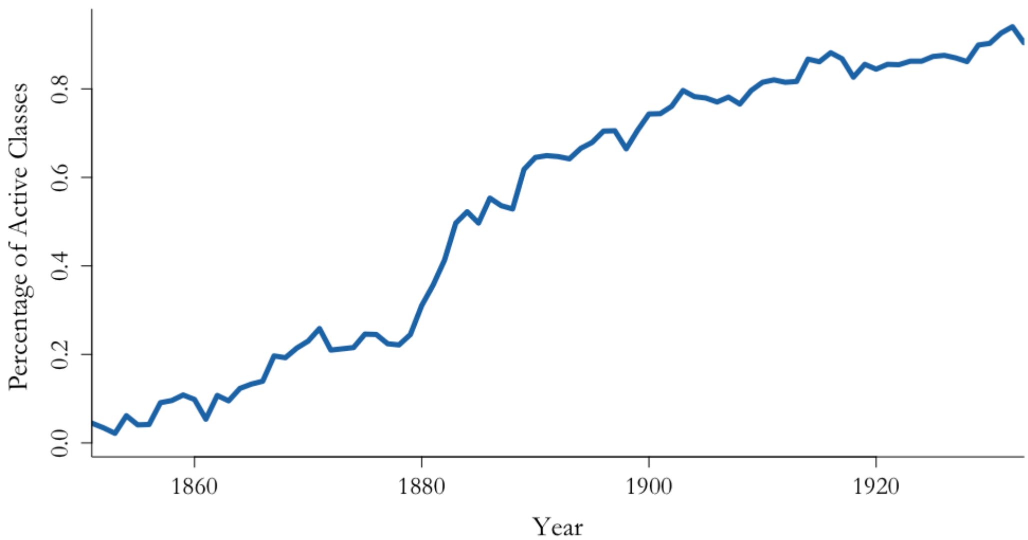
Figure 10 is a circuit diagram schematically illustrating the commutating means for making connections with the various magnetic heads.

Figure 11 is a diagram like Figure 10, but showing simplified connections.

Figure 12 is a plan view schematically illustrating a

Source: USPTO (Patent Number 2.956.114)

Figure 5.4: *Diffusion of E&E-Related Vocabulary*



Source: Own elaboration based on USPTO Patent Data

Figure 5.4 shows the impressive dissemination of E&E-related vocabulary throughout the spectrum of available technologies. It increased from representing a negligible share in 1850 to being present in approximately 90% of all non-E&E technological classes. This result points toward the wide variety of technologies using E&E technological components, principles, and notions as inputs in the innovation process. This suggests that electricity pervaded the whole inventive structure, affecting the entire nature of the technological production since its introduction.

If previous evidence showed that the use of electricity as a power source pervaded households (Jovanovic and Rousseau, 2005) and the productive structure of the U.S. economy (David, 1990; Goldfarb, 2005), Figure 5.4 adds a new dimension. It shows that electricity also had a simultaneous, and equally pervasive effect throughout the U.S. inventive structure. It changed the nature of technological advance, diffusing throughout the entire space of technologies.

5.3.3 Strong Complementaries with Existing and New Technologies

It is argued that GPT's can be considered "enabling technologies": this is because they provide a vast number of opportunities to adapt and modify existing products, processes, and organizational technologies. They expand the space of possible inventions and innovations, creating opportunities to develop new products, processes and technologies in combination with them. For instance, Bresnahan and Trajtenberg (1995) note that *"...the productivity gains associated with the introduction of electric motors in manufacturing were not limited to a reduction in energy costs. The new energy source fostered the more efficient design of factories, taking advantage of the new found flexibility of electric power."*

The far-reaching extent of its "innovation complementarity" (IC) is one of the most salient aspects of a GPT, as it is considered to be responsible for the creation and reinforcement of rapid technical advance and economic growth. Even though there is a vast literature collecting case-specific historical evidence (Bresnahan, 2010; David, 1990; DuBoff, 1979; Goldfarb, 2005; Helpman and Trajtenberg, 1998a; Lipsey et al., 2005; Rosenberg, 1998; **?**), there has not been any systematic and comprehensive empirical study on this subject.

This subsection proposes a way of measuring the extent of any technology's IC by examining the co-occurrence of different classes within patent documents. Whenever a patent is issued, several claims are made regarding the inventiveness and scope of the patent. These claims specify all the inventions contained within a particular patent for a product or process and are classified according to their technological characteristics into technological classes. Therefore, a patent can be classified into different technological categories, meaning that for this final product or process to work, inventions in different fields had to be realized.

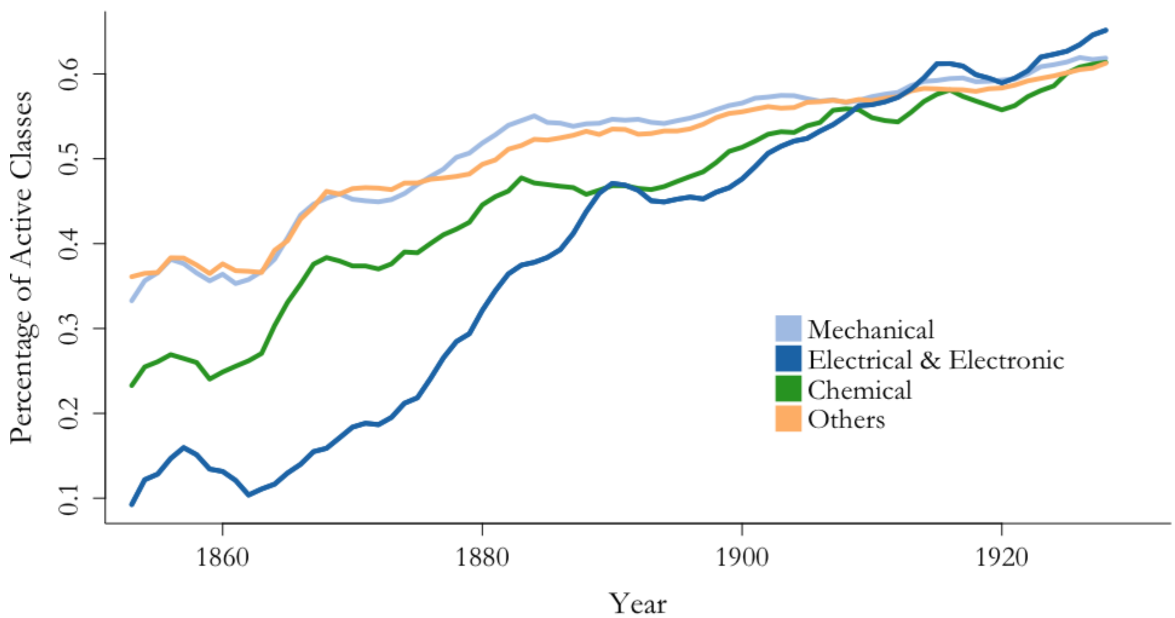
The extent and diversity of a technology's IC can be measured by examining the diversity of its co-occurrence profile.

For instance, consider the Ampex broad band magnetic tape system (1960) described in Section 5.3.2. It has claims in two different technological classes, class 360 (Dynamic Magnetic Information Storage or Retrieval) in C&C and class 386 (Motion Video Signal Processing for Recording or Reproducing) in E&E. This is because the patent introduces two main improvements: the first concerns a more efficient way of comprising and recording frequencies (class 360). For this improvement to be properly used, higher precision in the speed of the recording system needed to be achieved. This is when improvements in E&E technologies related to components regulating motion for recording systems had to be developed (class 386).

Therefore, the co-occurrence of technological classes within a patent can be used to measure ICs among different technologies. A GPT should co-occur with a wide variety of different technologies (classes); this is because its generality allows it to be re-combined with existing technologies to improve existing products (such as tape systems), as well as to develop new-to-the world and yet complementary technologies.

Figure 5.5 below measures the ICs of all the main technological categories by examining the number of classes that these technologies co-occur with, outside their own category. To avoid taking into account irrelevant, proximate combinations, co-occurrences of different classes within the same category are not counted. Thus, for instance, if a class within chemicals co-occurs with another class in the same category, then it is not considered an IC. This figure provides a clear and straightforward message: E&E technologies remarkably increased the variety and scope of their ICs since their introduction. They began as a very narrow technology and later became the most complementary technology by the beginning of the 1920s.

Figure 5.5: *Measuring Technologies' IC*



Source: Own elaboration based on USPTO Patent Data

To summarize, this section has addressed the issue of measuring empirically the main characteristics of GPTs. Using patent data, it first showed that E&E technologies were the fastest growing. Second, using the information contained in patent descriptions, it documented the pervasiveness of the use of E&E technologies as inputs in the innovation process. Finally, E&E technologies grew like no other category in terms of ICs. This meant that by the beginning of the 1920s, E&E technologies fulfilled the three main characteristics of a GPT. The timing and dynamics of these results are in line with previous evidence, which argues that the transformative power of electricity did not acquire momentum until after 1910s (David, 1990; Field, 2008; Greenwood, 1997; Lipsey et al., 2005).

The next section turns to evaluating whether E&E technologies, as has been argued, had a real transformative effect on the economy. To do so, we use county-level data on the geographical location of inventors and patentees for the period of the emergence and spread of E&E technologies (1860-1940) along with county data from the U.S Census Bureau on different economic indicators.

5.4 A GPT at Work

The previous section explored a way to empirically measure the key characteristics of a GPT, providing a means of identifying and distinguishing such technologies that proved effective in the case of E&E technologies. It remains to be answered, however, whether such a technology can have the alleged transformative effect on the economy. This section moves in that direction and evaluates whether the diffusion of E&E technologies between 1860 and 1930 affected output and wages as predicted by theory.

Theoretical models yield a clear set of stylized facts regarding the impact of a GPT. The diffusion of a GPT should eventually increase real output and and real wages as it becomes widely used. At early stages of the diffusion process however, while sectors adapt and develop complementarities, there may be periods of intermittent or even negative growth. (Aghion and Howitt, 2000; Helpman and Trajtenberg, 1998a,b)

There are two widely documented characteristics of technological diffusion that should be taken into account. First, technological change can take several decades to spread (Griliches, 1957), even in the case of revolutionary technologies (David, 1990). Second, processes of technological diffusion tend to be highly dependent on geography, as physical proximity and collocation are crucial in the dynamics of innovative processes (Audretsch and Feldman, 1996; Balland and Rigby, 2017; Feldman and Kogler, 2010; Feldman and Yoon, 2012; Jaffe et al., 1993). This effect is expected to be stronger for relatively new and unfamiliar technologies, such as electricity during the second half of the nineteenth century.

Even though theoretical models characterize the adoption of a GPT in a geographically deprived setup, we can profit from the uneven geographical spread of different technologies when empirically assessing their worth. Their unique place and time profile of adoption provides an exceptional opportunity to exploit time-place variation in the adoption of technologies, which is key in determining the realization of their benefits.

We trace the adoption of different technologies over space and time using the patenting activity of counties, assuming that the patenting activity of a place can be used as a suitable empirical counterpart of its capabilities to produce or use any given type of technology. This is equivalent to stating, in terms of theoretical models, that finding patenting activity represents having reached the component-finding stage of the Helpman and Trajtenberg (1998a) model or the discovery phase in Aghion and Howitt (2000).

It is to be expected that engaging or adopting the new GPT should have a positive overall effect on real wages and output, meaning that irrespective of the intrinsic characteristics of a place, the long-term effect of the adoption of E&E technologies should be positive. Additionally, this effect should be increasing over time (though probably negligible at the beginning), as the economy-wide (but also place-specific) amount of complementarities increases.

The main concern when evaluating the impact of technological adoption on economic outputs lies in the possibility that economic outputs may, in fact, have also a reinforcing effect on the adoption of technologies. If such feedbacks exist, Ordinary Least Squares (OLS) or Fixed Effects (FE) estimation methods would render biased estimates, as the error term will be correlated with the regressor(s). This is likely in the case of E&E technologies, as they required considerable initial investments and complementary infrastructure to work. For instance, an inventor of E&E technologies in the 1900s may have faced incentives to locate near a power station after developing an electrical device, as the destination market for his/her invention was there. If the choice of the location of power stations was related to the size and prosperity of the market they will serve, then the direction of the effect reverses.

The empirical strategy adopted here relies on using measures of the adoption of E&E technologies prior to the 1870s as an instrument to predict the adoption of E&E technologies between 1900 and 1930 to later evaluate their impact on per capita output growth and wages. This assumes that the early adoption of E&E technologies (prior to the invention of the electric lightbulb or the establishment of the first power plant) were crucial to the inventive structure of places 50 years later while not being correlated with the events that will determine growth and wages between 1900 and 1930.

Using the early adoption of E&E technologies as an instrument relies on three assumptions. First, prior to the 1870s it was nearly impossible to predict which future markets would be the most appropriate for E&E technologies, as they were in a very exploratory phase, such that their possible range of application and their scope for complementarities were barely understood. The history of the Thomson-Houston Electric Company, founded in 1883 in Lynn Massachusetts, constitutes an enlightening example. This company was a leader in the manufacture of products related to Elihu Thomson's inventions. These inventions included arc-lighting systems[19], dynamos[20] , and systems for electric distribution[21], among others. More than three decades after Elihu Thomson left his position (1880) as assistant professor of chemistry at Central High School in Philadelphia to pursue research on future applications of electricity (where he was a colleague of Edwin J. Houston), his company would be essential in the construction of the Panama canal, which was completed in 1914. General Electric (GE), created in 1892 through the merger of the Edison General Electric Company and

[19]See https://www.google.com/patents/US261790
[20]https://www.google.com/patents/US302963
[21]https://www.google.com/patents/US335159

Thomson-Houston Electric Company, supplied approximately 1500 electric motors and electromechanical control boards that not only operated the gates and valves of the canal, but also regulated the operations of the the hydroelectric dam at Gatun Lake, which provided electricity for the canal(Nebeker, 2009). The GE manufacturing plant at Lynn played an important role in the construction of the Panama canal, a role that could not have been envisioned at the time of the company's founding[22]. If these future developments could have been predicted back in the 1880s, the Thomson-Houston Electric Company should have been located near GE headquarters in Schenectady (New York), to profit from agglomeration economies.

The second assumption requires that the location of complementary infrastructure (present or future) was not a relevant factor influencing the location decisions of pioneering electrical entrepreneurs. Otherwise infrastructure-led growth might have been the responsible factor for the well-being of places, rather than their early adoption of any particular technology. Note, however, that electricity-related infrastructure was almost non-existent prior to the 1870s. In fact, Pearl Street Station (in Manhattan, New York) was the first central power plant in the US, which opened in 1882 and served only 82 customers (Orrok, 1930). Even the possibility of electrical illumination was uncertain at the end of the 1870s, as early experiments by Thomas Edison in 1879 produced light bulbs that lasted only 13.5 hours (Israel and Edison, 1998).

A clear example of this can be found by considering the development of the sector of electrical appliances in the US, which was one of the fastest growing industries after the 1920s. For instance, the production of refrigerators jumped from 5,000 units in 1920 to 1,000,000 units in 1930, reaching 6,000,000 units by 1936 (Nebeker, 2009). The highest share of this market was occupied by the

[22]In fact, the importance of this initial venture remains until today, as the factory is still an essential part of GE, producing helicopter and jet engines, among other things, and employing around 45.000 people

Kelvinator Company of Detroit, Michigan. This company, founded by engineer Nathaniel B. Wales in 1914, introduced the first refrigerator with automated control[23]. The story of this company exemplifies the extent to which success did not relate to the existence of complementary infrastructure. Nevertheless, at that time, the electrification of houses was primarily done for illumination purposes, such that at the beginning of the 20th centuryn most houses in Detroit (and for that matter in the entire U.S.) did not have wall sockets to connect appliances. This meant that appliances had to be wired to chandeliers to connect them to electrical current[24]. This indicates the lack of appropriate infrastructure that was present at the time, which was essential for the diffusion of electrical appliances such as refrigerators.

The third assumption requires that early developments of a technology can be used to predict future adoption. Therefore, the persistence over time and space of technological capabilities is essential. In this regard, there is a well-established literature demonstrating that processes of technological diffusion tend to be highly dependent on geography, as physical proximity and collocation are crucial in the dynamics of innovative processes (Audretsch and Feldman, 1996; Feldman and Kogler, 2010; Feldman and Yoon, 2012; **?**; **?**). This effect is expected to be stronger for relatively new and unfamiliar technologies. At early stages, technological diffusion is characterized by the importance of tacit knowledge; only when knowledge becomes standardized does geographical dispersion tend to occur(Feldman and Kogler, 2010).

We consider two key variables to measure the economic impact of the adoption of E&E technologies in a county, the *Average Wage* $(W_{c,t})$ paid to manufacturing workers and the *per capita Growth* $(\Delta y_{c,t})$ in those places. For each county, $\Delta y_{c,t}$ measures the log difference in per capita output in manufactures, while $W_{c,t}$ is

[23]See for instance https://www.google.com/patents/US1438178

[24]See for instance https://www.google.com/patents/US646179 for an example of a chandelier adapter.

calculated as the total expenditures on manufacturing wages divided by the total number of hands employed in manufactures.

To track the adoption of technologies, we create a set of dummies that take value 1 whenever a county had positive patenting activity in a particular technology in the five years prior the census year [25]. For instance,

$$Chemical_{c,t} = 1 \qquad if \qquad \sum_{t-5}^{t} Patents_{Chem} > 0 \; and \; zero \; otherwise$$

This implies that the entire technological profile of counties is characterized by a set of four dummy variables, $Electrical\ \&\ Electronics_{c,t}$, $Mechanical_{c,t}$, $Chemical_{c,t}$, and $Others_{c,t}$. Then the estimating equation can be summarized as follows:

$$DV_{c,t} = \beta_0 + \beta_1 T_{c,t} + \beta_2 X_{c,t} + \epsilon_{c,t} \tag{5.1}$$

Where DV (the dependent variable) will be either the *Average Wage* ($W_{c,t}$) or *per capita Growth* ($\Delta y_{c,t}$), and $T_{c,t}$ represents a vector of technological variables, which includes the full set of dummy variables described above ($Electrical\ \&\ Electronics_{c,t}$, $Mechanical_{c,t}$, $Chemical_{c,t}$, and $Others_{c,t}$.).

$X_{c,t}$ comprises a set of control variables that could be determinants of $W_{c,t}$ and $\Delta y_{c,t}$. This includes the share of the immigrant population (*Foreign Share*), as there is evidence of the importance of cultural diversity and immigration for growth (Ager and Brückner, 2013; Burchardi et al., 2016). Furthermore, two variables related to the availability and exploitation of natural resources, the share of primary inputs in manufacture production (*Primary Inputs*) and the number of actively worked mineral deposits are also included (*Working Deposits*)[26]. The implications and effects of natural resource (NR) exploitation during this period are

[25]We use a window of five years because patenting activity fluctuates greatly from year to year.

[26]This variable was created by geolocating all mineral deposits within counties' boundaries. Data on geo-located mineral deposits, including their operational status, can be found at https://www.usgs.gov.

discussed in detail by Wright (1990) . Additionally, we include a set of variables related to the state's presence at the county level, given the critical role that has been attributed to institutional factors in the early development process of the US (Khan, 2005; Sokoloff, 1988). We use the share of federal, state, and local public employees as a share of total population (*Federal Employment*, *State Employment*, *Local Employment*, respectively). Additionally, and following Acemoglu et al. (2016), we include the number of post offices per county *(PO)* as a proxy for the state's capacity on places. Depending on the specification, $X_{c,t}$ will include county, state, and year dummy variables.

Regression table number 5.4 below reports two regressions that evaluate the effect of technological adoption on $W_{c,t}$ and $\Delta y_{c,t}$: they include county and year fixed effects, and standard errors are clustered by county and year according to Cameron et al. (2012). These regressions evaluate, as a starting point, the extent to which differences in technological adoption can be related to economic outcomes, above and beyond what can be explained by specific characteristics of places. They include the years between 1860 and 1930, covering the entire cycle of the emergence, development, and diffusion of E&E technologies. Note that counties that adopted E&E technologies over-performed others in terms of both average wages and per capita growth. This provides evidence that the adoption of E&E technologies can be related to a positive and significant differential in terms of wages and output per capita during this period.

Additionally, the degree of foreign immigration in a county was positively related to the growth of counties but negatively correlated with wages. This is consistent with previous findings; see, for instance Ager and Brückner (2013) and Hatton and Williamson (2008) . The intensity of the use of primary inputs is positively related to growth, which is in line with previous findings on the role of the exploitation of geological potential and non-reproducible natural resources in U.S. manufacturing (Wright, 1990). Regarding the role of the state's presence, it only seems to be associated with a positive shift in wages in places with a higher number of post offices. This is not surprising since they represented a high share of the local employment, thus placing pressure on local labor markets. For instance, by 1831, postal employees already accounted for 76% of the civilian federal workforce, and postmasters outnumbered soldiers, making the former the most widespread representatives of the federal government (Acemoglu et al., 2016; Service, 2007).

As discussed earlier, joint determination of economic outcomes and technological adoption, as well as the omission of any relevant explanatory variable correlated with technological adoption, would yield biased estimates. Table 5.5 below shows the result of estimating a 2SLS model using early adoption of E&E technologies as an excluded instrument to predict the adoption of E&E technologies between 1900 and 1930. In particular, we construct two instruments:

$$E\&E_{c,1870} = 1 \qquad if \qquad \sum_{1866}^{1875} Patents_{E\&E} > 0, \; and \; zero \; otherwise,$$

and

$$E\&E_{c,1860} = 1 \qquad if \qquad \sum_{1856}^{1865} Patents_{E\&E} > 0, \; and \; zero \; otherwise$$

For comparison, the FE estimates of table 5.4 are reported jointly with the 2SLS estimates of each specification. The 2SLS estimates of the effects of E&E adoption on both wages and growth are several times higher than the corresponding FE estimates. Note however that there is a difference between specifications regarding the time span considered, as the FE estimates are obtained using the entire sample (1860-1930), while the 2SLS estimates only over the period 1900-1930.

Table 5.4: *Effect of the Adoption of E&E Technologies*

	Growth ($\Delta y_{c,t}$) (FE) (1)	*Average Wage* ($W_{c,t}$) (FE) (2)
Electrical and Electronics	0.109**	45.903***
	(0.043)	(16.126)
Mechanic	0.048	27.974***
	(0.031)	(7.400)
Chemical	0.052*	15.069***
	(0.027)	(5.794)
Others	0.037*	10.473*
	(0.022)	(5.903)
Foreign (Share)	0.002***	−0.385*
	(0.001)	(0.207)
Establishments per Capita (in logs)	0.741***	−33.974
	(0.043)	(21.197)
Primary Inputs	1.376***	−56.063
	(0.178)	(36.628)
Working Deposits (in logs)	0.045	−8.226
	(0.034)	(15.904)
Federal Employment (Share)	−0.001	0.173
	(0.001)	(0.228)
State Employment (Share)	−0.001	0.699
	(0.003)	(0.675)
Local Employment (Share)	0.002	0.309
	(0.003)	(0.553)
PO (in logs)	−0.006	28.383*
	(0.067)	(15.190)
$y_{c,t-1}$	−0.807***	
	(0.075)	
Period	1860-1930	1860-1930
Year FE	Yes	Yes
County FE	Yes	Yes
Observations	14,866	16,762
Adjusted R^2	0.568	0.866

*p<0.1; **p<0.05; ***p<0.01

Figure 5.6 below provides a characterization of the dynamics, which allows for more meaningful comparisons. Nonetheless, 2SLS estimates suggest that the adoption of E&E technologies during this period is associated with a higher steady-state level of income per capita. Additionally, places adopting E&E technologies also paid, on average, higher wages. Other coefficients have similar values when comparing FE with 2SLS estimates, provided that they are statistically significant in both cases.

Note that the instruments are highly correlated with the adoption of E&E technologies, as the F-test on the joint significance strongly rejects the null hypothesis. Additionally, as the 2SLS estimates are over-identified, the Sargan test of over-identified restrictions is reported at the bottom of the table. In both cases, the test does not the reject the null hypothesis of the instrument being orthogonal to the error term.

According to theoretical models, the beneficial effect of adopting the new GPT should be the highest when enough complementarities have been developed. Historical evidence suggests that in the case of electricity, this occurred after the 1910s (David, 1990; Lipsey et al., 2005), which is consistent with the economy-wide diffusion of use and innovative complementarities reported in subsections 5.3.2 and 5.3.3. It is therefore worth exploring whether we see a similar pattern here.

Figure 5.6 shows the result of interacting the E&E variable with a dummy for each census year [27]. It corroborates what David (1990); Greenwood (1997); Lipsey et al. (2005) and Field (2008) suggest: the impact of E&E technologies increased over time and is significant only after the 1910s. This indicates that the economy-wide expansion in the number of E&E complementarities strengthened the effect of adopting E&E technologies on economic outcomes.

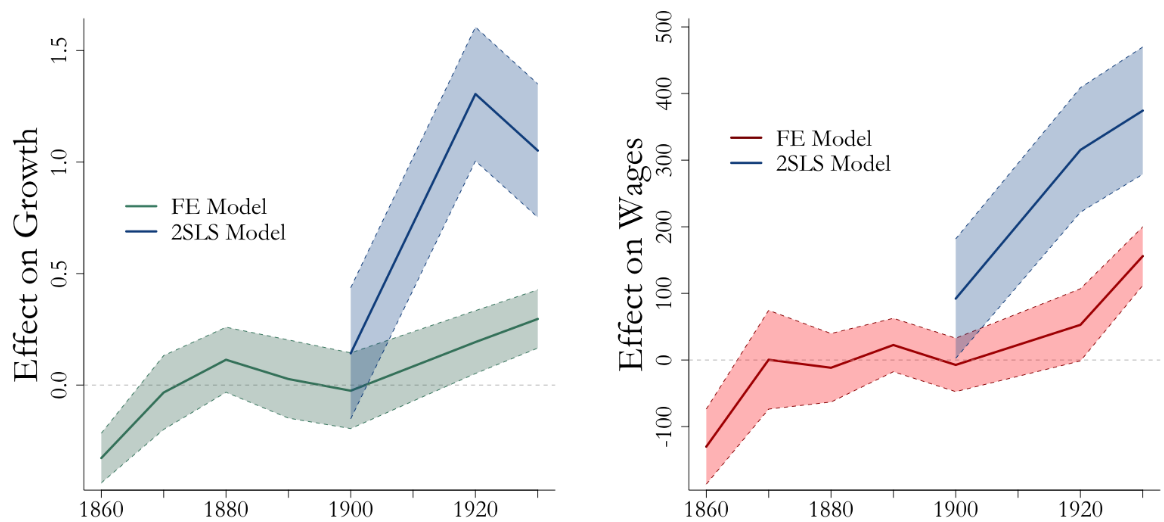
[27]The regression table can be found in the appendix.

Table 5.5: *Effect of the Adoption of E&E Technologies*

	Growth ($\Delta y_{c,t}$)		*Average Wage* ($W_{c,t}$)	
	(FE)	(2SLS)	(FE)	(2SLS)
	(1)	(2)	(3)	(4)
Electrical and Electronics	0.109**	0.816***	45.903***	293.907***
	(0.043)	(0.158)	(16.126)	(43.588)
Mechanic	0.048	0.092***	27.974***	26.378***
	(0.031)	(0.026)	(7.400)	(7.514)
Chemical	0.052*	0.014	15.069***	−16.381
	(0.027)	(0.038)	(5.794)	(11.980)
Others	0.037*	0.049*	10.473*	13.721*
	(0.022)	(0.028)	(5.903)	(7.632)
Foreign (Share)	0.002***	0.001***	−0.385*	−0.004
	(0.001)	(0.0002)	(0.207)	(0.060)
Establishments per Capita (in logs)	0.741***	0.644***	−33.974	−45.709***
	(0.043)	(0.025)	(21.197)	(6.423)
Primary Inputs	1.376***	1.392***	−56.063	−115.387***
	(0.178)	(0.074)	(36.628)	(19.906)
Working Deposits (in logs)	0.045	−0.013	−8.226	1.226
	(0.034)	(0.012)	(15.904)	(3.412)
Federal Employment (Share)	−0.001	−0.001	0.173	0.206
	(0.001)	(0.001)	(0.228)	(0.200)
State Employment (Share)	−0.001	0.001	0.699	1.722
	(0.003)	(0.004)	(0.675)	(1.377)
Local Employment (Share)	0.002	0.001	0.309	1.704***
	(0.003)	(0.002)	(0.553)	(0.572)
PO (in logs)	−0.006	−0.017	28.383*	1.135
	(0.067)	(0.020)	(15.190)	(4.748)
$y_{c,t-1}$	−0.807***	−0.465***		
	(0.075)	(0.015)		
Period	1860-1930	1900-1930	1860-1930	1900-1930
Year FE	Yes	Yes	Yes	Yes
County FE	Yes	No	Yes	No
State FE	No	Yes	No	Yes
Weak Instruments (F-statistic)		52.702		72.338
Sargan Test (p-value)		0.758		0.3
Observations	14,866	7,547	16,762	8,059
Adjusted R^2	0.568	0.272	0.866	0.725

*p<0.1; **p<0.05; ***p<0.01

Figure 5.6: *The Effect of E&E Adoption over Time*



There are a few threats to the identification strategy that should be taken into account. First, it is possible that when properly estimated, the effects of adopting other technologies are higher than those of E&E technologies. If that is the case, the proposed E&E differential would not exist.

The validity of the instrument used here relies heavily on characteristics and historical features that are particular to E&E technologies alone. For instance, because Chemical and Mechanical technologies were already mature technologies prior to the 1870s, their adoption may be correlated with characteristics of a place that remain relevant years later. In the case of chemical technologies, adoption may be correlated with the availability of natural resources, which probably had long-lasting implications for growth.

To address this first issue, we use lagged values of technological adoption as instruments for technologies other than E&E. In the appendix, we show that the results do not change in any significant way. Although the magnitude of the E&E coefficient does not change significantly after instrumenting for the adoption of other technologies, its significance decreases.

Another potential threat concerns the fact that using lagged variables of E&E adoption may actually be capturing characteristics of places that remain relevant 50 years later. For instance, early E&E adoption may be correlated with initial characteristics of places such as a higher level of development, better infrastructure, or better institutions, which can have a long-lasting effect on growth rates and wages. More generally, this concern relates to the fact that the success of the empirical strategy relies on the maintained assumption that county-specific characteristics, included in the error term in the 2SLS specification, are not correlated with the instrument.

We address this issue in two robustness checks. First, we use the Blundell and Bond (1998) GMM procedure, which allows us to construct a time-varying instrument, similar to that used in the baseline estimation, and estimate the effect of E&E adoption after controlling for county-level heterogeneity. Using the proposed set of GMM instruments, we instrument for both the adoption of E&E technologies and the lagged dependent variable ($y_{c,t-1}$) in the growth regression with their lagged values ($E\&E_{c,t-r}$ and $y_{c,t-s}$ for $s > 2$ and $r > 3$, respectively). Even though this GMM specification does not pass the Sargan test of overidentifying restrictions, it yields estimates of the effect of technological adoption that are consistent with previous findings, after controlling for county-specific characteristics.

Additionally, we check the robustness of the results to the inclusion of additional regressors to control for initial county-specific characteristics. In the appendix, we provide further results using lagged values of regressors and output levels as additional explanatory variables. Because the coverage of the census is more limited prior to the 1870s, this implies the looss of some observational units. As before, the results do not change significantly.

Finally, estimating a growth equation using OLS or FE methods can potentially bias the coefficients of the technological variables. Nickell (1981) shows that using the within-effect estimator will produce inconsistent estimates of the lagged dependent variable in dynamic models. If $y_{c,t-1}$ is correlated with other regressors, the estimated coefficients could be biased even if the regressors are exogenous.

To address this concern, we use alternative GMM methods designed to properly identify the effect of $y_{c,t-1}$, such as those described in Blundell and Bond (1998). Using the proposed set of instruments, which also rely on lagged values of the endogenous variables, we instrument for both the adoption of E&E technologies and the effect of $y_{c,t-1}$. Additionally, and because exogeneity tests cast doubt on the validity of the GMM instruments, we provide an additional robustness check using lagged values of E&E adoption and $y_{c,t-1}$ as instruments in a 2SLS framework. In both cases, the results are consistent with the previous findings.

5.5 Conclusions

This chapter discusses, on the one hand, a simple way of characterizing GPTs using patent data. By relying on historical patent documents, it provides a comprehensive view of the emergence, evolution, development, and diffusion of E&E technologies in their historical context. It shows that the behavior of E&E technologies between 1860 and 1930 is in line with what can be expected from a GPT, namely: above-average growth rates in patenting activity, the development of a wide variety of innovation complementarities, and a high degree of pervasiveness in the U.S. inventive structure.

On the other hand, we present evidence that the adoption of E&E technologies is related to higher per capita growth and higher wages. Identification of the impact is achieved using 2SLS methods and instrumenting E&E adoption with lagged values of it. Even if potential threats to the identification strategy exist, as in any empirical endeavor, the results are robust to changes in the model specification and to alternative adoption measures, among other factors.

Having a simple and useful means of characterizing GPTs has considerable policy implications, for instance, to identify technologies that are currently showing qualities of a GPT, which would be capable of fostering growth and generating spillovers. Additionally, the results suggest that the uneven diffusion of revolutionary technologies may generate inequality across places in terms of growth and earnings.

It is also important to emphasize the limitations of this study. Because it relies solely on collecting the traces of information that have been left behind in patent documents, it is not possible to draw any general conclusion on the full extent of the effect of GPTs on the economy. In the case of E&E technologies, it is widely known that their impact went beyond what can be traced in patent documents. It had implications for the organization of factories, the transportation systems of cities, among other developments. Additionally, this study does not account for the benefits that mere adopters may have experienced.

A natural step forward would be to test whether the characterization of GPTs provided here holds similarly for ICTs after the 1970s. Focusing on a more recent period would provide the opportunity to use more disaggregated and detailed data; which could be exploited to design better identification strategies and/or explore additional aspects of the diffusion of GPTs.

5.6 Appendix

5.6.1 List of Key-Words Used in Section 5.3.2

- Catelectrode
- Catelectrotonic
- Catelectrotonus
- Dielectric
- Dynamoelectric
- Electre
- Electrepeter
- Electress
- Electric
- Electrical
- Electrically
- Electricalness
- Electrician
- Electricities
- Electricity
- Electrifiable
- Electrification
- Electrified
- Electrifying
- Electrify
- Electrine
- Electrition
- Electrization
- Electrized
- Electrizing
- Electrize
- Electrizer
- Electro
- Electroballistic
- Electroballistics
- Electrobiologist
- Electrobiology
- Electrobioscopy
- Electrocapillarity
- Electrocapillary
- Electrochemical
- Electrochemistry
- Electrochronograph
- Electrochronographic
- Electrocute
- Electrode
- Electrodynamic
- Electrodynamical
- Electrodynamics
- Electrodynamometer
- Electroengraving
- Electroetching
- Electrogenesis
- Electrogenic
- Electrogeny
- Electrogilding
- Electrogilt
- Electrograph
- Electrokinetic
- Electrokinetics
- Electrolier
- Electrology
- Electrolysis
- Electrolyte
- Electrolytic
- Electrolytical
- Electrolyzable
- Electrolyzation
- Electrolyzed
- Electrolyzing
- Electrolyze
- Electromagnet
- Electromagnetic
- Electromagnetism
- Electrometallurgy
- Electrometer
- Electrometric
- Electrometrical
- Electromotion
- Electromotive
- Electromotor
- Electromuscular
- Electron
- Electronegative
- Electropathy
- Electrophone
- Electrophori
- Electrophorus
- Electrophysiological
- Electrophysiology
- Electroplating
- Electroplate
- Electroplater

- Electropolar
- Electropositive
- Electropuncturation
- Electropuncturing
- Electropuncture
- Electroscope
- Electroscopic
- Electrostatic
- Electrostatics
- Electrostereotype
- Electrotelegraphic
- Electrotelegraphy
- Electrotherapeutics
- Electrothermancy
- Electrotint
- Electrotonic
- Electrotonize
- Electrotonous
- Electrotonus
- Electrotype
- Electrotyped
- Electrotyping
- Electrotyper
- Electrotypic
- Electrotypy
- Electrovital
- Electrovitalism
- Electrum
- Hydroelectric
- Idioelectric
- Magnetoelectric
- Magnetoelectrical
- Magnetoelectricity
- Parelectronomic
- Parelectronomy
- Photoelectric
- Photoelectrotype
- Pyroelectric
- Pyroelectricity
- Resinoelectric
- Stereoelectric
- Thermoelectric
- Thermoelectricity
- Thermoelectrometer
- Voltaelectric
- Voltaelectrometer

5.6.2 Regression Table with Year Interactions

In section 5.4, Figure 5.6 shows how the effect of E&E adoption changes over time for both models, FE and 2SLS. Here, we report the tables on which this figure is based, which consist of analogous sets of regressions to those reported in Section 5.4, Tables 5.4 and 5.5, but including an interaction between the variable *Electrical and Electronics* (E&E) and a set of dummies for each census year. In the case of the 2SLS model, we do not treat each interaction as an endogenous regressor, as we have only two instruments. We proceed in two steps: first, we obtain the predicted values of E&E adoption as in the baseline case, to later plug in the generated regressor $\widehat{E\&E}$ interacted with the set of year dummies in the second-stage OLS regression.

Table 5.6: *FE Estimates with Year Interactions*

	Growth ($\Delta y_{c,t}$)	*Average Wage* ($W_{c,t}$)
	(1)	(2)
Electrical and Electronics * Year=1860	-0.327^{***}	-130.091^{***}
	(0.057)	(28.750)
Electrical and Electronics * Year=1870	-0.034	0.428
	(0.085)	(37.814)
Electrical and Electronics * Year=1880	0.114	-11.642
	(0.074)	(26.237)
Electrical and Electronics * Year=1890	0.027	22.515
	(0.089)	(20.455)
Electrical and Electronics * Year=1900	-0.025	-7.373
	(0.087)	(20.458)
Electrical and Electronics * Year=1920	0.193^{***}	52.788^{*}
	(0.072)	(27.647)
Electrical and Electronics * Year=1930	0.296^{***}	155.966^{***}
	(0.067)	(22.394)
Mechanic	0.050^{*}	28.815^{***}
	(0.027)	(6.169)
Chemical	0.054^{**}	15.440^{***}
	(0.027)	(5.679)
Others	0.041^{**}	12.112^{**}
	(0.020)	(4.920)
Foreign (Share)	0.002^{***}	-0.355^{*}
	(0.001)	(0.192)
Establishments per Capita (in logs)	0.752^{***}	-29.800
	(0.041)	(18.698)
Primary Inputs	1.375^{***}	-55.563
	(0.173)	(34.244)
Working Deposits (in logs)	0.048	-6.874
	(0.033)	(15.623)
Federal Employment (Share)	-0.001	0.143
	(0.001)	(0.227)
State Employment (Share)	-0.003	0.135
	(0.003)	(0.662)
Local Employment (Share)	0.001	0.042
	(0.002)	(0.395)
PO (in logs)	-0.002	30.597^{**}
	(0.068)	(14.872)
$y_{c,t-1}$	-0.808^{***}	
	(0.073)	
Period	1860-1930	1860-1930
Year FE	Yes	Yes
County FE	Yes	Yes
Observations	14,866	16,762
Adjusted R^2	0.573	0.871

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 5.7: *2SLS Estimates with Year Interactions*

	Growth ($\Delta y_{c,t}$)	*Average Wage* ($W_{c,t}$)
	(1)	(2)
Electrical and Electronics * Year=1900	0.098	291.910***
	(0.146)	(29.226)
Electrical and Electronics * Year=1920	1.266***	507.814***
	(0.148)	(34.248)
Electrical and Electronics * Year=1930	1.014***	577.477***
	(0.147)	(34.835)
Mechanic	0.089***	8.686
	(0.025)	(6.507)
Chemical	0.020	−54.366***
	(0.033)	(8.315)
Others	0.048*	1.552
	(0.026)	(6.844)
Foreign (Share)	0.001***	−0.102*
	(0.0002)	(0.054)
Establishments per Capita (in logs)	0.652***	−58.086***
	(0.024)	(6.005)
Primary Inputs	1.356***	−137.503***
	(0.071)	(18.854)
Working Deposits (in logs)	−0.015	1.386
	(0.012)	(3.065)
Federal Employment (Share)	−0.001	0.140
	(0.001)	(0.193)
State Employment (Share)	0.001	1.384
	(0.004)	(1.098)
Local Employment (Share)	0.001	0.837
	(0.002)	(0.708)
PO (in logs)	−0.012	−9.882**
	(0.018)	(4.718)
$y_{c,t-1}$	−0.461***	
	(0.014)	
Period	1900-1930	1900-1930
Year FE	Yes	Yes
State FE	Yes	Yes
Observations	7,547	7,545
Adjusted R^2	0.375	0.789

*p<0.1; **p<0.05; ***p<0.01

5.6.3 Controlling for Initial Characteristics of Places

One potential threat to the identification strategy is that using lagged variables of E&E adoption could also be capturing the initial conditions of places. If these initial conditions were important enough to establish a virtuous cycle of long-lasting growth and wealth accumulation, then the instrument would lose its validity. The early characteristics of places such as having higher levels of development, better infrastructure, or a greater state presence may have had long-lasting implications for the productive structure of places and their growth.

We use two additional variables to capture the initial conditions of places: On the one hand, we use output per capita in 1870 ($y_{c,1870}$) to proxy for initial levels of development and infrastructure and on the other, we include the number of post offices in 1870 to proxy for state presence ($PO_{c,1870}$).

Using 1870's per capita output as an additional regressor requires constraining the sample to include places that were covered by the census of 1870. This may result in a non-trivial selection of the sample, biasing the results. Therefore, we approach the problem by checking the robustness of the results in two steps: First, we compare two sets of FE estimates, with the only difference between them being the number of counties included in the sample. This is to assess the effect that constraining the sample has on the estimates independent of initial conditions, as FE estimates already control for them. Next, we include $y_{c,1870}$ and $PO_{c,1870}$ as additional regressors in the 2SLS estimations and assess whether the results change.

Columns (1) and (3) in Table 5.8 below present the baseline results of Table 5.4; while columns (2) and (4) replicate the procedure for the subset of counties that had output data by 1870. The results do not change significantly between specifications, especially for the variables capturing technological adoption.

Table 5.8: *FE Estimates on the Restricted Sample*

	Growth ($\Delta y_{c,t}$)		*Average Wage* ($W_{c,t}$)	
	(1)	(2)	(3)	(4)
Electrical and Electronics	0.109**	0.109**	45.903***	47.676***
	(0.043)	(0.048)	(16.126)	(18.198)
Mechanic	0.048	0.050	27.974***	23.826***
	(0.031)	(0.031)	(7.400)	(8.017)
Chemical	0.052*	0.050*	15.069***	12.183**
	(0.027)	(0.027)	(5.794)	(6.015)
Others	0.037*	0.036	10.473*	9.118
	(0.022)	(0.024)	(5.903)	(6.578)
Foreign (Share)	0.002***	0.001**	−0.385*	−0.331
	(0.001)	(0.001)	(0.207)	(0.216)
Establishments per Capita (in logs)	0.741***	0.735***	−33.974	−37.419*
	(0.043)	(0.048)	(21.197)	(20.796)
Primary Inputs	1.376***	1.276***	−56.063	−74.758**
	(0.178)	(0.174)	(36.628)	(33.394)
Working Deposits (in logs)	0.045	0.049	−8.226	−14.651
	(0.034)	(0.033)	(15.904)	(16.171)
Federal Employment (Share)	−0.001	−0.001	0.173	0.280
	(0.001)	(0.001)	(0.228)	(0.207)
State Employment (Share)	−0.001	0.0001	0.699	1.887***
	(0.003)	(0.003)	(0.675)	(0.565)
Local Employment (Share)	0.002	0.002	0.309	0.519
	(0.003)	(0.003)	(0.553)	(0.726)
PO (in logs)	−0.006	−0.005	28.383*	24.782
	(0.067)	(0.073)	(15.190)	(15.941)
$y_{c,t-1}$	−0.807***	−0.780***		
	(0.075)	(0.076)		
Period	1860-1930	1860-1930	1860-1930	1860-1930
Year FE	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	14,866	12,917	16,762	13,775
Adjusted R^2	0.568	0.547	0.866	0.857

*p<0.1; **p<0.05; ***p<0.01

Table 5.9 below compares the 2SLS baseline estimates of Table 5.5, in columns (1) and (4), with two analogous estimates: Columns (2) and (5) replicate the baseline estimates but on the reduced sample, while columns (3) and (6) do the same but also include $y_{c,1870}$ and $PO_{c,1870}$ as additional regressors.

The results are robust to the subsetting of the sample and the inclusion of proxies for the initial conditions of places. They display the same consistent pattern across specifications, with instrumented adoption of E&E technologies showing a strong effect on wages and output. Additionally, note that both tests for the validity of the instruments also work on the alternative specifications: The F-test of joint significance always strongly rejects the null hypothesis while the Sargan test of over-identified restrictions does not reject the null.

Table 5.9: *2SLS Estimates Proxying for Initial Conditions*

	Growth ($\Delta y_{c,t}$)			*Average Wage* ($W_{c,t}$)		
	(1)	(2)	(3)	(4)	(5)	(6)
Electrical and Electronics	0.816^{***}	0.756^{***}	0.658^{***}	293.901^{***}	306.323^{***}	312.954^{***}
	(0.158)	(0.171)	(0.182)	(43.429)	(49.693)	(56.627)
Mechanic	0.092^{***}	0.066^{**}	0.073^{**}	26.548^{***}	21.677^{**}	21.734^{**}
	(0.026)	(0.031)	(0.032)	(7.470)	(9.199)	(9.499)
Chemical	0.014	0.023	0.037	-16.352	-21.814	-23.463
	(0.038)	(0.041)	(0.042)	(11.942)	(13.588)	(14.407)
Others	0.049^{*}	0.063^{**}	0.067^{**}	14.265^{*}	18.016^{**}	17.725^{**}
	(0.028)	(0.031)	(0.031)	(7.538)	(8.425)	(8.515)
Foreign (Share)	0.001^{***}	0.001^{**}	0.001^{**}	-0.011	0.094	0.035
	(0.0002)	(0.0002)	(0.0002)	(0.059)	(0.077)	(0.080)
Establishments per Capita (in logs)	0.644^{***}	0.575^{***}	0.570^{***}	-44.850^{***}	-60.795^{***}	-62.043^{***}
	(0.025)	(0.029)	(0.029)	(6.337)	(7.293)	(7.215)
Primary Inputs	1.392^{***}	1.154^{***}	1.140^{***}	-113.816^{***}	-130.347^{***}	-127.800^{***}
	(0.074)	(0.089)	(0.088)	(19.620)	(23.920)	(23.853)
Working Deposits (in logs)	-0.013	-0.008	-0.011	1.115	-0.577	-0.001
	(0.012)	(0.013)	(0.013)	(3.404)	(3.911)	(3.968)
Federal Employment (Share)	-0.001	-0.001	-0.001	0.200	0.403	0.379
	(0.001)	(0.001)	(0.001)	(0.200)	(0.308)	(0.308)
State Employment (Share)	0.001	0.002	0.001	1.737	1.719	1.635
	(0.004)	(0.004)	(0.004)	(1.374)	(1.555)	(1.596)
Local Employment (Share)	0.001	0.003	0.003	1.569^{***}	2.166^{**}	1.933^{*}
	(0.002)	(0.003)	(0.003)	(0.562)	(1.003)	(1.008)
PO (in logs)	-0.017	-0.031	-0.035	1.298	4.796	17.945^{**}
	(0.020)	(0.023)	(0.025)	(4.709)	(6.200)	(7.092)
$y_{c,t-1}$	-0.465^{***}	-0.425^{***}	-0.425^{***}			
	(0.015)	(0.017)	(0.017)			
$y_{c,1870}$			0.024^{**}			8.131^{**}
			(0.012)			(3.926)
$PO_{c,1870}$ (in logs)			0.028			-24.454^{***}
			(0.021)			(6.146)
Period	1900-1930	1900-1930	1900-1930	1900-1930	1900-1930	1900-1930
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Weak Instruments (F-statistic)	52.702	41.556	35.087	72.917	56.26	43.797
Sargan Test (p-value)	0.758	0.56	0.566	0.295	0.304	0.343
Observations	7,547	5,893	5,893	8,129	5,983	5,983
Adjusted R^2	0.272	0.237	0.265	0.727	0.724	0.722

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

5.6.4 Robustness to Alternative Estimation Methods for Dynamic Panels

As mentioned at the end of Section 5.4, estimating a growth equation using FE methods can potentially bias the coefficients of the technological variables. Nickell (1981) shows that using the within-effect estimator will produce inconsistent estimates of the effect of the lagged dependent variable in dynamic models. If $y_{c,t-1}$ is correlated with other regressors, the estimated coefficients could be biased even if they are exogenous. This means that results in Table 5.4 could be a reflection of the underlying correlation structure between technological adoption variables and $y_{c,t-1}$. In this subsection, we check the robustness of the results in Table 5.4 using an alternative GMM procedure, described in Blundell and Bond (1998), to appropriately estimate the effect of $y_{c,t-1}$ in growth regressions.

Additionally, as this procedure uses lagged values and lagged differences of endogenous variables as instruments, it could be used to test the robustness of the results in Table 5.5 in a setup in which instruments for E&E adoption vary over time. This would make it possible to account for county-specific characteristics in the estimation.

Note that equation (5.2) can be re-written as,

$$y_{c,t} = \beta_0 + \beta_1 T_{c,t} + \beta_2 X_{c,t} + \rho y_{c,t-1} + \gamma_c + \gamma_t + \epsilon_{c,t} \tag{5.2}$$

after adding $y_{c,t-1}$ on both sides. For the sake of exposition, let $y_{c,t-1}$ and the county and time dummy variables (γ_c, γ_t) be expressed separately from $X_{c,t}$. Nickell (1981) shows that if we estimate this equation by FE, β_1 would be biased even if $T_{c,t}$ are exogenous.

To address this concern, we use the GMM procedure described in Blundell and Bond (1998), which uses lagged values of $y_{c,t-s}$ and $\Delta y_{c,t-s}$ as instruments for $y_{c,t-1}$. In this way we seek to eliminate the possibility that the alleged E&E

differential we see in Table 5.4 is caused by the “Nickel bias” mentioned above by appropriately estimating the effect of $y_{c,t-1}$ on $y_{c,t}$ (ρ). Table 5.10 below reports the results of this exercise.

Note that the GMM estimates of the effect of technological adoption are consistent with those estimated by FE. They maintain the same relative rankings but are higher than their FE counterparts. The adoption of E&E technologies is still associated with higher growth rates under this alternative estimation procedure. This result could serve also as a partial explanation for why the FE estimates are downward biased with respect to 2SLS estimates, as they incorporate “Nickel bias”. Additionally, note that $\widehat{\rho} = 0.52$ and lies between the FE (0.193) and OLS (0.65) estimates, as expected (Blundell and Bond, 2000).

The pattern of serial correlation in the first-differenced residuals is consistent with the assumptions, showing only a significant first-order serial correlation. However, the validity of the instruments is clearly rejected by the Sargan test of overidentifying restrictions. This could be explained by the presence of measurement errors, or simply because there is a component of the error term that is correlated with output and persistent enough to invalidate the use of lagged output values as instruments. Unfortunately, the result of the Sargan test is robust to different model specifications and variations in the construction of the instruments.

As mentioned above, it could be argued that the results reported in Table 5.5 are partially explained by county-specific heterogeneity, if the instrument is somehow related to a particular, time-invariant, characteristic of places. Despite the effort to address this issue in Section 5.6.3, the alleged differential of E&E adoption may be the reflection of some intrinsic characteristic of places that could not be captured by including proxies for initial conditions ($y_{c,1870}$ and $PO_{c,1870}$).

Table 5.10: *GMM vs FE Estimates*

	Growth ($\Delta y_{c,t}$)	
	(FE)	(GMM)
	(1)	(2)
Electrical and Electronics	0.109**	0.155***
	(0.043)	(0.021)
Mechanic	0.048	0.076***
	(0.031)	(0.022)
Chemical	0.052*	0.128***
	(0.027)	(0.019)
Others	0.037*	0.048**
	(0.022)	(0.023)
Foreign (Share)	0.002***	0.001***
	(0.001)	(0.0001)
Establishments per Capita (in logs)	0.741***	0.671***
	(0.043)	(0.025)
Primary Inputs	1.376***	0.800***
	(0.178)	(0.079)
Working Deposits (in logs)	0.045	−0.019
	(0.034)	(0.013)
Federal Employment (Share)	−0.001	−0.001
	(0.001)	(0.001)
State Employment (Share)	−0.001	−0.001
	(0.003)	(0.005)
Local Employment (Share)	0.002	0.002
	(0.003)	(0.002)
PO (in logs)	−0.006	0.043**
	(0.067)	(0.017)
$y_{c,t-1}$	0.193***	0.519***
	(0.075)	(0.019)
Period	1860-1930	1860-1930
Year FE	Yes	Yes
County FE	Yes	Yes
Sargan Test (p-value)		0
Autocorrelation Test (1) (p-value)		0
Autocorrelation Test (2) (p-value)		0.836
Observations	14,866	16,795

*p<0.1; **p<0.05; ***p<0.01

Using the proposed set of GMM instruments, we instrument both the adoption of E&E technologies and the lagged dependent variable ($y_{c,t-1}$) with their lagged values ($E\&E_{c,t-r}$ and $y_{c,t-s}$ for $s > 2$ and $r > 3$, respectively). In this way, we test the robustness of the results in Table 5.5 not only to the inclusion of county-specific FE but also to an alternative but similar construction of the instrument for E&E adoption.

Table 5.11 reports, in the first column, a replication of the baseline 2SLS estimation of Table 5.5 and, in the second column, its GMM counterpart instrumenting E&E adoption and $y_{c,t-1}$ as described above. Note that estimated coefficients of E&E adoption under the GMM and 2SLS procedures are quite similar. The adoption of E&E technologies is still associated with high growth rates under both estimation procedures, even though the periods considered are different and the way in which instruments are constructed also differs. Additionally, note that $\widehat{\rho} = 0.424$ lies in between the FE (0.193) and OLS (0.65) values, as expected.

Again, the pattern of serial correlation in the first-differenced residuals is consistent with the assumptions, showing only a significant first-order serial correlation. However, the validity of the instruments is clearly rejected by the Sargan test of overidentifying restrictions. Because the exogeneity tests cast doubt on the validity of the GMM instruments, we provide an additional robustness check using lagged values of E&E adoption and $y_{c,t-1}$ as instruments but in a 2SLS framework (third column). This specification is analogous to that reported in column (1) but treats $y_{c,t-1}$ as endogenous and instruments it with $y_{c,1880}$. Note that these results are also consistent with previous findings, showing a strong effect of E&E adoption on growth rates[28]. The Sargan test of overidentifying restrictions does not reject the null hypothesis that the error term is orthogonal to the instruments, and the F-test of joint significance strongly rejects the null hypothesis for both first stages.

[28]As in Section 5.6.3 by using $y_{c,1880}$ as an instrument, we are making a non-trivial selection of the sample. This seems to bias the coefficient of E&E adoption downward by approximately 0.1

Table 5.11: *GMM vs 2SLS Estimates*

	Output per Capita ($y_{c,t}$)		
	(2SLS)	(GMM)	(2SLS)
	(1)	(2)	(3)
Electrical and Electronics	0.816***	0.897***	0.597***
	(0.180)	(0.095)	(0.213)
Mechanic	0.092***	0.060***	0.070***
	(0.027)	(0.022)	(0.027)
Chemical	0.014	0.020	0.024
	(0.041)	(0.023)	(0.041)
Others	0.049*	0.019	0.075***
	(0.027)	(0.023)	(0.027)
Foreign (Share)	0.001***	0.001***	0.0005**
	(0.0002)	(0.0001)	(0.0002)
Establishments per Capita (in logs)	0.644***	0.701***	0.548***
	(0.021)	(0.022)	(0.028)
Primary Inputs	1.392***	0.933***	1.168***
	(0.064)	(0.080)	(0.073)
Working Deposits (in logs)	−0.013	−0.010	−0.009
	(0.012)	(0.013)	(0.012)
Federal Employment (Share)	−0.001	−0.0005	−0.001
	(0.001)	(0.001)	(0.001)
State Employment (Share)	0.001	0.0004	0.004
	(0.004)	(0.005)	(0.004)
Local Employment (Share)	0.001	0.003	0.001
	(0.002)	(0.002)	(0.002)
PO (in logs)	−0.017	−0.019	−0.007
	(0.021)	(0.019)	(0.023)
$y_{c,t-1}$	0.535***	0.424***	0.641***
	(0.014)	(0.021)	(0.039)
Period	1900-1930	1860-1930	1900-1930
Weak Instruments E&E (F-statistic)	52.702		75.925
Weak Instruments $y_{c,t-1}$ (F-statistic)			315
Sargan Test (p-value)	0.758	0	0.725
Autocorrelation Test (1) (p-value)		0	
Autocorrelation Test (2) (p-value)		0.345	
Observations	7,547	16,795	6,607

*p<0.1; **p<0.05; ***p<0.01

In sum, this section provides further evidence that the baseline results are robust to alternative statistical methods and econometric specifications, in particular to the possibility that bias in the estimation of $\widehat{\rho}$ or the inability to properly control for county-specific characteristics may have influenced the results. Even though neither of the GMM specifications passed the Sargan test of overidentifying restrictions, they resulted in estimates of the effect of technological adoption that are consistent with previous findings.

5.6.5 Including Additional Endogenous Variables

It is possible that when properly estimated, the effects of adopting other technologies are higher than those of E&E technologies. If this were the case, the alleged E&E differential would not exist.

The validity of the instruments used in the baseline estimation relies heavily on characteristics and historical features that are specific to E&E technologies. For instance, because Chemical and Mechanical technologies were already mature technologies prior to the 1870s, their early adoption may be correlated with characteristics of the place that remain relevant years later. In the case of Chemical technologies, adoption may be correlated with the availability of natural resources, which probably had long-lasting implications for growth. Even though using early adoption as an instrument may not be optimal for Mechanical and Chemical technologies, we test the robustness of the results to treating them as endogenous regressors.

In this subsection, we report how the results change after instrumenting Mechanical and Chemical technologies with lagged values of their adoption. In particular, Tables 5.12 and 5.13 show how the baseline estimates of the effect of E&E adoption on growth and wages change, respectively, as other technologies are instrumented. Column (1) replicates the results obtained in Table 5.5, while columns (2) and (3) sequentially allow Mechanical and Chemical technologies to be treated as endogenous, respectively. In all cases, adoption after the 1900s is instrumented with early adoption (before 1975), as in the baseline case.

Note that in both tables, the baseline results on the effect of E&E adoption on growth and wages barely change after treating Mechanical technologies as endogenous (column two). Additionally, the Sargan test of overidentifying restrictions does not reject the null hypothesis that the error term is orthogonal to the instruments, and the F-test of joint significance strongly rejects the null hypothesis in all cases. The results suggest that the diffusion of Mechanical technologies is associated with lower average wages. Regarding their effect on growth, the results are not statistically significant.

When Chemical technologies are also treated as endogenous (column three), we observe a generalized loss of statistical significance for all estimates of technological adoption. However, note that in the case of the growth regressions, Table 5.12, the estimates lie very close to previous values for both E&E and Mechanical technologies. In Table 5.13, the estimates confirm and amplify the previous differences.

In summary, the baseline results are robust to treating the adoption of other types of technologies as endogenous. They are robust in the sense that estimates of the effect of E&E adoption are stable in magnitude after the inclusion of additional endogenous variables. However, their statistical significance decreases with the number of instrumented variables.

Table 5.12: *Allowing for Multiple Endogenous Variables*

	Growth ($\Delta y_{c,t}$)		
	(1)	(2)	(3)
Electrical and Electronics	0.816***	0.760***	0.901
	(0.158)	(0.145)	(1.104)
Mechanic	0.092***	0.249	0.300
	(0.026)	(0.243)	(0.291)
Chemical	0.014	0.007	−0.164
	(0.038)	(0.033)	(1.027)
Others	0.049*	0.004	−0.001
	(0.028)	(0.075)	(0.078)
Foreign (Share)	0.001***	0.001***	0.001***
	(0.0002)	(0.0002)	(0.0003)
Establishments per Capita (in logs)	0.644***	0.641***	0.641***
	(0.025)	(0.025)	(0.027)
Primary Inputs	1.392***	1.383***	1.373***
	(0.074)	(0.075)	(0.090)
Working Deposits (in logs)	−0.013	−0.015	−0.014
	(0.012)	(0.013)	(0.018)
Federal Employment (Share)	−0.001	−0.001	−0.001
	(0.001)	(0.001)	(0.001)
State Employment (Share)	0.001	0.0002	0.0002
	(0.004)	(0.004)	(0.005)
Local Employment (Share)	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)
PO (in logs)	−0.017	−0.022	−0.021
	(0.020)	(0.021)	(0.022)
$y_{c,t-1}$	−0.465***	−0.469***	−0.468***
	(0.015)	(0.015)	(0.017)
Period	1900-1930	1900-1930	1900-1930
Year FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Weak Instruments E&E (F-statistic)	52.702	52.091	67.933
Weak Instruments Mechanical (F-statistic)		21.303	22.147
Weak Instruments Chemical (F-statistic)			88.65
Sargan (p-value)	0.758	0.722	0.796
Observations	7,547	7,547	7,547
Adjusted R^2	0.272	0.282	0.234

*p<0.1; **p<0.05; ***p<0.01

Table 5.13: *Allowing for Multiple Endogenous Variables*

	Average Wage ($W_{c,t}$)		
	(1)	(2)	(3)
Electrical and Electronics	293.901***	304.277***	482.037
	(43.429)	(43.694)	(426.726)
Mechanic	26.548***	−218.384***	−175.542*
	(7.470)	(76.826)	(106.440)
Chemical	−16.352	18.915*	−166.658
	(11.942)	(11.245)	(408.621)
Others	14.265*	97.711***	92.693***
	(7.538)	(24.492)	(27.892)
Foreign (Share)	−0.011	0.040	0.015
	(0.059)	(0.058)	(0.100)
Establishments per Capita (in logs)	−44.850***	−29.841***	−31.328***
	(6.337)	(7.253)	(8.647)
Primary Inputs	−113.816***	−80.500***	−90.671***
	(19.620)	(22.732)	(32.177)
Working Deposits (in logs)	1.115	4.547	6.808
	(3.404)	(3.843)	(7.232)
Federal Employment (Share)	0.200	0.150	0.293
	(0.200)	(0.241)	(0.395)
State Employment (Share)	1.737	2.928**	2.902*
	(1.374)	(1.468)	(1.635)
Local Employment (Share)	1.569***	1.803***	1.711**
	(0.562)	(0.604)	(0.706)
PO (in logs)	1.298	13.620**	13.457**
	(4.709)	(5.462)	(6.157)
Period	1900-1930	1900-1930	1900-1930
Year FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Weak Instruments E&E (F-statistic)	72.917	69.996	95.788
Weak Instruments Mechanical (F-statistic)		26.496	35.377
Weak Instruments Chemical (F-statistic)			125.872
Sargan (p-value)	0.295	0.656	0.793
Observations	8,129	8,129	8,129
Adjusted R^2	0.727	0.669	0.567

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

CHAPTER 6

CONCLUSION

6.1 Main Findings

Broadly, this dissertation has studied how technological change operates to determine the fortune of places, thus linking the global urgency to understand the current level of disparities in wealth and income (WEF 2017, UNESCO 2017) with the idea that these disparities are, to a large extent, determined by the capacity of individuals to collectively accumulate knowledge and know-how. Ultimately, although technological innovation is the main driving force behind economic growth, industrial development, and increased living standards, only a handful of countries are actively developing new technologies. The United States (US), Western European countries, Japan and South Korea host a small fraction of the world's population but are responsible for most technological advances.

Such a venture required overcoming two major obstacles. On the one hand, it required developing advanced modeling tools capable of characterizing the mechanisms and channels that relate technological change with economic progress. On the other hand, assessing the impact of the emergence, development, and diffusion of key technologies over time and space required constructing a unique database with disaggregated information on the inventive activities of places for the last 150 years. As a result, this dissertation shed light on a number of issues that are crucial to understanding the uneven consequences of technological progress, across and within economies.

Chapter 2 begins by characterizing current state of the global technological production, showing how unequally distributed technological capabilities are. This characterization is comprehensive and detailed enough to identify key obstacles and opportunities for technological diversification.

This chapter fills a gap in the innovation literature, which still has a limited understanding of the process though which countries accumulate new technological capabilities along different stages of their economic development. In fact, cross-country quantitative studies exploring patterns of technological diversification and specialization have been very scarce and are often restricted to the analysis of a handful of developed economies (see, for instance, Archibugi and Pianta (1994) and Cantwell and Vertova (2004)). As a result, there is a need for robust and comprehensive evidence providing a general characterization of the types of technologies that countries are more likely to produce, whether they tend to follow coherent patterns of technological specialization as they develop, and the extent to which technological change is bounded to pre-existing technological capabilities.

The results show that the development of new technologies is a highly cumulative and path-dependent process, in which technological upgrading emerges out of pre-existing knowledge bases and patterns of specialization.The likelihood of diversification into a new technological activity is higher for those domains that are related to countries' existing profile of competences. Countries climb the ladder of technological development rung by rung, as new capabilities have to be accumulated gradually. Additionally, patterns of technological development change in two important ways as countries develop. First, diversification is more heavily constrained by related indigenous capabilities at early stages of development. At later stages of the development process, countries are able to make greater leaps and develop new technologies that are less and less related to their current knowledge bases. Furthermore, the types of technologies in which developed and developing countries specialize are different. During their economic development, countries tend to upgrade their technological structure by specializing in increasingly complex and more valuable technologies. These findings contribute to discussions spanning several fields and can be related to studies such as Bell (2009); Bell and Pavitt (1992, 1997); Breschi et al. (2003); Jaffe (1986) in the innovation literature, Hausmann and Hidalgo (2011); Hausmann et al. (2007); Hidalgo and Hausmann (2009); Hidalgo et al. (2007); Lall (2000) in international trade, or Boschma et al. (2014); Neffke et al. (2011) in economic geography.

As mentioned above, the absence of a long series of disaggregated data on inventive activity has thus far imposed an immediate limitation on the study of the emergence, development, and diffusion of key technologies over time and space. After all, technological diffusion may take decades, and the most interesting cases accurred long before data started to be collected systematically. Chapter 3 addresses this issue by developing a sophisticated machine learning algorithm that is able to identify the locations of inventors and/or assignees within historical patent documents, information that is thus far only available as PDF files. The result of that endeavor is a novel and unique database containing detailed technological information on the inventive activities of places in the United States dating back to 1836. This open source initiative was made freely available to the research community and has been subject to continuous improvements since its introduction, mainly by other researchers also working with it. Therefore, the development of this database provided a unique opportunity for the entire research community to study the evolution of technologies in their historical context.

Chapter 4 takes full advantage of these new data possibilities. It develops a theoretical Ricardian trade model that allows for cross-country heterogeneities in both, the amount of accumulated technological knowledge and its dispersion across sectors. This model extension is empirically tested using the novel dataset developed in Chapter 3.

This chapter is an effort to overcome the other main limitation of the literature, the unsatisfactory role assigned to technological change in economic modeling; which is not rich enough to understand the mechanisms and channels through which technological change operates. It contributes to an extensive literature concerned with the role of technological advance on international trade that dates back to David Ricardo's famous 1817 model, whose recent exponents include Eaton and Kortum (2002), with their general equilibrium multi-country model (EK model henceforth), and the multi-sector extensions of Caliendo and Parro (2014), Chor (2010), Costinot et al. (2012), and Shikher (2011).

Chapter 4 documents the existence of a high degree of heterogeneity in how countries distribute their technological capabilities across sectors, as reflected by their patent activity. More important, this heterogeneity appears to be highly correlated with international trade flows. After incorporating within-country technological dispersion into an EK model, Chapter 4 demonstrates the crucial role of the allocation of technological capabilities in determining trade patterns. Unlike in the EK model, a country's overall comparative advantage (which determines overall exports) can depend on technological dispersion. Countries with lower input costs and low levels of technological accumulation can benefit (increase their overall exports) by lowering the dispersion of their technological activities. In other words, what matters is the covariance between relative input costs and technological dispersion. A country exports more when this covariance is negative, meaning that competitors with higher costs have lower technological dispersion.

Finally, and also exploiting the wealth of information contained in the database developed in Chapter 3, the last chapter of this dissertation investigates how the adoption of disruptive technologies can generate uneven and durable disparities in income and wages. To do so it evaluates the effect of the emergence, development, and geographical diffusion of one of the most revolutionary technologies of the 20th century, electricity.

While theoretical models have advanced greatly, providing a precise and coherent characterization of the economic implications of the diffussion of a GPT (Bresnahan and Trajtenberg (1995), Helpman and Trajtenberg (1998b) ,Helpman and Trajtenberg (1998a), and Aghion and Howitt (2000)), the lack of convincing and comprehensive empirical evidence has called into question the relevance and usefulness of the notion of GPTs (Field, 2008). Therefore, Chapter 5 fills one of the main empirical gaps in the innovation literature by providing comprehensive empirical evidence on the effect of the diffusion of a GPT across the economy. By relying on historical patent documents, this chapter provides a comprehensive view of the emergence, evolution, development, and diffusion of Electrical & Electronic (E&E) technologies in their historical context. It shows that the behavior of E&E technologies between 1860 and 1930 is in line with what can be expected from a GPT, namely: Above-average growth rates in patenting activity, the development of a wide variety of innovation complementarities, and a high degree of pervasiveness in the U.S. inventive structure.

More important, and after combining the database developed in Chapter 3 with economic and demographic data at the county level, this chapter shows that the adoption of E&E technologies had a positive effect on the wages and the income per capita growth of places after the 1900s, generating uneven and durable disparities across counties. Jointly, these two results provide a comprehensive picture of the emergence, evolution, and diffusion of E&E technologies during this period and their effect on the economy.

6.2 Implications for Future Research

Even though this dissertation provided evidence and tools to understand the uneven economic consequences of disruptive technological change, there are still many relevant and urgent questions that remain open. Particularly important among these are questions related to the impact of technological change on the distribution of income within places. A recent article in MIT Technology Review by David Rotman discusses the apparent relationship between disruptive technological change and the increase in income inequality within places. He argues that it does not seem coincidental that the largest homeless camp in the US is located 20 minutes away from the most technologically advanced city in the country, within waking distance of its city hall. The occupation of public space by homeless and marginalized people seems like an undesired consequence of a vertiginous process of technological change, where "prodigious" prosperity meets marginalization (according to Russell Hancock, president of Joint Venture Silicon Valley, the wealth generated in Silicon Valley is "as prodigious as it has ever been").

There is an understudied and promising area of research related to the effect of technological change on income inequality, a natural step forward from this dissertation, which has focused primarily on the disparities generated between rather than within places. In fact, it is not even clear in the innovation literature whether there is a relationship between technological progress and income inequality. If this is the case, the positive between-place impact found in this dissertation could be offset, to some extent, by strong enough negative distributional effects. This also raises the question of whether the mechanisms, channels, and consequences of technological progress differ as we narrow down the unit of analysis.

Another promising avenue for future research is to question the extent to which the effects found here are applicable to all disruptive technologies. Should we expect a different impact when studying other GPT candidates, such as Information and Communication Technologies (ICTs)? Is there something particular to the nature of the technology in question that may generate more or less virtuous or vicious effects? Did these technologies complement or substitute skilled labour and/or unskilled labour? Fortunately, HistPat grants a unique opportunity to study these questions, by providing more than 150 years of disaggregated and detailed data on inventive activity in the US. It is now possible and relatively straightforward to conduct a similar study for the steam engine and ICTs, which will definitely shed light on many of the questions raised previously.

Adopting a broader perspective, HistPat is a great example of how one of the most primitive forms of artificial intelligence, machine learning algorithms, can be used in research and exemplifies the immense amount of possibilities that further developments in this area will bring to research in general. Apart from the immediate research opportunities mentioned above, similar algorithms can be applied to retrieve all sorts of information from historical documents. In the particular case of patent documents, the identification of firms and the disambiguation of inventors is an immediate step forward; such information can be used to study the evolution of inventor and firm networks over time. Additionally, the recent availability of personal information on historical censuses creates the possibility to match the inventive activity of people with all sorts of demographic and personal characteristics. This will allow researchers, among other things, to study the impact of migration flows on the economic and inventive activity of places, people, and firms. Such a level of detail would be unprecedented in research, and it is only a few steps away from this dissertation.

Another area that deserves closer attention involves the role of institutions in fostering or deterring technological progress. Even though the inventive process manifests itself at an individual level (or group of individuals), the institutional environment is a determinant of its success. It does so by aligning incentives, protecting intellectual creations, providing resources, and much more (Jones and Williams, 2000; North, 1990). Recent history provides numerous examples for studying how different institutional, legal, and cultural frameworks relate to innovation and technological progress. These previous experiences can provide answers to contemporaneous, relevant, and urgent questions. They could also be used to understand how institutional frameworks can be designed to mitigate the potential negative consequences of technological progress.

Two hundred years ago, a group of English textile workers and weavers, the Luddites, began destroying weaving machinery as a form of protest. They considered the use of machinery to be a "fraudulent and deceitful manner" to circumvent standard labour practices, fearing that their skills would go to waste as machines would replace them. To what extent are current institutional environments equipped to cope with the uneven consequences of disruptive technological change? Do we have an institutional environment that is strong and flexible enough to respond to the victims of the creative destruction currently being left behind by automation? How better equipped are we today to answer the concerns that the Luddites had exactly 200 years ago?

6.3 Policy Implications

Even though none of the chapters explicitly included policy as a main subject of study, several results of this dissertation have important policy implications. Particularly relevant are those emerging from Chapter 2, which derives valuable insights into countries' development strategies. This chapter shows that developing strong capabilities and leadership in technologies out of the blue is nearly impossible. It is tempting to attempt to become a leader in information technologies or biotechnologies because these technologies are fashionable and profitable, but such a technology-targeted strategy is a lottery. What if the specific knowledge and capabilities required to become a successful producer of the targeted technology are simply not not available or extremely difficult to obtain? In many cases, such a blinded policy will end up in wasted taxes. This issue is particularly important for developing countries, where the possibilities for diversification are even more heavily constrained by indigenous capabilities. The results suggest that a more efficient policy strategy consists of carefully assessing the pre-existing knowledge bases of a country, treating indigenous capabilities as a starting point. The specific analysis of the technological strengths and weaknesses of a country can reveal unexploited opportunities, which may lead to alternative development paths.

It is important to note that a policy supporting only technologies that are very closely related to existing capabilities is also quite likely to fail. Such a development strategy will not be risky in the short run, but it can lock-in the technological development of countries in the long run. This is especially true for developing economies, which tend to make shorter technological jumps (i.e., more related) in less sophisticated activities (i.e., less complex). Therefore, there is a threat for developing countries of becoming locked-in to the production of less-sophisticated technologies, and innovation policies should not reinforce this path-dependent process by narrowing technological opportunities. It is important for developing countries to continuously seek to produce more complex technologies to upgrade their technological structure.

Similar policy implications can be derived from Chapter 4, which argues that the optimal strategy for a country depends on where it stands in the world scenario. Low-cost countries are better off competing on costs, while high-cost countries are better off competing on comparative advantage. Depending on its relative position in the international markets, a country's optimal strategy could be to diversify or to concentrate its innovative activities.

Chapter 5 also provides useful tools that have considerable policy implications. Having a simple and useful means of characterizing GPTs could be used, for instance, to identify technologies that are currently showing above-average growth and complementarity. The early identification and adoption of these sorts of technologies has the potential to foster growth and generate considerable spillovers. Therefore, if Chapter 2 recommends a continuous upgrading of countries' technological portfolios toward more complex and valuable technologies, Chapter 5 also highlights the importance of identifying technologies with the greatest potential for spillovers.

An important caveat of this dissertation is that it focuses mostly on patent activity. Patents are praised as the only systematic measure of invention but are also criticized because they only capture specific types of innovation and technologies. Many generic forms of innovation, especially in developing countries, will not appear in patent data. Similarly, patents do not capture the innovation outcomes that are generated by the imitative activities of firms. These activities can range from pure imitation of existing technologies to the creative design of novel processes and products, the innovative content of which can be considered equally relevant as that in patents. For example, Chapter 5 relies on collecting the traces of information that have been left behind in patent documents, while the full extent of the impact of E&E technologies remains to be documented. This is because the empirical strategy used here does not capture the entire diffusion of E&E technologies. It is widely known that their impact went beyond what can be traced in patent documents. It had implications for the organization of factories, the transportation systems of cities, and other factors.

It is important to bear these critiques in mind when interpreting the results and assessing the policy implications that this dissertation draws. In particular, it is not clear whether the results reported here will also hold true for non-manufacturing sectors, or for sectors relying less on patenting activity. How transferable to the service sector are the experiences documented here for manufacturing? This is a very important question, given the prominent role of services in advanced economies. Additionally, to what extent do technical advances in manufacturing and services depend on, complement, or substitute for each other?. Is this different for the case of disruptive technologies? This could be addressed in future research by combining patents with innovation surveys, both in the manufacturing and service industry. An interesting question here is whether service sectors can also serve as catalyzers, enhancing technological diversification possibilities and magnifying and multiplying their disruptive effects, as they generally span across different types of productive activities.

Bibliography

Acemoglu, D., Moscona, J., and Robinson, J. A. (2016). State capacity and american technology: Evidence from the nineteenth century. *The American Economic Review*, 106 (5): 61–67.

Ager, P. and Brückner, M. (2013). Cultural diversity and economic growth: Evidence from the us during the age of mass migration. *European Economic Review*, 64: 76–97.

Aghion, P. and Howitt, P. (2000). On the macroeconomic effects of major technological change. In *The economics and econometrics of innovation*, 31–53. Springer.

Ahuja, G. (2000). The duality of collaboration: Inducements and opportunities in the formation of interfirm linkages. *Strategic management journal*, 317–343.

Anastasiadis, A. D., Magoulas, G. D., and Vrahatis, M. N. (2005). New globally convergent training scheme based on the resilient propagation algorithm. *Neurocomputing*, 64: 253–270.

Archibugi, D. and Pianta, M. (1994). Aggregate convergence and sectoral specialization in innovation. *Journal of Evolutionary Economics*, 4 (1): 17–33.

Audretsch, D. B. and Feldman, M. P. (1996). R&d spillovers and the geography of innovation and production. *The American economic review*, 86 (3): 630–640.

Balland, P.-A. and Rigby, D. (2017). The geography of complex knowledge. *Economic Geography*, 93 (1): 1–23.

Bell, M. (2009). Innovation capabilities and directions of development.

Bell, M. and Pavitt, K. (1992). Accumulating technological capability in developing countries. *The World Bank Economic Review*, 6 (suppl 1): 257–281.

Bell, M. and Pavitt, K. (1997). Technological accumulation and industrial growth: contrasts between developed and developing countries. *Technology, globalisation and economic performance*, 83137: 83–137.

Bloom, N., Draca, M., and Van Reenen, J. (2015). Trade induced technical change? the impact of chinese imports on innovation, it and productivity. *The Review of Economic Studies*, rdv039.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of econometrics*, 87 (1): 115–143.

Blundell, R. and Bond, S. (2000). GMM estimation with persistent panel data: an application to production functions. *Econometric Reviews*, 19 (3): 321–340.

Bolatto, S. (2013). Trade across countries and manufacturing sectors with heterogeneous trade elasticities. *Centro Studi Luca d'Agliano Development Studies Working Paper*, (360).

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*, dtu012.

Breschi, S., Lissoni, F., and Malerba, F. (2003). Knowledge-relatedness in firm technological diversification. *Research policy*, 32 (1): 69–87.

Breschi, S., Malerba, F., and Orsenigo, L. (2000). Technological regimes and schumpeterian patterns of innovation. *The economic journal*, 110 (463): 388–410.

Bresnahan, T. (2010). General purpose technologies. *Handbook of the Economics of Innovation*, 2: 761–791.

Bresnahan, T. F. and Trajtenberg, M. (1995). General purpose technologies 'engines of growth'? *Journal of econometrics*, 65 (1): 83–108.

Burchardi, K., Chaney, T., and Hassan, T. A. (2016). Migrants, ancestors, and foreign investments.

Bustos, P. (2011). Trade liberalization, exports, and technology upgrading: Evidence on the impact of mercosur on argentinian firms. *The American economic review*, 101 (1): 304–340.

Caliendo, L. and Parro, F. (2014). Estimates of the trade and welfare effects of nafta. *The Review of Economic Studies*, 229–242.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, 29 (2): 238–249.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2012). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*.

Cantwell, J. and Vertova, G. (2004). Historical evolution of technological diversification. *Research Policy*, 33 (3): 511–529.

Chor, D. (2010). Unpacking sources of comparative advantage: A quantitative approach. *Journal of International Economics*, 82 (2): 152–167.

Cohen, W. M. and Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative science quarterly*, 128–152.

Colombelli, A., Krafft, J., and Quatraro, F. (2014). The emergence of new technology-based sectors in european regions: A proximity-based analysis of nanotechnology. *Research Policy*, 43 (10): 1681–1696.

Costinot, A., Donaldson, D., and Komunjer, I. (2012). What goods do countries trade? a quantitative exploration of ricardo's ideas. *The Review of Economic Studies*, 79 (2): 581–608.

Crafts, N. (2004). Steam as a general purpose technology: a growth accounting perspective. *The Economic Journal*, 114 (495): 338–351.

Crafts, N. and Mills, T. C. (2004). Was 19th century british growth steam-powered?: the climacteric revisited. *Explorations in Economic History*, 41 (2): 156–171.

Dahlman, C. J., Ross-Larson, B., and Westphal, L. E. (1987). Managing technological development: lessons from the newly industrializing countries. *World development*, 15 (6): 759–775.

Daniels, R. (1990). *Coming to America*. Harper Collins Publishers.

David, P. and Wright, G. (1999). General purpose technologies and surges in productivity: Historical reflections on the future of the ict revolution. Technical report, Economics Group, Nuffield College, University of Oxford.

David, P. A. (1990). The dynamo and the computer: an historical perspective on the modern productivity paradox. *The American Economic Review*, 80 (2): 355–361.

Dornbusch, R., Fischer, S., and Samuelson, P. A. (1977). Comparative advantage, trade, and payments in a ricardian model with a continuum of goods. *The American Economic Review*, 823–839.

Dosi, G., Freeman, C., Nelson, R., Silverberg, G., Soete, L., et al. (1988). *Technical change and economic theory*, volume 988. Pinter London.

DuBoff, R. (1979). *Electrical Power in American Manufacturing 1889-1958*. New York: Arno Press.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 1741–1779.

Eaton, J. and Kortum, S. (2010). Technology in the global economy: A framework for quantitative analysis. *Unpublished manuscript.*

Edquist, C. and Lundvall, B.-A. (1993). Comparing the danish and swedish systems of innovation. *National innovation systems: A comparative analysis*, 265–298.

Enos, J. L. (1991). *The creation of technological capability in developing countries.* Thomson Learning.

Essletzbichler, J. (2015). Relatedness, industrial branching and technological cohesion in us metropolitan areas. *Regional Studies*, 49 (5): 752–766.

Feldman, M. P. and Kogler, D. F. (2010). Chapter 8 - stylized facts in the geography of innovation. In Hall, B. H. and Rosenberg, N., editors, *Handbook of The Economics of Innovation, Vol. 1*, volume 1 of *Handbook of the Economics of Innovation*, 381 – 410. North-Holland.

Feldman, M. P. and Yoon, J. W. (2012). An empirical test for general purpose technology: an examination of the cohen–boyer rdna technology. *Industrial and Corporate Change*, 21 (2): 249–275.

Field, A. J. (2003). The most technologically progressive decade of the century. *The American Economic Review*, 93 (4): 1399–1413.

Field, A. J. (2006). Technological change and us productivity growth in the interwar years. *The Journal of Economic History*, 66 (01): 203–236.

Field, A. J. (2008). Does economic history need gpts? *Available at SSRN 1275023*.

Fleming, L. (2001). Recombinant uncertainty in technological search. *Management science*, 47 (1): 117–132.

Freeman, C. (1989). *Technology policy and economic performance*. Pinter Publishers Great Britain.

Frenken, K., Van Oort, F., and Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional studies*, 41 (5): 685–697.

Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The elements of statistical learning*, volume 1. Springer series in statistics Springer, Berlin.

Furman, J. L., Porter, M. E., and Stern, S. (2002). The determinants of national innovative capacity. *Research policy*, 31 (6): 899–933.

Goldfarb, B. (2005). Diffusion of general-purpose technologies: understanding patterns in the electrification of us manufacturing 1880–1930. *Industrial and Corporate Change*, 14 (5): 745–773.

Greenwood, J. (1997). *The third industrial revolution: technology, productivity, and income inequality*. Number 435. American Enterprise Institute.

Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change. *Econometrica*, 25 (4): 501–522.

Günther, F. and Fritsch, S. (2010). neuralnet: Training of neural networks. *The R journal*, 2 (1): 30–38.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001a). The nber patent citation data file: Lessons, insights and methodological tools. Technical report, National Bureau of Economic Research.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001b). The nber patent citation data file: Lessons, insights and methodological tools. Technical report, National Bureau of Economic Research.

Hall, B. H., Trajtenberg, M., et al. (2006). Uncovering general purpose technologies with patent data. *Chapters*.

Hatton, T. J. and Williamson, J. G. (2008). The impact of immigration: Comparing two global eras. *World Development*, 36 (3): 345 – 361.

Hausmann, R. and Hidalgo, C. A. (2011). The network structure of economic output. *Journal of Economic Growth*, 16 (4): 309–342.

Hausmann, R., Hwang, J., and Rodrik, D. (2007). What you export matters. *Journal of economic growth*, 12 (1): 1–25.

Helpman, E. and Trajtenberg, M. (1998a). Diffusion of general purpose technologies. In Helpman, E., editor, *General purpose technologies and economic growth*, 85. MIT Press.

Helpman, E. and Trajtenberg, M. (1998b). A time to sow and a time to reap: Growth based on general purpose technologies. In Helpman, E., editor, *General purpose technologies and economic growth*, 55. MIT Press.

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.

Israel, P. B. E. and Edison, A. (1998). A life of invention.

Jaffe, A. B. (1986). Technological opportunity and spillovers of r&d: evidence from firms' patents, profits and market value.

Jaffe, A. B., Trajtenberg, M., and Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *the Quarterly journal of Economics*, 108 (3): 577–598.

Jeffrey, R. S. (2002). 'new and improved' direct marketing: A non-parametric approach. In *Advances in Econometrics*, 141–164. Emerald Group Publishing Limited.

Jones, C. I. and Williams, J. C. (2000). Too much of a good thing? the economics of investment in r&d. *Journal of Economic Growth*, 5 (1): 65–85.

Jovanovic, B. and Rousseau, P. L. (2005). General purpose technologies. *Handbook of economic growth*, 1: 1181–1224.

Katila, R. and Ahuja, G. (2002). Something old, something new: A longitudinal study of search behavior and new product introduction. *Academy of management journal*, 45 (6): 1183–1194.

Kerr, W. R. (2013). Heterogeneous technology diffusion and ricardian trade patterns. Technical report, National Bureau of Economic Research.

Khan, B. Z. (2005). *The Democratization of Invention: patents and copyrights in American economic development, 1790-1920*. Cambridge University Press.

Kim, L. (1999). Building technological capability for industrialization: analytical frameworks and korea's experience. *Industrial and corporate change*, 8 (1): 111–136.

King, K. and Fransman, M. (1984). *Technological capability in the third world.* Springer.

Klevorick, A. K., Levin, R. C., Nelson, R. R., and Winter, S. G. (1995). On the sources and significance of interindustry differences in technological opportunities. *Research policy*, 24 (2): 185–205.

Koren, M. and Tenreyro, S. (2007). Volatility and development. *The Quarterly Journal of Economics*, 122 (1): 243–287.

Kortum, S. S. (1997). Research, patenting, and technological change. *Econometrica: Journal of the Econometric Society*, 1389–1419.

Lall, S. (1992). Technological capabilities and industrialization. *World development*, 20 (2): 165–186.

Lall, S. (2000). The technological structure and performance of developing country manufactured exports, 1985-98. *Oxford development studies*, 28 (3): 337–369.

Laursen, K. (1999). The impact of technological opportunity on the dynamics of trade performance. *Structural Change and Economic Dynamics*, 10 (3): 341–357.

LeCun, Y., Boser, B. E., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W. E., and Jackel, L. D. (1990). Handwritten digit recognition with a back-propagation network. In *Advances in neural information processing systems*, 396–404.

Lee, K. (2013). *Schumpeterian analysis of economic catch-up: Knowledge, path-creation, and the middle-income trap.* Cambridge University Press.

Lee, K. and Lim, C. (2001). Technological regimes, catching-up and leapfrogging: findings from the korean industries. *Research policy*, 30 (3): 459–483.

Leromain, E. and Orefice, G. (2014). New revealed comparative advantage index: dataset and empirical distribution. *International Economics*, 139: 48–70.

Levchenko, A. A. and Zhang, J. (2016). The evolution of comparative advantage: Measurement and welfare implications. *Journal of Monetary Economics*, 78: 96–111.

Levinthal, D. A. and March, J. G. (1993). The myopia of learning. *Strategic management journal*, 14 (S2): 95–112.

Levitt, B. and March, J. G. (1988). Organizational learning. *Annual review of sociology*, 14 (1): 319–338.

Li, G.-C., Lai, R., D'Amour, A., Doolin, D. M., Sun, Y., Torvik, V. I., Amy, Z. Y., and Fleming, L. (2014). Disambiguation and co-authorship networks of the us patent inventor database (1975–2010). *Research Policy*, 43 (6): 941–955.

Lileeva, A. and Trefler, D. (2010). Improved access to foreign markets raises plant-level productivity...for some plants. *The Quarterly Journal of Economics*, 125 (3): 1051–1099.

Lipsey, R. G., Carlaw, K. I., and Bekar, C. T. (2005). *Economic transformations: general purpose technologies and long-term economic growth.* OUP Oxford.

Lundvall, B.-Å. (2010). *National systems of innovation: Toward a theory of innovation and interactive learning*, volume 2. Anthem Press.

MacDonald, J. M. (1985). R & d and the directions of diversification. *The Review of Economics and Statistics*, 583–590.

Malerba, F. (1992). Learning by firms and incremental technical change. *The economic journal*, 102 (413): 845–859.

Malerba, F. (2002). Sectoral systems of innovation and production. *Research policy*, 31 (2): 247–264.

Malerba, F. and Orsenigo, L. (1995). Schumpeterian patterns of innovation. *Cambridge Journal of Economics*, 19 (1): 47–65.

Marin, A. and Petralia, S. (2015). Sources and contexts of inter-industry differences in technological opportunities: the cases of argentina and brazil. *Documento de Trabajo CENIT*.

Metcalfe, J. S. (1995). Technology systems and technology policy in an evolutionary framework. *Cambridge Journal of Economics*, 19 (1): 25–46.

Moser, P. and Nicholas, T. (2004). Was electricity a general purpose technology? In *The American Economic Review, Papers and Proceedings*, volume 94, 388–394.

Nebeker, F. (2009). *Dawn of the electronic age: Electrical technologies in the shaping of the modern world, 1914 to 1945*. John Wiley & Sons.

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.

Nelson, R. R. (2008). Economic development from the perspective of evolutionary economic theory. *Oxford development studies*, 36 (1): 9–21.

Nelson, R. R. and Rosenberg, N. (1993). Technical innovation and national systems. *National innovation systems: A comparative analysis*, 322.

Nelson, R. R. and Winter, S. G. (1997). An evolutionary theory of economic change. *Resources, Firms, and Strategies. A Reader in the Resource-Based Perspective. Oxford ua S*, 82–99.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, 1417–1426.

Niosi, J., Saviotti, P., Bellon, B., and Crow, M. (1993). National systems of innovation: in search of a workable concept. *Technology in society*, 15 (2): 207–227.

North, D. C. (1990). *Institutions, institutional change and economic performance.* Cambridge university press.

Orrok, G. A. (1930). Pearl street station the first central station in the world. *Transactions of the Newcomen Society*, 11 (1): 98–108.

Park, K.-H. and Lee, K. (2006). Linking the technological regime to the technological catch-up: analyzing korea and taiwan using the us patent data. *Industrial and Corporate Change*, 15 (4): 715–753.

Patel, P. and Pavitt, K. (1994). National innovation systems: why they are important, and how they might be measured and compared. *Economics of innovation and new technology*, 3 (1): 77–95.

Patel, P. and Pavitt, K. (1997). The technological competencies of the world's largest firms: complex and path-dependent, but not much variety. *Research policy*, 26 (2): 141–156.

Penrose, E. (1959). The theory of the firm. *NY: John Wiley & Sons*.

Perez, C. and Soete, L. (1988). Catching up in technology: entry barriers and windows. *Technical change and economic theory. CF Giovanni Dosi, Richard Nelson, Gerald Silverberg, Luc Soete. London, Pinter*, 458–479.

Petralia, S., Balland, P.-A., and Rigby, D. L. (2016). Unveiling the geography of historical patents in the united states from 1836 to 1975. *Scientific Data*, 3.

Phene, A., Fladmoe-Lindquist, K., and Marsh, L. (2006). Breakthrough innovations in the us biotechnology industry: the effects of technological space and geographic origin. *Strategic Management Journal*, 27 (4): 369–388.

Piketty, T. (2014). *Capital in the Twenty-First Century*. Harvard University Press.

Riedmiller, M. and Braun, H. (1993). A direct adaptive method for faster backpropagation learning: The rprop algorithm. In *Neural Networks, 1993., IEEE International Conference on*, 586–591. IEEE.

Rigby, D. L. (2015). Technological relatedness and knowledge space: entry and exit of us cities from patent classes. *Regional Studies*, 49 (11): 1922–1937.

Rodrik, D. (2008). *One economics, many recipes: globalization, institutions, and economic growth*. Princeton University Press.

Rosenberg, N. (1998). Chemical engineering as a general purpose technology. *General purpose technologies and economic growth*, (7): 167–192.

Rosenberg, N. and Trajtenberg, M. (2004). A general-purpose technology at work: The corliss steam engine in the late-nineteenth-century united states. *The Journal of Economic History*, 64 (01): 61–99.

Rowley, T., Behrens, D., and Krackhardt, D. (2000). Redundant governance structures: An analysis of structural and relational embeddedness in the steel and semiconductor industries. *Strategic management journal*, 369–386.

Saviotti, P. P. and Frenken, K. (2008). Export variety and the economic performance of countries. *Journal of Evolutionary Economics*, 18 (2): 201–218.

Schmookler, J. (1966). Invention and economic growth.

Service, U. S. P. (2007). *The United States Postal Service: An American History, 1775-2006*. Government Relations, US Postal Service.

Shikher, S. (2011). Capital, technology, and specialization in the neoclassical model. *Journal of international Economics*, 83 (2): 229–242.

Simonovska, I. and Waugh, M. (2011). The elasticity of trade: Evidence and estimates. Technical report, CESifo Working Paper Series.

Soete, L. and Wyatt, S. (1983). The use of foreign patenting as an internationally comparable science and technology output indicator. *Scientometrics*, 5 (1): 31–54.

Sokoloff, K. L. (1988). Inventive activity in early industrial america: evidence from patent records, 1790-1846.

Steinwender, C. (2015). The roles of import competition and export opportunities for technical change. *Centre for Economic Performance (CEP) Discussion Paper 1334*.

Teece, D. J. (1982). Towards an economic theory of the multiproduct firm. *Journal of Economic Behavior & Organization*, 3 (1): 39–63.

Teece, D. J., Rumelt, R., Dosi, G., and Winter, S. (1994). Understanding corporate coherence: Theory and evidence. *Journal of Economic Behavior & Organization*, 23 (1): 1–30.

Tibshirani, R., Hastie, T., Narasimhan, B., and Chu, G. (2003). Class prediction by nearest shrunken centroids, with applications to dna microarrays. *Statistical Science*, 104–117.

Vertova, G. (1999). Stability in national patterns of technological specialisation: some historical evidence from patent data. *Economics of innovation and new technology*, 8 (4): 331–354.

Vertova, G. (2001). National technological specialisation and the highest technological opportunities historically. *Technovation*, 21 (9): 605–612.

Wooldridge, J. M. (1995). Selection corrections for panel data models under conditional mean independence assumptions. *Journal of econometrics*, 68 (1): 115–132.

Wright, G. (1990). The origins of american industrial success, 1879-1940. *The American Economic Review*, 651–668.

Yip, G. S. (1982). Diversification entry: Internal development versus acquisition. *Strategic Management Journal*, 3 (4): 331–345.

Nederlandse samenvatting

In grote lijnen bestudeerde deze dissertatie hoe technologische verandering werkt om het fortuin van plaatsen te bepalen. Bijgevolg koppelt de wereldwijde urgentie om het huidige niveau van verschillen in rijkdom en inkomen te verstaan met het idee dat deze verschillen grotendeels worden bepaald door de capaciteit van individuen om samen kennis en knowhow te accumuleren.

Voor een dergelijke onderneming vereiste het overwinnen van twee belangrijke obstakels.Enerzijds vereiste het het ontwikkelen van geavanceerde modelistischen instrumenten die de mechanismen en kanalen kunnen karakteriseren die technologische veranderingen met de economische vooruitgang verbinden. An de andere kant, was het voor het beoordeling van de impact van de opkomst, ontwikkeling en verspreiding van belangrijke technologieÃ«n in de tijd en ruimte nodig om een unieke database met geaggregeerde informatie over inventieve activiteiten van plaatsen voor de laatste 150 jaar op te bouwen.

Uit de resultaten blijkt, dat de ontwikkeling van nieuwe technologieÃ«n een zeer cumulatief en padafhankelijk proces is, waarbij technologische upgrading uit bestaande kennisgebieden en specialisatiepatronen voortkomt. De kans op diversificatie in een nieuwe technologische activiteit is hoger voor die domeinen die verband houden met het bestaande competentieprofiel van de landen. Landen beklimmen de ladder van technologische ontwikkeling rang op rang, aangezien nieuwe mogelijkheden geleidelijk opgebouwd worden moeten. Bovendien wordt de technologische diversificatie in de vroege ontwikkelingsfasen door gerelateerde inheemse capaciteiten zwaarder beperkt.

In dit dissertatie wordt ook het bestaan van een hoge graad van heterogeniteit ten aanzien van hoe landen hun technologische capaciteiten over sectoren verspreiden beschreven. Belangrijker nog is dat deze heterogeniteit zeer gecorreleerd te zijn lijkt met de internationale handelsstromen.

Om dit uit te leggen behandelt hoofdstuk 4 de cruciale rol van de toewijzing van technologische mogelijkheden om handelspatronen door het integreren van technologische binnenlandse verspreiding in een handelsmodel van Eaton-Kortum te bepalen. Als een resultaat kan kan dit model uitleggen hoe het algemene comparatieve voordeel van de landen (die de totale uitvoer bepaalt) van de technologische verspreiding afhankelijk zijn. Landen met lagere invoerkosten en een laag niveau van technologische accumulatie kunnen door de verspreiding van hun technologische activiteiten profiteren (hun totale uitvoer verhogen).

Hoofdstuk 5 vult een van de belangrijkste empirische leemten in de innovatieliteratuur omdat uitgebreide empirische bewijzen over het effect van de diffusie van een algemene gebruikstechnologie (GPT) over de economie verstrekt worden. Door het vertrouwen op historische patentendocumenten, geeft dit hoofdstuk een uitgebreid beeld van de opkomst, evolutie, ontwikkeling en verspreiding van Elektrische & Elektronische (E&E) technologieÃ«n in hun historische context. Het blijkt dat het gedrag van E&E technologieÃ«n tussen 1860 en 1930 in lijn is met wat van een GPT kan worden verwacht, namelijk: bovengemiddelde groeipercentages bij octrooienactiviteiten, de ontwikkeling van een breed scala aan innovatie complementariteiten en een hoge graad van doordringing in de Amerikaanse inventieve structuur.

Belangrijker nog, en na het combineren van de database die is ontwikkeld in Hoofdstuk 3 met economische en demografische gegevens op provinciaal niveau, blijkt hoofdstuk 5 dat de goedkeuring van E&E technologieÃ«n positief effect heeft op lonen en de groei per inwoner per inwoner na de jaren 1900, waardoor ongelijk en duurzaam verschillen in de provincies worden gegenereerd.

ACKNOWLEDGEMENTS

This dissertation is the ultimate result of a collective effort of many people who, conciously or not, contributed to it. The latter group includes my family and friends, José, Marta, Daiana, David, Pablo, Julieta, Machi, Carolina, José Maria, Silvia, Carito, Tamy, Ariel, Marcos, Silvio, Pablito, Vale, and Tango. Even though Tango was a dog, he was the only one who believed I was pursuing a Ph.D; the rest still think that I am living off welfare benefits somewhere in Europe.

In the first group, there are many people who were essential to this never-ending learning process. In chronological order, I will always be indebted to Alejo Macaya and Pablo Fajfar from the University of Buenos Aires, whose generous and engaging approach to education were crucial for my future academic development. I am also indebted to all my mates and professors at the master's program at the Universidad de San Andrés, in particular to Mica, Mili, Gonzo, Rama, Virginia, Diego, Angelita, Renzo, El Hongo, El Colo, Maris, Ivo, Santi, El Negro, Anita, Iva, Manu, Enrique, Tamara, Alicia, and Federico. Walter Sosa-Escudero, Marcela Svarc, and Juan Carlos Hallak deserve a special acknowledgement, as they were the ones who, probably against the odds, decided to invest much of their time to get me as far as they could.

I learnt immensely during my time at Penn State University, mostly thanks to Hiroki, Younghwa, Wang, Bulat, Zhang, Jan, Leo, Haerim, Berni, Gustavo, Martin, Bin, Talat and all my mates there. Special thanks go to Stephen Yeaple, Alexander Monge-Naranjo, Neil Wallace, Paul Grieco, Vijay Krishna, Peter Newberry, and John Riew.

I enjoyed the stimulating company of many great people at Utrecht University, in particular that of Nicola, Mathieu, Teresa, Dario, Nynke, Marijke, Marianne, Niccolò, Rodrigo, Martijn, Ricardo, Joan, Wendy, the new acquisitions Andrea and Tom, and all my colleagues at the department who were always very supportive, especially Chris, who always backed up our ambitious ideas with funding. The Netherlands is a wonderful place to live, and I will always be grateful for having shared my time there with Ornella, Javi, Peter, Christian, Daniel, Patty, Maureen, Lisanne, Mathilde, Fer, Michelle, Nadia, Horacio, Mariano, and Marjet. I leave Utrecht feeling indebted, especially after having met Anne.

I would also like to acknowledge the vital role of my co-authors and mentors, who were the main reason that this was such an enjoyable process. I take with me the open-mindedness of David and Pierre-Alex, who are always willing to play against the odds, the patience, generosity, and understanding attitude of Andrea, who managed an incredibly complicated project wonderfully, the far-reaching academic vision and the considerate advice of Ron, the disorganized creativity of Anabel, and the determination of Caro and Fer.

Finally, I feel incredibly fortunate to have met such wonderful colleagues. Maryann Feldman, Koen Frenken, Cesar Hidalgo, Alessandro Nuvolari, Francesco Lissoni, and Frank Van Oort brought distinction to this dissertation by being part of the evaluation comitee. Always ready to help were Julio Raffo, Petra Moser, Frank Nefke, Jérôme Vicente, Chris Esposito, Frank Van der Wouden, Christoph

Roesler, Elisabeth Perlman, Rosina Moreno, Elena Mas Tur, Andrés López, Lilia Stubrin, and all the members of the CENIT in Buenos Aires.

Thanks to you all.

Curriculum Vitae

Sergio Petralia was born in Buenos Aires, Argentina. He holds a licenciatura en economia from the University of Buenos Aires (2007) and a masters degree in economics from the University of San Andrés (2009) in Buenos Aires, and from Penn State University (2013) in the United States. In both master programs he was awarded with a scholarship to pursue his studies.

Before joining Utrecht University in 2013, he participated in several research projects for institutions such as the Inter-American Development Bank, the International Development Research Centre, and the University of San Andrés. He received several grants and scholarships, most notably, a full Doctoral Fellowship at Penn State University (2011) and Utrecht University (2013), and a position of Visiting Scholar at the Science Policy Research Unit (SPRU) of the University of Sussex (2010), and at the University of California at Los Angeles (2016).

He finishes his Ph.D. with two academic publication, both part of his thesis. "*Unveiling the geography of historical patents in the United States from 1836 to 1975*" was published in Scientific Data, Nature Publishers, in 2016, while "*Climbing the Ladder of Technological Development*" was published in Research

Policy in 2017. During this period he also contributed in two book chapters that are already published. "*Evaluating the Impact of Open Access Initiatives within Academia and Beyond. Made in Latin America: Open Access, Scholarly Jounals, and Regional Innovations.*" was published in 2016 while "*Anatomy of Distributive Changes in Argentina. The Economies of Argentina and Brazil: A Comparative Perspective.*" was published in 2011.