# Three Essays in the Economics of Innovation: Geography, Green Technologies and Policy Uncertainty

Enrico Bergamini

University of Turin & Collegio Carlo Alberto

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics

Thesis supervisor: Francesco Quatraro

Vilfredo Pareto Doctorate in Economics, Cycle XXXVI

Academic years: 2020-2024

September 2024

# Abstract

This thesis investigates technological change through three essays, focusing on geography, environmental innovation, and policy. The first essay explores the role of inward and outward Foreign Direct Investments (FDIs) in fostering recombinant novelty at the regional level, focusing both technological distance and functional characteristics of FDIs in generating local novelty. The second essay investigates green diversification in U.S. cities, emphasizing how local skill composition interacts with green FDIs on branching into environmentally sustainable technologies. The third essay studies the effects of climate policy uncertainty on environmentally directed technical change across European firms, highlighting the importance of clear policies to direct technological efforts toward low-carbon innovation. Together, these essays contribute to understanding how external connectivity, local capacities, and policy signals shape innovation, providing empirical evidence for regional development and green industrial policy.

# Contents

# Acknowledgments

The Ph.D. journey is full of challenges, and even though very often is a solitary journey, it is never a path that one walks alone. This thesis is not only the product of my efforts, but also of the wonderful community of academics that improved this work, through their feedback, suggestions and questions. I am greatly thankful to all of you that inspired the writing of these Chapters in between Turin and Barcelona.

First, I want to extend my gratitude to my supervisor Francesco Quatraro, for his guidance and patience over these four years, for making me grow as a researcher and including me in his team. I also want to thank my co-authors at the University of Turin, Fabrizio Fusillo, Gianluca Orsatti and Alessandra Scandura.

I am very grateful to Jeroen van den Bergh, for his motivation and encouragement, and warm welcome at ICTA in Barcelona. I also want to thank Rosina Moreno, for her kind and warm hospitality at the AQR research group at the UB. I am also deeply grateful to Emmanuel Mourlon-Druol and Georg Zachmann for their friendship and mentorship; to Jacob Jordaan and Francesco Papadia for their kind and generous words over the years.

During my first two years in Turin, I have been lucky enough to encounter great colleagues at Collegio Carlo Alberto. I want to express my gratitude to all the staff at Collegio. This experience would not have been the same without a fantastic group of PhDs at Collegio. I am deeply grateful to Elisa, with whom I shared much of the ups and downs of this journey, for her sincere friendship; and to Guido, who is a constant source of laughter and support, a fantastic friend and mentor. My experience in Barcelona was shaped by the gift of a supportive and inspiring community, I am immensely thankful to all who contributed to making these times so enriching.

There are never words that can quite summarize the meaning of the unconditional love and support that one is lucky enough to receive from friends and family. I feel privileged for having everyday on my side Catarina, Giovanni, Gaston and Sveva, no matter kilometers and years. To my family, Marinella, Marco and Fabio: you are the roots of all I am. Thank you for always reminding me of what is so important and simple, that I sometimes forget it. You are my inspiration, this thesis is dedicated to you.

# Chapter 1

# Introduction

This thesis is composed of three essays exploring technological change from different angles. Long recognized as the engine of long-term economic growth, technological change - and its spatial and environmental dimensions - are now central to the policy agenda. The three essays are interconnected by their focus on understanding the mechanisms and dynamics of technological change in the context of global and local forces. Together, they investigate how regions and firms navigate the interplay between external knowledge flows and internal capabilities to foster innovation, with particular attention to the green transition.

The chapters collectively emphasize the importance of geography in shaping innovation outcomes, while also addressing distinct perspectives. The first two chapters are connected by a focus on geography, studying the relationship between regional innovation performance, and the external connectivity to global knowledge pools through Foreign Direct Investments (FDIs). The first chapter focuses on the determinants of recombinant novelty at the local level, while the second on the process of diversification into green technological domains. The second two chapters come together around the theme of the climate transition, and the development of environmentally-sustainable technologies, from the different perspective of cities in the United States and firms in Europe. In the context of addressing the climate crisis, environmental innovation is the fundamental element at the crossroads of growth and emissions reduction. In times of renewed interest in green industrial policy and the need for implementing strong climate policies - bearing in mind the potential geographical inequalities generated by the transition - understanding the determinants of environmentally-directed technological change is crucial.

The first essay is coauthored with Francesco Quatraro and Alessandra Scandura, and focuses on recombinant novelty from a regional perspective. The motivation for this chapter lies in the need to understand the spatial dynamics of recombinant innovation, as a result of the external connectivity of European regions. Bringing together different literature streams on the geography of innovation, recombinant growth, and evolutionary economic geography, we investigate the role of FDIs in fostering the recombination of knowledge at the local level. Studying patenting activity in European NUTS3 provinces, between 2003 and 2017, we take a recombinant approach and explore how the connectivity brought by greenfield FDIs might favour regions in accessing foreign knowledge elements, in turn stimulating their capacity for recombination and potentially leading to new technological trajectories. We identify patents which make unprecedented combinations at the local level, and study their relationship with both inward and outward FDIs. In addition, we employ a measure of proximity between FDIs and the local knowledge base of the home regions, as technological distance between internal and external knowledge might play a significant role in expanding the local technological search space, and in turn foster recombinant innovation.

The main findings suggest that inward FDIs (IFDIs) are positively associated with local recombinant novelty, while OFDIs are negatively associated with it, although with more nuanced results. We uncover a high degree of geographical heterogeneity by exploring the different associations of FDIs depending both on the origin and destination of the flows, and by observing different samples for EU15 countries and non EU15 countries. We also find that proximity of FDIs (in both directions) to the the local knowledge base is negatively associated with recombinant novelty, suggesting that more distant knowledge elements available to regions through external linkages might be more effective in expanding local recombination possibilities. By employing spatial econometric models, we also find evidence of different effects in terms of spatial spillovers, contributing to the idea that the concentration of knowledge returns and recombinant capabilities in space is fostered by the presence of external connectivity. Finally, we also unpack the functional characteristics of FDIs, finding that greenfield FDIs focused on Research and Development

(R&D) projects are more positively associated to recombinant novelty, and particularly so in terms of OFDIs.

The second essay is coauthored with Fabrizio Fusillo, Gianluca Orsatti, Francesco Quatraro and Alessandra Scandura, and investigates the interplay between the local endowments of skills, and the presence of external knowledge linkages proxied by inward FDIs. This chapter focuses on the determinants of diversification into green technological domains, contributing to the literature of evolutionary economic geography and eco-innovation. We study a panel of 287 cities (Metropolitan Statistical Areas, MSAs) in the United States, for the period 2003-2018 and consider IFDIs as agents of technological change, driving the process of branching into green technologies together with the local skills composition. We specifically consider FDIs in green-intensive sectors, and explore the moderating roles of abstract and routine skills in the relationship between FDIs and green diversification of cities. The interactions between FDIs and the local skills composition reveal how absorptive capacities, both potential and realized, contribute to the integration and exploitation of external knowledge, particularly in green technological development. The findings suggest that MSAs with higher levels of green FDIs and abstract skills have a greater likelihood of developing new green technological specializations. Moreover, we show how abstract skills are more important in regions with fewer green FDIs, acting as a compensatory factor, while routine skills enhance the effects of FDI on green technological diversification.

The third essay instead takes on the perspective of firms, and investigates empirically the relationship between climate policy uncertainty (CPU) and directed technical change (DTC) across European firms. This study is motivated by the need to understand how uncertainty around climate policy-making influences firms' innovation choices. By collecting a novel dataset of newspaper articles, and using text-as-data techniques, I develop new indexes for CPU in four European countries: France, Germany, Italy, and Spain. I explore how CPU can influence the direction of technological innovation considering both low-carbon and polluting technologies patents filed by firms. The analysis builds on the DTC literature, and connects it with that of the empirics of policy uncertainty. I build sub-indexes for both positive and negative-leaning uncertainty, namely pointing towards increased or decreased probability of future stringency, and test econometrically the effects of CPU on patenting. I adopt an empirical model of DTC and firm-level data matched with patent portfolios, controlling for the path-dependency in the environmental innovation process. The findings show that firms respond to CPU's direction in terms of their technological efforts. Policy uncertainty implying an increase in regulatory stringency, is positively associated to green innovation, and negatively to polluting patents. On contrary, when policy uncertainty indicates a weakening or a setback climate policy-making firms stick to polluting technologies and divest from green ones. The results highlight that clear and consistent climate policies are crucial to pushing firms towards green innovation, and away from polluting innovation, bearing important policy implications for steering the direction of technological change.

Evolutionary Economic Geography and Economic Complexity are the theoretical and empirical backbone for the thesis. While the first two essays share a regional perspective in terms of the empirical analysis, all the chapters focus on innovation as the engine for economic growth. Methodologically, this work is in line with the tradition of analyzing innovation employing patent data as its proxy. The thesis investigates how regions and firms can adapt to the challenges of globalization and climate change, attempting to provide empirical evidence to policymakers on urgent issues shaping the current policy debate.

# Chapter 2

# Recombinant novelty and Foreign Direct Investments: evidence from European Regions

**Chapter co-authored with Francesco Quatraro and Alessandra Scandura**

## Abstract

This work investigates the determinants of technological recombinant novelty at the local level across European regions, focusing on the role of inward and outward greenfield Foreign Direct Investments (FDIs). We conduct an empirical analysis on a panel of European NUTS3 regions observed from 2003 to 2017. The results show that inward FDIs are positively associated with regional technological novelty consistently across several estimations. Outward FDIs are generally negatively associated with technological novelty but the relationship changes when we account for the geographical distribution of FDIs and the knowledge intensity of the investments. We also employ a proximity measure between FDIs and the local knowledge base, which we find to be negatively associated with novel recombination efforts. The results of this work are relevant both for the academic discourse on local and non-local determinants of regional technological performances and for the policy discussion on the relevance of global connectivity for regional economic development.

## 2.1 Introduction

Technological novelty has long attracted academic research interest in innovation studies, given its relevance for understanding the rate and direction of innovation and technological change. The wide body of literature that has focused on the dynamics underlying the introduction of novelty has proposed two distinct views grounded respectively in Marshallian's gradualism and Schumpeter's saltationism. Accordingly, on the one hand, novelty is regarded as the outcome of the slow accumulation of variations and improvements in knowledge and technologies; on the other hand, the evolutionary approach stresses how it stems from a recombination process suddenly leading to the emergence of new paradigms and major technological breakthroughs (Antonelli 2007; Strumsky and Lobo 2015).

In recent times, a new wave of empirical studies has focused on analyzing the economic effects of technological novelty. Understanding technological novelty is especially important from a geographical perspective because of its impact on the change of the local economic structure (Quatraro 2012; Antonelli 2014). Extant literature has documented a positive link between technological novelty and a firm's ability to generate impactful innovation and sustain persistent, innovative efforts (Arts and Veugelers 2015; Carnabuci and Operti 2013a). More recently, technological novelty has been found to be positively connected with an increased likelihood of generating environmentally friendly patents and developing regional specialization in green technological domains (Orsatti, Quatraro and Pezzoni 2020b; Orsatti, Quatraro and Scandura 2023).

Because of the proven significance of technological novelty, recent efforts have also been devoted to enhancing the measurement of technological novelty (Verhoeven et al. 2016) and comprehending potential drivers from a geography of innovation perspective. Within this line of analysis, considerable attention has been given to the composition of local knowledge pools, particularly the diversity of knowledge and the influence of related versus unrelated variety, as well as to the presence of local highly fungible knowledge pools, like those characterizing the generation of key enabling technologies (Castaldi, Frenken et al. 2015; Castaldi and Los 2017; Mewes 2019; Berkes and Gaetani 2021; De Noni and Belussi 2021; Montresor, Orsatti et al. 2023).

Yet, the literature developed so far has overlooked the role of non-local drivers for regions' capability to develop novelty in local contexts. A recent debate has been connecting the literature in evolutionary economic geography (EEG) and that of Global Production Networks (GPN), highlighting gaps and potential complementarities (Yeung 2021; Boschma 2022a). Recent works in EEG have started to address the gaps, stressing the importance of foreign-owned firms for regional structural change, measured as the capability of regions to diversify into new unrelated industrial activities (Neffke et al. 2018a; Elekes et al. 2019). The present study aims to contribute to this field of inquiry by connecting the analysis of technological novelty in regional contexts to the literature on non-local agents of structural change. In particular, we follow Iammarino (2018) and focus on the impact of regional connectivity, i.e. *the exposure of a place to the inflows and outflows of assets, knowledge, capabilities and expertise from and towards the rest of the world* (Iammarino 2018, p.157), which we measure with global foreign direct investments (FDIs). These latter can be considered as carriers of different forms of knowledge through which Multinational enterprises (MNEs) create economic connections between territories characterized by heterogeneous technological, industrial and scientific structures (Iammarino and McCann 2013; Castellani, Marin et al. 2022a).

Combining the well-established recombinant knowledge approach (Weitzman 1998) with the economic geography and geography of innovation literature on MNEs (Iammarino and McCann 2013; J. Cantwell and Iammarino 2005), we contend that FDIs affect the opportunities for local agents to access heterogeneous pools of knowledge and hence to implement unprecedented combinations of ideas and knowledge components. In doing so, we discriminate between the role of incoming and outgoing greenfield FDIs for regional novelty. Furthermore, we postulate that the potential impact on the generation of recombinant novelty is related to the extent of similarity between FDIs and the economic activities of local firms.

Our empirical analysis focuses on European NUTS3 regions over the period included between 2003 and 2017. We employ panel data models along spatial econometric ones, uncovering the role of spatial spillovers and adding geographical nuance to the analysis. Our results show a positive association between inward FDIs and recombinant novelty of European regions. Outward FDIs are generally negatively associated with local novelty, although our evidence on this matter is more mixed. We uncover heterogeneity in the relationship between novelty and FDIs, both due to the knowledge-intensity of investments, and due to geographical differences within Europe. In addition, we employ a measure of relatedness-density between FDIs and the local knowledge base, found to be negatively associated with novel recombination efforts at the regional level.

The contribution of this work is manyfold. In the first place, our theoretical framework and empirical analysis shift the focus from the usual innovation output measures (e.g. patent counts) to a more comprehensive and representative measure of technological novelty at the regional level. The latter mirrors regions' capabilities to put in place innovative efforts directed at generating effective new knowledge, which we measure by looking only at patents entailing unique recombination of technological classes. Secondly, we stress the relation between regions' innovative efforts and the extent of their global connectivity throughout FDIs, hence zooming in at the sub-national scale to look closer at local technological dynamics, as advocated by recent works on the topic (see e.g. Iammarino (2018) and Elekes et al. (2019)). Thirdly, we include in our study both incoming and outgoing FDIs. In particular, outward FDIs have been largely overlooked both in scholarly communities and in the policy debate for a number of reasons mostly revolving around the argument about employment destruction and consequent wage depression at home (Iammarino 2018). Yet, offshoring is increasing steeply, notably in advanced economies, hence definitely raising a call for further research on this topic. Fourthly, we explore the heterogeneity of both inward and outward FDIs, notably with respect to their geographical location and direction (extra- versus intra-EU) and type of activities (R&D versus non-R&D activities), which allow to better qualify the relation between regions' novelty and their global connectivity.

The rest of the paper proceeds as follows. In Section 2.2 we review the background literature and develop hypotheses. In Section 2.3 we present the datasets used for the analysis, the variables of interest, and our estimation strategy. Section 2.4 presents and discusses the results while Section 2.5 provides the concluding remarks.

## 2.2 Background literature

### 2.2.1 The role of non-local drivers for regions' technological novelty

Following the recombinant approach, innovation can be conceived as the outcome of the incremental search process through which agents recombine knowledge elements (Weitzman 1998; Nelson and Winter 1982). Novel inventions can be byproducts of existing components recombined in new ways that, if successful, open further technological trajectories (Weitzman 1998; Fleming 2001). As a matter of fact, a rich empirical literature has operationalized indicators of recombinant novelty, capturing unprecedented combinations of technologies within patents (Arts and Veugelers 2015; Verhoeven et al. 2016; Arts, Hou et al. 2021).

Technological novelty, defined as the creation of unprecedented combinations of knowledge elements, offers a deeper understanding of innovation quality compared to standard metrics like patent counts. Berkes and Gaetani (2021) underlines that unconventional innovations often emerge in densely populated urban areas, where diverse knowledge pools and frequent interactions foster the recombination of disparate ideas. Similarly, Mewes (2019) demonstrates that larger metropolitan areas exhibit superlinear scaling in the production of atypical knowledge combinations, reflecting the enhanced capacity of these regions to generate novel technologies. These findings reinforce the idea that fostering technological novelty allows regions to enhance their recombinant capabilities, ultimately promoting economic growth and diversification.

Indeed, the search process over the knowledge space results from the capabilities of firms and regions to channel knowledge internally (Carnabuci and Operti 2013a) and on the availability of external knowledge elements. Likewise, Boschma (2017a) underscored the importance of recombinant innovations resulting from the successful combination of both local and external knowledge, which can help regions to exit stagnant technological paths and result in a process of unrelated diversification. Recent empirical works at the regional level have focused on the role of relatedness and regional capabilities for recombinant novelty within patents. In addition to the role of relatedness, scholars have investigated the role of universities (Plunket and Starosta de Waldemar 2023; Giorgi et al. 2024) and academic inventors (Orsatti, Quatraro and Scandura 2024) as agents of technological change, enhancing the recombination of local and non-local knowledge.

Over the past two decades, the EEG literature has largely investigated the drivers of growth, innovation and structural change from a geographical perspective. Starting from Frenken et al. (2007), much work has focused on the path-dependency of technological and industrial change. The concept of relatedness has been recognized as a fundamental driver of diversification into new industrial and technological specializations, from numerous different perspectives (Hidalgo et al. 2018). Traditionally, the focus of this literature has been on the role of regional internal capabilities for developing novel knowledge elements. Recent critical readings of this literature have highlighted the need to connect these findings with the International Business literature (Iammarino 2018) and that of Global Production Networks (Yeung 2021). Until recently, the literature had overlooked the staple role external connectivity, through which foreign agents act as non-local drivers of regional growth and innovation (Iammarino 2018; Boschma 2017a). As explained by Iammarino (2018): *Connectivity, an essential but somehow disregarded di-*

*mension of territorial equity and economic development policy, extends far beyond the idea of 'attractiveness': connected places are flows recipients as well as senders.*

Local actors embedded in a territory do not only interact within it, but in a global, networked and complex context that spans well beyond their geographical boundaries (Yeung 2021). The role of the global connectivity and its interaction with each peculiar territorial condition has famously been described as a "local buzz" (Storper and Venables 2004) with "global pipelines" (Bathelt, Malmberg et al. 2004a). A variety of non-local agents are responsible for connecting the local economy to global markets, opening up the possibility to tap into heterogeneous pools of knowledge. Recently, scholars have started to consider the role of non-local actors in affecting the local economy, with notable examples including MNEs and migrant inventors (e.g. Neffke et al. (2018a), Elekes et al. (2019), Crescenzi, Di Cataldo et al. (2021) and Miguelez and Morrison (2023)).

Iammarino and McCann (2013) highlight that the access to GPNs, facilitated by FDIs, enable regions to access external knowledge and resources, enhancing local innovation capabilities and enabling knowledge spillovers. FDIs are carriers of different forms of knowledge through which MNEs create economic connections between territories characterized by heterogeneous technological, industrial and scientific structures (Iammarino and McCann 2013; Castellani, Marin et al. 2022a). The connections between firms and regions through FDIs can be established in several different forms, rendering their effects on innovation complex and multifaceted. External linkages can be created by non-local actors entering the region or by the internationalization of firms outside the home region. Importantly, the directionality of flows - inward and outward - exposes regions to two types of co-existing knowledge flows that may have heterogeneous effects on local innovation dynamics.

### 2.2.2 Inward FDIs

Inward FDIs have long been recognized as a key driver of regional growth (Iammarino and McCann 2013). The positive relationship with regional productivity and innovation can be generated through technological transfers and spillovers to the local innovation system. Traditionally, MNEs have also been considered as more advanced firms compared to firms which are not internationalizing. MNEs spend more in R&D, are more productive, and characterized by a technological superiority and broader expertise if compared to the local ones (Castellani and Zanfei 2007).

Knowledge transfers from the home to the host regions can be of tacit or non-tacit nature. The latter is the explicit knowledge introduced through the flow of novel products and services. Domestic firms can learn directly from MNEs through vertical or horizontal linkages, as non-local MNEs can integrate in the territory through the local supply chain, directly interacting with local suppliers, both downstream and upstream (Castellani, Meliciani et al. 2016). Tacit knowledge, instead, concerns all those components that are "embedded" and cannot be simply inferred by reverse-engineering products:

know-how, expertise, company culture, technical standards, requirements for local producers (Amendolagine, Presbitero et al. 2019). Among others, García et al. (2013) showed how technological spillovers from MNEs to local firms are a critical component of innovation dynamics and technological upgrading in local firms present in the host regions. Through FDIs, MNEs can be vectors of tacit and explicit knowledge components that domestic firms can absorb, adapt and, crucially for our application, recombine.

Tacit knowledge is often disseminated through informal mechanisms such as on-the-job training, mentorship, and collaborative problem-solving, as emphasized by Ernst and Kim (2002). These informal mechanisms complement formal channels like structured training programs and standardized procedures in contributing to knowledge transfers. The effectiveness of these mechanisms largely depends on the absorptive capacity of local firms—their ability to recognize, assimilate, and apply new knowledge. FDIs, therefore, can stimulate regional innovation through local knowledge spillovers, which can by direct or indirect (labour mobility, buyer-supplier networks, etc.; see for example Antonietti, Bronzini et al. (2015)). Through these mechanisms of knowledge transfers, MNEs could therefore increase the local technological search space, fostering the creation of atypical recombination.

Furthermore, extensive literature has investigated the effect of FDIs on innovation through agglomerations (Burger and Meijers 2016). The co-location of local and non-local actors within the same region can give rise to knowledge externalities of both Jacobian and Marshallian nature, from which regions can benefit (Iammarino 2018). Technology transfers, and the ability to generate (and profit from) local spillovers, also depends on firms' capacity to assimilate and apply new technologies and on the economic environment. Crescenzi, Gagliardi et al. (2015) find evidence for intra-industry knowledge spillovers to local firms in the United Kingdom, as a result of greater investments by MNEs. More recently, Jin et al. (2019) emphasized that regions with high absorptive capacity—defined as the ability to recognize, assimilate, and apply external knowledge are more likely to benefit from these technological spillovers. Absorptive capacity, hence, enables local firms to incorporate the new knowledge into their own innovation processes, leading to technological diversification and a better innovative performance. Y. Huang and Yan Zhang (2020) also find evidence of firm spillovers in innovation for IFDI in China.

Outside of knowledge transfer and collaborative effects, another channel through which inward FDIs are expected to impact innovation in the host regions concerns the pro-competitive effects in the receiving region. Competition arises over local inputs for the innovation process, for example specialized labour force, necessary to gain a competitive advantage and increase productivity (Castellani, Castellani et al. 2006). As reviewed in Antonietti, Bronzini et al. (2015), many works have focused on the effect of inward FDIs (both greenfield and M&As) onto the performance of local firms. Increased competition can push incumbents to a reallocation of resources. Critically in terms of long and short-term effects, FDIs may have negative effects in the short term if the capital reallocation is not fast enough (Antonietti, Bronzini et al. 2015). At the country level, evidence in this direction is found between inward FDIs and the complexity index by Antonietti and Franco (2021). In our case we expect that a new player entering the region could increase competitive pressure by reducing the space for technological exploitation (especially given MNEs' technological superiority), and in the longer term force local competitors to increase exploration over the technological space, producing novel technologies to remain competitive on the market.

Recent empirical evidence from regional-level studies on connectivity and innovation has been investigating the effects of IFDIs on a variety of outcomes in Europe. Antonietti, Bronzini et al. (2015) and Ascani and Gagliardi (2015) found evidence for Italy, showing that IFDIs in services can stimulate patenting at the provincial level. Elekes et al. (2019) bring forward evidence for the specialization patterns of Hungarian provinces, showing that foreign-owned firms contribute to unrelated diversification. From a network perspective, Ascani, Bettarelli et al. (2020), reconstruct the ownership structure of firms present in Italian regions, finding evidence for the role of connectivity in fostering regional innovation, with important nuances in terms of the structure of the networks. Ascani, Balland et al. (2020a) add on the spatial and sectoral heterogeneity in the relationship between FDIs and local innovation. Observing Italian provinces, they do not find evidence of spatial spillovers from FDIs to neighbourhood regions. They add nuance to the sectoral spillover effects, finding that only IFDIs in knowledge-intensive branches of manufacturing ("Science-based" and "Specialised supplier") positively impact other sectors with inter-industry linkages.

Crescenzi, Dyèvre et al. (2022) sheds light on the effect of spillovers through greenfield IFDIs in R&D intensive activities, on the emergence of high-innovation clusters in regions. They find evidence on two mechanisms: both knowledge spillovers to local firms and the attraction of subsequent IFDIs, showing that MNEs can not only directly affect technological change but also act as catalysts for attracting other foreign players. Similarly, Bello et al. (2023) also study knowledge-intensive FDIs through both greenfield investments and M&As. They study the potential increase in availability of knowledge components due to external linkages, specifically focusing on green and digital technologies, finding that FDIs can indeed expand the knowledge base of these regions. In terms of spillovers, two recent papers add nuance on the internal characteristics that can moderate the effects external knowledge on local innovation, employing data on Brazilian regions. Garcia, Araujo, Mascarini, Gomes Santos et al. (2023) find a positive role for FDIs and patenting, while showing that the effect on innovation performances are higher in regions characterized by a more diverse industrial composition. Garcia, Araujo, Mascarini, Santos et al. (2024) instead elaborate on the role of two types of absorptive capacities, showing how both institutional and industrial capacities (respectively proxied by R&D in universities and firms) can positively moderate spillover effects.

While evidence about the IFDIs and innovation is generally positive, recently more mixed results have added nuance in terms of competition effects, the timing of the dynamics on regional innovation, and the sectoral aspect of knowledge spillovers (Rojec and Knell 2018). Damioli and Marin (2024), in Europe, explore this heterogeneity in terms of entry mode (greenfield vs M&A for European regions) finding a negative effect between greenfield IFDI and the total amount of patenting. They provide evidence on how greenfield entry can displace local teams, by hiring more senior inventors, resulting in a detrimental effect on more junior ones.

Another channel through which IFDIs could be acting on the regional possibility for recombination, is by introducing competitive pressures, through market-stealing and resource reallocation. As explained in Aitken and Harrison (1999) foreign firms, due to their superior technology and scale, can crowd out domestic firms by capturing market share and limiting access to critical inputs. Similarly, Ascani and Gagliardi (2020) emphasize the role of absorptive capacity, where regions and firms with lower capabilities are more vulnerable to these effects, as they struggle to compete with technologically advanced foreign entrants. This displacement may on the one hand hinder local firms' productivity and innovation potential, but on the other put pressure for a more exploratory innovation, resulting in atypical combinations in inventions.

Nevertheless, much of the extant evidence points towards a positive, though nuanced, effect of greenfield FDIs onto local innovation. Many of these empirical exercises have focused on aggregate count of patents as an outcome variable proxying for innovation performance. In contrast with the extant literature on IFDIs and regional innovation, here we consider the recombinant characteristics of this patents, and how FDIs might be acting as external linkages for regions, providing them with a larger set of recombination possibilities. Hence, in the case of novelty, we expect that IFDIs, through knowledge spillovers and technological transfers, are able to introduce novel knowledge components, broadening the technological search space available for recombination. Moreover, the pro-competitive effects spurred by foreign entry may also push the incentives of local firms to recombine, increasing explorative innovation, and in turn driving the generation of local recombinant novelty. In line with these arguments, we test the following hypothesis for IFDIs:

> *Hypothesis 1:* Inward greenfield FDIs are positively associated with local recombinant novelty in EU regions.

### 2.2.3 Outward FDIs

In contrast with IFDIs, OFDI are commonly perceived as a negative phenomenon for the home economy, as linked to the offshoring of production activities resulting in a destruction of local employment. OFDI policies have historically been leaning towards the maximisation of net flows, attracting investments while retaining firms as much as possible within the territory (Iammarino 2018). However, the effects of OFDIs have been subject to much less academic attention in comparison with IFDIs, especially when considering innovative outcomes and at the regional level (Bathelt and Buchholz 2019). Increasingly, opposing evidence has been brought forward, suggesting that OFDIs might have a positive impact on the home economy. While on the one hand offshoring of production and job loss could be responsible for negative outcomes at the regional level, a more nuanced vision of connectivity as a two-way flow of knowledge has been emerging in the economic geography literature, pointing towards positive effects of internationalization on the home economy (Crescenzi and Iammarino 2018).

In recent work on firms, Valacchi et al. (2021) finds a positive effect of internationalization on the patenting activities of MNEs. Another stream of literature finds similar evidence for the internationalization of Chinese firms, particularly in the context of the Belt and Road initiative, (e.g. Fu et al. (2018), Zhou et al. (2019), Xing Li et al. (2021) and Yongmin Zhang et al. (2024)), and for that of Indian MNEs (Amendolagine, Piscitello et al. 2022; Reddy et al. 2022). The effects of internationalization, however, are also subject to a large degree of sectoral and geographical heterogeneity. Recent critical readings of this literature have cautioned about the ambiguity of these results in terms of the direction of the effects for Chinese firms (R. Yang and Bathelt 2022) as a result of empirical limitations, and a lack of focus in the data on knowledge dynamics.

However, empirical evidence in terms of regional innovation is still scant and the effect of OFDIs can be ambiguous. Castellani and Pieri (2016) analyze the impact of OFDIs on regional productivity in Europe, finding that OFDIs in manufacturing are negatively associated with productivity growth, while investments in sales, distribution, and marketing enhance local productivity, especially when directed outside the EU. Opposing forces affect home regions' performance after firms internationalize, largely depending on the type of investment, and on the geographical origin and destinations of the investments. In terms of mechanisms, the first channel between OFDIs and local innovation is that of reverse knowledge transfers, positively affecting home regions through backward linkages. Similarly to the case of IFDI, outward connectivity can create the opportunity for tapping into a more diverse foreign knowledge base, with internationalizing firms learning abroad and bringing back novel knowledge components.

Connectivity through internationalization of local firms can open up new markets for them, increasing efficiency and productivity, exerting a positive effect not only on the MNEs investing outside of the region, but also in aggregate for home regions and cities (Bathelt, Buchholz and J. A. Cantwell 2023). MNEs themselves can become more productive, learn on international markets, and become by consequence more innovative and productive within their home economy. OFDIs are in fact associated with higher economies of scale and scope, that incentivise investment in R&D activities (Petit and Sanna-Randaccio 2000). OFDIs also open up the chance for regions to improve their ability to source and exploit foreign knowledge (Fosfuri and Motta 1999). By adapting and interacting in the new environments, MNEs learn from the foreign knowledge base and might absorb and recombine in novel and unexpected ways, increasing the likelihood of discovery (Buchholz et al. 2020). Therefore, as MNEs tend to gain in size and productivity compared firms remaining in the domestic market (Helpman et al. 2004; Bannò et al. 2014; Bertrand and Capron 2015; Cozza et al. 2015), they might directly affect the home region themselves through reverse knowledge transfers.

At the regional level, Ascani, Bettarelli et al. (2020) find evidence for internationalisation of firms as a key mechanism for local learning opportunities through reverse knowledge transfers. However, evidence on OFDIs and aggregate regional innovation performance is still scant (Castellani, Mancusi et al. 2015; Iammarino, McCann and Ortega-Argilés 2018; Bathelt and Buchholz 2019) and results are mixed. A recent stream of empirical exercises, from an economic geography perspective, study the relationship between OFDIs and income levels in the United States. Bathelt and Buchholz (2019) provides fresh evidence on OFDIs from US cities, showing that they exert a positive effect on median incomes. One of the channels investigated to explain the positive role of OFDIs is that reverse knowledge transfers. OFDIs provide access to foreign knowledge bases, and knowledge flowing backwards not only enhances own-firm performance but spills over to local agents at home, which in turn can absorb it and recombine novel elements, in turn increasing median incomes.

Buchholz et al. (2020) adds on this results by looking at OFDIs' effects on inequality and aggregate income distribution rather then at income only. They argue that exposure to OFDIs can increase geographical inequality by the increasing returns to the stock of knowledge in a location. Increasing returns are also depending on labour markets: the presence of MNEs successfully internationalizing exerts a positive effect on the local innovation also through higher attractiveness to high-skill workers, in turn reinforcing local knowledge spillovers. If higher-income cities (regions) are apparently disproportionately benefiting from knowledge spillovers, the virtuous (vicious) circles of accumulation through increasing returns will increase the polarization of local labour markets and drive up inequality (Bathelt, Buchholz and J. A. Cantwell 2023). The authors underscore that knowledge dynamics at the regional level have been under-investigated, and can vary depending on host location. These aspects provide further motivation for our focus on the geography of recombinant knowledge, adding evidence on its spatial dynamics.

In contrast to the idea of reverse knowledge transfers and spillovers, the literature has investigated a competing effect from OFDIs: a "hollowing-out" of local resources (for a comprehensive review in the case of knowledge-intensive FDIs, see D'agostino (2015)). The "hollowing-out" occurs when critical innovation resources, such as R&D and skilled labor, are relocated abroad, leaving the home region with diminished capacities to generate technological advancements. Recent empirical literature shows that the effect could be negative if the firm-level gains do not offset the aggregate loss of value added due to offshoring abroad (e.g. Castellani and Pieri (2016)). In such case, outward FDIs may adversely affect the overall balance of payments and exports as well as domestic employment and skills (e.g. Crinò (2009) and Gagliardi et al. (2021)). The offshoring of R&D can result in a gradual substitution of locations in the innovation process, with demand for high-skill competence moving outside of the home region. In addition, a detrimental effect of outward connectivity might take place in case of barriers to (backward) knowledge flows, impeding the transmission of knowledge from host to home regions. On the one hand the subsidiary in the host region is exposed to the novel knowledge elements of the subsidiary, and act as a bridge transferring them back to the headquarters, which in turn could absorb them and recombine them into novel technologies. On the other, barriers could stop this knowledge flows if the subsidiary is not sufficiently embedded in the MNE network, or if cultural and institutional differences between locations (and companies) is too high (D'agostino 2015).

As in the case of IFDIs, geographical heterogeneity also increases the complexity of these mechanisms (Castellani, Castellani et al. 2006; Castellani and Pieri 2016). Evidence on the geographical spillovers from OFDIs have been also underinvestigated. In addition, in terms of geographical heterogeneity, Damijan et al. (2017) also adds nuance when considering emerging regions in Eastern Europe, in conjunction to the scope of OFDIs. Flows focused on low-cost production often fail to generate meaningful improvements in parent firm performance or regional innovation. Their study on new EU member states shows that while OFDIs can enhance productivity when targeting advanced economies, those aimed at cost reduction in less developed countries often result in minimal knowledge spillovers and limited positive outcomes for the home economy. In addition, the degree of relatedness and complementarity can also affect significantly the effects of OFDIs, as shown for the case of manufacturing and services in European by Ascani, Bettarelli et al. (2020). Finally, the heterogeneity found in the literature in terms of effects and mechanisms, depending on the knowledge intensity of sectors, and the scope of OFDI flows (D'agostino 2015; Cozza et al. 2015; Valacchi et al. 2021; Bathelt, Buchholz and J. A. Cantwell 2023) are very relevant in our case, considering recombinant novelty.

In summary, we expect two opposite mechanisms in OFDIs to mainly affect regional recombinant novelty. In the case of a successful and effective transfer of knowledge from host to home, regions might be tapping into foreign knowledge pools, and able to recombine foreign knowledge components locally. On the contrary, OFDIs might also put regions at risk of a hollowing-out of innovation processes to foreign locations. Hence, our expectation is that both phenomena might be happening, and there can be competing explanations on the potential impact of OFDIs on recombinant novelty. Therefore, both positive and negative outcomes likely expected:

> *Hypothesis 2a*: Outward greenfield FDIs are positively associated with (local) recombinant novelty in EU regions.

*Hypothesis 2b*: Outward greenfield FDIs are negatively associated with (local) recombinant novelty in EU regions.

### 2.2.4 Proximity between FDIs and local capabilities

Having outlined the possible channels through which IFDIs and OFDIs can affect recombinant novelty, we add on the distance between external and internal knowledge components. As mentioned, the internal variety and diversity characterising regional knowledge composition affects the possibilities for new combinations, leading to breakthrough innovations (Fleming 2001; Verhoeven et al. 2016). Castaldi, Frenken et al. (2015) have shown how both related and unrelated variety, within regions, are the building blocks for technological recombination. Regions with a higher degree of unrelated knowledge might be better able to explore new recombination, opening technological trajectories connecting previously disconnected fields. The transfer of external knowledge affects this knowledge composition, exposing regions to a broader set of innovations. Hence, in the context of FDIs, we expect that the (dis)similarity between local and non-local knowledge components to be a relevant dimension for the generation of recombinant novelty. As highlighted by Boschma (2017a), the gradient of distance between incoming and local knowledge can influence the potential for knowledge spillovers. Also in the case of regional recombinant novelty, we consider the distance between internal knowledge and FDIs as a factor at play. If the knowledge is too similar, it may not provide enough novelty to spur potential for recombination: connecting to more distant technologies could expand the search space and positively affect recombination probabilities. Elekes et al. (2019) find that foreign firms showing a higher deviation from the region's capabilities, induce more unrelated diversification in regions, as compared with domestic firms. Increased distance indicates that foreign firms are more likely to induce structural change by introducing unrelated diversification in regions, and particularly so in peripheral, rather than capital, regions. They find that a higher distance increases the probability for breakthrough, as IFDIs are allowing access to knowledge distinct from the region's knowledge base.

In our context, we also expect that a larger expansion in the search space can lead to the generation of novel recombination. Building on these insights, we propose the following hypothesis to test the relationship between the proximity of IFDIs to the local knowledge base and the generation of recombinant novelty:

> *Hypothesis 3*: The proximity between IFDIs and the local knowledge base is negatively associated with recombinant novelty.

A symmetric reasoning could be applied in the case of the proximity between the *host* knowledge base and OFDIs: the distance between OFDI and foreign knowledge would be expected to be positive, as the home region would be increasing the space for exploration tapping into a more distant knowledge base abroad. However, as explained in more detail in the next Section, we can only observe the relationship between *home* knowledge base and the OFDIs originating from there. While the role of distance is rather intuitive in the case of IFDIs, that between *home* knowledge base and OFDI is less so. A higher technological proximity between the OFDI and the local knowledge base indicates that the MNE is more embedded in its regional innovation system, playing a more central role within the technological structure. Firms' centrality within the regional productive structure, on the one hand, might mean that the MNEs going on the international markets through OFDIs is more productive, and therefore more likely to effectively transfer knowledge through backward linkages (Bathelt and Buchholz 2019). In this case, the proximity of OFDIs to the regional knowledge base would exert a positive effect on recombinant novelty. On the other hand, if the dominating effect is instead the "hollowing-out" of local resources, a higher degree of proximity between MNEs and their region's technological production could increase the negative effects on the home economy. In this case, in terms of incremental innovation, loosing more central elements would deplete even more knowledge resources in favour of foreign regions. As mentioned in the previous section, a vicious circle (Bathelt, Buchholz and J. A. Cantwell 2023) could be triggered within the home labour market. Triggered by OFDI of (technologically) highly embedded firms, closer to the core of the technological structure of the region, the "hollowing-out" could be further reinforced by loss of high-skilled workers, decreasing local knowledge spillovers and the possibility of effectively recombine knowledge. For these reasons, the expectations in terms recombinant novelty on the distance between OFDIs and regional knowledge are less clear-cut, and mirror our expectations about the role of OFDIs, leading us to the following alternative hypotheses:

*Hypothesis 4a*: The proximity between OFDIs and the local knowledge base is positively associated with recombinant novelty.

*Hypothesis 4b*: The proximity between OFDIs and the local knowledge base is negatively associated with recombinant novelty.

## 2.3 Empirical framework

### 2.3.1 Data and variables

We test the hypotheses developed in the previous Section, and investigate the relationship between recombinant novelty and FDIs, building a balanced panel dataset for 1136 European NUTS3 regions, over the period 2003-2017. The choice of this time-frame is related to the availability of FDI data, which are only available starting from 2003. We employ patent data to proxy for the innovative activity of regions, and construct indicators of recombinant novelty. In particular, we rely on the OECD REGPAT database, collecting patent applications to the European Patent Office (EPO). REGPAT provides geolocated data, at the NUTS3 European level, for both patent inventors and applicants. Based on the location of inventors, we assign patents to NUTS3 provinces, double-counting patents collaboratively produced across different regions.

We construct our dependent variable as the number of patents that combine originally pairs of technologies, following Verhoeven et al. (2016) and Orsatti, Quatraro and Scandura (2023). A patent can be filed to the EPO under a set of different technological codes, classified through the Cooperative Patent Classification (CPC). First, we compare all the unique combinations of 4-digit CPC codes attributed to a given patent with all the unique combinations appeared before that patent, within that same region. The logic of local novelty is that a patent can be recombining for the first time a technology for that region, and does not necessarily need to be novel for the whole set of patents observed. We rely on 4-digit CPC codes rather than lower digits, in order to avoid inflating the novel patents that combine very similar technologies. Based on this notion, we construct the indicator of recombinant novelty (variable name: $Novelty$), which will serve as our dependent variable throughout the analysis, as the number of novel patents for NUTS3 regions, between 2003 and 2017.

In terms of indicators of recombinant novelty, similar methodologies have been applied by Mewes (2019) and Berkes and Gaetani (2021), who analyzed atypical combinations in metropolitan areas, and by Montresor, Orsatti et al. (2023), who explored regional recombinant capabilities to generate technological novelty. The choice to focus on a local indicator of recombinant novelty reflects the need to capture how regions adapt and recombine knowledge in ways that are unique to their contexts. This approach builds on recent works emphasizing the importance of measuring innovation quality rather than quantity, such as through the creation of unprecedented technological combinations within a region. Local measures capture the unique interplay between external flows and the existing regional knowledge base, driving place-specific innovation (Iammarino 2018) and the potential emergence of new local varieties. Future extensions could incorporate broader indicators, such as those capturing novelty at the national or global levels could provide further insights into this dynamics.

We create FDI indicators using the Financial Times' fDiMarkets database. This database contains information about greenfield FDI flows, observed at bilateral investment level, collected from the year 2003 onward. Greenfield FDIs are cross-border investment projects for which the firm is opening a new establishment in a foreign country, rather than acquiring or merging with already existing firms. As discussed in Castellani and Pieri (2016), our main estimates refer to the association between novelty and newly established projects, one of several types of external linkages through which firms can tap into foreign knowledge. For each project, fDiMarkets provides information on the level of capital investment (CAPEX), as well as a sectoral classification, which we reconcile with the North American Industry Classification (NAICS) classification.

We exploit information about FDIs addresses (provided at the city-level) for investing companies and target locations in order to assign each project with its target (IFDI) or source (OFDI) in a NUTS3 region. We aggregate information about FDI flows into several variables. We work out our main FDI regressors as counts of projects, according to the permanent inventory method, in order to account for the fluctuations, spuriousness and seasonality in yearly FDI flows for NUTS3 regions (Albulescu and Goyeau 2019). In addition, as mentioned in the previous Section, this also allow us to capture the compounding effects of different projects, which are likely interacting and accumulating over time, impacting the regional innovation system at different speeds. In this sense, this also help us reducing the measurement noise from fDiMarkets, projects are registered in the database when announced: each project might take different time to be completed and exert an effect on local knowledge recombination.

While our focus is on on greenfield FDIs as drivers of recombinant novelty, the distinction between greenfield FDIs and M&As warrants further discussion. Greenfield FDIs are typically associated with the introduction of entirely new operations, which may offer opportunities for knowledge creation and recombination from the ground up. However, M&As could also play a significant role in facilitating knowledge spillovers (Javorcik 2004; Valacchi et al. 2021; Damioli and Marin 2024). The choice to prioritize greenfield FDIs in this analysis stems from their data availability and clarity as direct indicators of external knowledge flows. Nonetheless, future research could explore the comparative relationship of greenfield and M&A investments with innovation quality, such as technological novelty (Mewes 2019; Berkes and Gaetani 2021).

**Table 2.1:** Variables' description

| Variable | Description | Source |
|---|---|---|
| Novelty | Count of patents with (locally) novel recombinations of CPC4 technological codes | REGPAT (Spring 2022) |
| IFDI | Count of flows of inward FDIs into focal region, cumulative stocks. | FDIMarkets (FT Intelligence) |
| OFDI | Count of flows of outward FDIs out of focal region, cumulative stocks. | FDIMarkets (FT Intelligence) |
| PopDensity | Population Density (persons per square kilometer) | ARDECO |
| GDPpercapita | Gross Domestic Product per capita, in euros per person | JRC Urban Data Platform |
| Reldens In | Relatedness Density of the local technological base around inward FDIs | FDIMarkets (FT Intelligence) and REGPAT (Spring 2022) |
| Reldens Out | Relatedness Density of the local technological base around outward FDIs | FDIMarkets (FT Intelligence) and REGPAT (Spring 2022) |
| KnowledgeStock | Knowledge Stock - cumulative stock of patents | REGPAT (Spring 2022) |
| Variety | Ratio between Unrelated and Related Variety (Castaldi et al. 2015) | REGPAT (Spring 2022) |
| RegSpecialization | Regional specialization, calculated as revealed comparative advantage, into employment in manufacturing sectors (B-E), expressed as a dummy variable (above or below 1) | ARDECO |

Our key explanatory variables, therefore, are based on the number of projects flowing into (or out of) European regions (respectively $IFDI$ and $OFDI$). An advantage of this measure for our framework, as opposed to summing up capital expenditures, is that counts better able to capture the number of distinct linkages established by regions, rather than capturing the size of the investments. While these are the preferred explanatory variables, relying on counts of projects does not allow for depreciation or disinvestments in the capital stock of the region. We ensure that our results are not influenced by this construction, by running a battery of robustness checks with

FDI variables constructed in different ways[1], confirming our main estimates.

Next, in order to investigate the role of distance between FDIs and the knowledge base of home regions, and test hypotheses 3 and 4, we build a measure of relatedness density (Hidalgo et al. 2018) around FDIs. This measure reflects the average share of local technologies that are related to the cumulative number of FDI projects in that region. We work out two variables for inward and outward flows (variable names: $ReldensIn$, $ReldensOut$). Given an investment entering or leaving a region, our relatedness measures take into consideration the share of technologies within that region which are related to that investment. In order to compare the technological content of FDIs to the local technological capabilities, we employ crosswalks between sectors and technologies provided by Lybbert and Zolas (2014b), and map the sector of each FDI to 4-digit technological codes. In turn, we build a symmetric proximity matrix between technologies, in which we rely on all the patents and CPC codes in REGPAT to calculate pairwise proximities. Based on all the patent applications in the database, we calculate the minimum conditional probability of co-occurrences:

$$\phi_{i,j} = min\{P(RCAx_i|RCAx_j), P(RCAx_j|RCAx_i)\} \tag{2.1}$$

$\phi_{i,j}$ is a proximity matrix of technology-technology occurrences, and symmetrically defines the technological-space Hidalgo et al. (2018). We define

[1]Robustness checks include discounted stocks of capital expenditures and stocks of inflows net of outflows

the matrix $\psi_{i,r,t}$ for technology $i$ in region $r$ at year $t$ as proposed in Boschma, Heimeriks et al. (2014), obtaining the relatedness measures for each region-technology pairs:

$$\psi_{i,r,t} = \frac{\sum_{j \in r, j \neq i} \phi_{i,j}}{\sum_{j \neq i} \phi_{i,j}} \tag{2.2}$$

$\psi_{i,r,t}$ represents the relatedness between each technology and the rest of the technologies in the region (i.e. its knowledge base). In turn, we aggregate $\psi_{i,r,t}$ as to the region-year level, in order to capture the average relatedness-density around FDIs. In Equation 2.3 we calculate $\omega_{i,r,t}$ as the number of FDIs in technological category $i$, for that region and year. Finally, we calculate the relatedness-density vector as the average relatedness density around the FDIs. For region $r$ at time $t$ this will be expressed as:

$$Reldens_{r,t} = \frac{\sum_{j \neq i} \omega_{i,r,t} \psi_{i,r,t}}{\sum_{i \neq j} \omega_{i,r,t}} \tag{2.3}$$

The count is based on the count of (cumulative) FDI projects, consistently with our main regressors for both $IFDI$ and $OFDI$, and captures the evolution over time of the relatedness density around FDIs. In summary, the region-year vector that we obtain represents the average relatedness that the technological basis of a given region has with respect to inward and outward FDIs. Therefore, it ranges between 0 and 1, where the zeros might also be due to the absence of FDIs. As mentioned in Section 2.2.4, this measure is more intuitive for the case of $IFDI$, as it measures the average share of technologies related to the technologies of the incoming FDIs. In the case of $OFDI$, instead, the measure reflects how related the knowledge base in the home region is to the knowledge of the OFDI moving outside of it. In this sense, we rely on the diversity of FDIs in interacting with regional knowledge bases, without making assumptions about the composition of the structure of the foreign knowledge-base. This measure reflects the distance between FDIs and regions, as opposed to the bilateral distance between region-region knowledge bases. By adopting a similar approach to relatedness-density and worldwide regionalized patents, future work could explore the (bilateral) technological distance between host and home regions, connected through FDIs, in spurring recombinant novelty.

Additionally to recombinant novelty and FDI variables, we also build a set of controls. First, we create a variable for the total stock of patents in the region (variable name: $KnowledgeStock$), calculated using the inventory rule with no discount rates.[2] This allows us to control for the total amount of patents, with respect to the novel-only patents as well as for the size-effect of the innovation system. Next, we build a control for population density ($PopDensity$, population per square kilometer), capturing the effect of agglomeration economies. Furthermore, we control for GDP per capita in order to account for the level of development of the region ($GDPpercapita$). We also create a variable for the composition of knowledge variety at the local level. We follow previous work (Frenken et al. 2007; Castaldi, Frenken et al. 2015) in order to build entropy-based measures of unrelated variety. In

[2] In robustness checks, we test different discount rates, at 10% and 20%.

our models explaining recombinant novel patents, it controls for the size of technological classes that are distant enough from each other and that are available in the local innovation system, affecting the probability to generate recombiant novelty. Regions might combine to different degrees both related and unrelated variety (Castaldi, Frenken et al. 2015). We create an indicator reflecting the relative importance of unrelated variety (UV) over related variety (RV) expressed as the ratio of UV over RV. Finally, in order to control for the regional industrial base composition ($RegSpecialization$), we build a control for specialization in manufacturing employment in the region. We calculate the revealed comparative advantage, using employment data, from ARDECO database, for the NUTS3 regions in manufacturing sectors (B-E). We dichotomize this indicator as a dummy variable, capturing regions above or below median values for specialization in manufacturing. In Table 2.1 we provide a summary of the main variables employed in the analysis, while in Table 3.2 we provide the descriptive statistics for our final dataset. In Figure 2.1 we show the geographical distribution of $Novelty$, $IFDI$ and $OFDI$.

### 2.3.2 Methodology

In order to test our hypotheses, we estimate the following baseline model, regressing the recombinant novel patents on FDI variables and a set of controls:

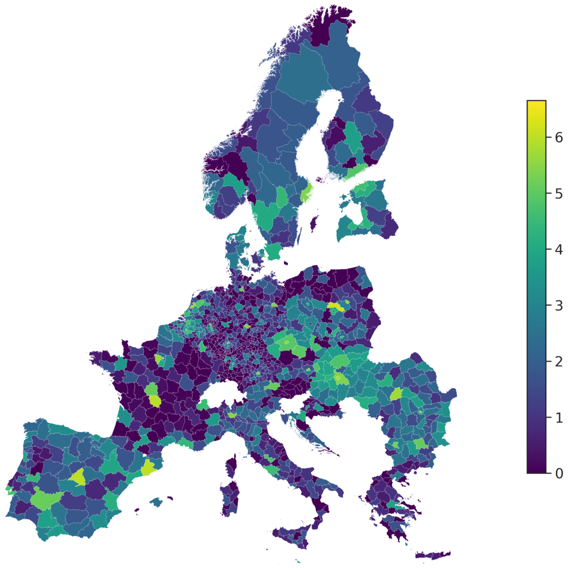
$$y_{i,t} = \alpha + \beta_1 FDI_{i,t-5} + \mathbf{X}_{i,t-5}\boldsymbol{\beta} + \gamma_i + \theta_t + \epsilon_{i,t} \tag{2.4}$$

where:

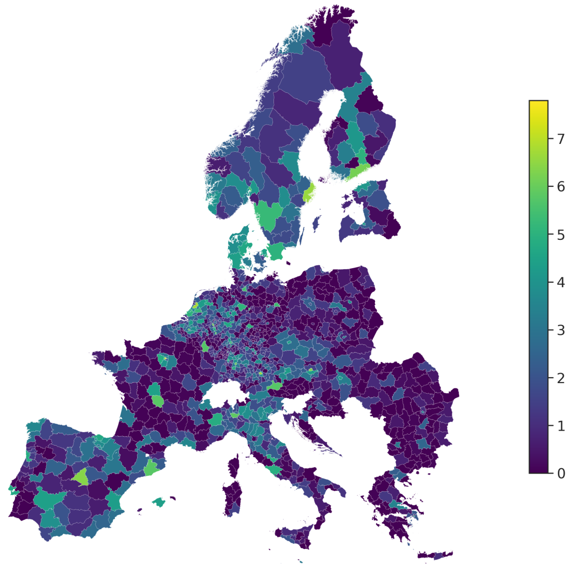
- $\gamma_i$ are time-fixed effects;
- $\theta_t$ are region-fixed effects;
- $\mathbf{X}$ is a vector of control variables;
- $\epsilon_{i,t}$ is the idiosyncratic error term

Our estimation strategy relies on a set of fixed effects in order to control for region and time-specific factors affecting the generation of recombinant novelty, in which we test the battery of FDI variables constructed in the previous subsection. Standard errors are clustered the NUTS3 level. Given that both patents-based indicators FDIs variables suffer from volatility, especially at a highly granular level like NUTS3 regions. In order to account for the high volatility of patents at such a fine-grained geographical level, we run our main estimates by employing a 4-years moving average of both dependent and independent variables. In robustness checks, we test different windows for moving averages.[3] All the variables are constructed at the

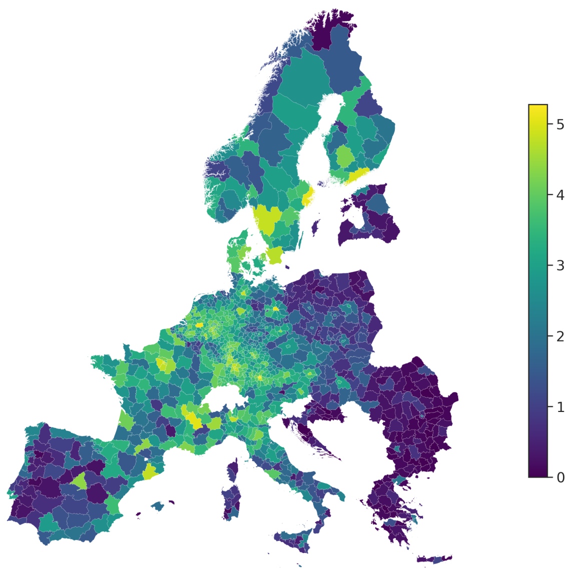
[3] As robustness, we also run models with an aggregation in rolling-sums of novel patents, avoiding the use of moving averages on dependent and independent variables, finding consistent results.




**(a)** Stock of IFDIs 2003-2017 (log scale)




**(b)** Stock of OFDIs 2003-2017 (log scale)




**(c)** Recombinant Novel patents (log scale)

**Figure 2.1:** IFDIs, OFDIs and local recombinant novelty in NUTS3 regions (2003-2017)

NUTS3 level over the period 2003-2017, and enter the estimation in log form.

Importantly, all the variables on the right-hand side are lagged by five years. The timing of effects when considering the nexus between FDI and innovation might be crucial, leading us to choose this specification for three reasons. First, while greenfield FDIs are registered in the fDiMarkets at announcement, it takes time for them to be realized, and start affecting the local innovation system. Second, recombinant innovation arguably takes some time to materialize, as knowledge has to be absorbed and further re-combined. From a theoretical point of view, as detailed in Section 2.2, the effects of FDI projects are also likely to compound in time, and the interaction between external and internal knowledge within the home innovation system is likely to require time to materialize and translate into recombinant novelty, leading us to assume at least a 5-years lag. From an econometric standpoint, lagging independent variables also reduces potential endogeneity problems due to the simultaneity bias and of reverse causality. Regions with a more diverse knowledge pool and characterized by higher levels of local recombinant innovation are also more likely to attract IFDIs, or be home economies for more competitive firms engaging in OFDIs. Despite theoretical and econometric reasons in running our baseline estimates with a 5 years lag structure, in robustness checks we dive deeper into the timing of effects and run several specifications with different lag structures.

## 2.4 Results

### 2.4.1 Main results

We estimate our baseline models by means of two-way fixed effects OLS regressions, with both time and NUTS-3 region fixed effects, in 1096 NUTS3 regions between 2003 and 2017. Following Aghion, Akcigit et al. (2019), we correct for the zero inflation in FDIs by building a dummy variable that turns positive if the number of FDIs is equal to 0, and having the continuous FDI variable turn 0 to avoid logarithms of null values. This dummy captures the possible structural difference between NUTS3 regions lacking FDIs or having them, always significant and not reported.

In Table 2.2 we regress recombinant-novel patents against $IFDI$, $OFDI$ and the set of control variables detailed in the previous Section. Columns 1-5 add stepwise the set of controls for $IFDI$: in line with hypothesis 1, inward FDIs are positively associated with recombinant novelty, suggesting that external knowledge flows to the focal regions. In columns 6-10, instead, we regress recombinant novelty against $OFDI$, and find a negative and significant association, in favor of hypothesis 2b and in contrast with hypothesis 2a. The magnitude of the coefficients suggests a small but significant elasticity, with a 1% increase in the IFDIs being associated with a 0.04% increase in the number of novel patents.

Also the rest of predictors are in line with our expectations. $PopDensity$, capturing the agglomeration dynamics, is always positive and significant. The measure for $Variety$ (the ratio between UV and RV) also seems positively associated with novelty. This reflects the idea that the higher the relative importance of unrelated variety in the regional knowledge composition, the larger the possibilities for novel and unprecedented combinations. Also the control for GDP per capita is positive and significant in all specifications. Finally, our control for regional specialization ($RegSpecialization$) is also positively and significantly associated with local recombinant novelty.

**Table 2.2:** OLS two-way FE

| | Novelty | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| IFDI | 0.0512*** | 0.0506*** | 0.0487*** | 0.0435*** | 0.0434*** | | | | | |
| | (0.0163) | (0.0162) | (0.0162) | (0.0159) | (0.0159) | | | | | |
| OFDI | | | | | | -0.0422*** | -0.0415*** | -0.0395*** | -0.0440*** | -0.0441*** |
| | | | | | | (0.0146) | (0.0145) | (0.0145) | (0.0142) | (0.0142) |
| PopDensity | 0.7052** | 0.6794** | 0.7075** | 1.254*** | 1.262*** | 1.080*** | 1.049*** | 1.060*** | 1.632*** | 1.642*** |
| | (0.2986) | (0.2983) | (0.2981) | (0.3062) | (0.3077) | (0.2997) | (0.2991) | (0.2986) | (0.3085) | (0.3101) |
| KnowledgeStock | | 0.0420 | 0.0466 | 0.0030 | 0.0029 | | 0.0401 | 0.0448 | -0.0012 | -0.0013 |
| | | (0.0328) | (0.0325) | (0.0318) | (0.0318) | | (0.0329) | (0.0326) | (0.0318) | (0.0318) |
| Variety | | | 0.7116** | 0.5026 | 0.5034 | | | 0.6782** | 0.4517 | 0.4525 |
| | | | (0.3163) | (0.3102) | (0.3102) | | | (0.3151) | (0.3073) | (0.3073) |
| GDPpercapita | | | | 0.9357*** | 0.9361*** | | | | 0.9730*** | 0.9735*** |
| | | | | (0.1405) | (0.1406) | | | | (0.1406) | (0.1406) |
| RegSpecialization | | | | | 0.1198*** | | | | | 0.1285*** |
| | | | | | (0.0252) | | | | | (0.0249) |
| Observations | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 |
| $R^2$ | 0.98222 | 0.98225 | 0.98231 | 0.98283 | 0.98283 | 0.98219 | 0.98222 | 0.98228 | 0.98284 | 0.98284 |
| Within $R^2$ | 0.01218 | 0.01403 | 0.01764 | 0.04635 | 0.04643 | 0.01095 | 0.01262 | 0.01588 | 0.04706 | 0.04715 |
| region FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Clustered (region) standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dep var: Number of novel patents in NUTS3 regions. $IFDI$ is the cumulative count of inward FDI projects in the region. $OFDI$ is the cumulative count of number of outward FDI projects from the region. Explanatory variables are lagged by five years. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models are estimated through two-way fixed effects OLS estimators. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the NUTS3 level.

We turn to testing hypotheses 3 and 4a/4b, concerning the distance between external knowledge (respectively of $IFDI$ and $OFDI$) and the local knowledge base, as proxied by the relatedness-density measure constructed in Section 2.3. In Table 2.3 we show the results of the estimations around both $IFDI$ and $OFDI$, in columns 1-5 and 6-10 respectively. In line with the expectations laid out in hypothesis 3, relatedness around $IFDI$ is negatively associated with local recombinant novelty. This indicates that FDIs are likely to feed the generation of novelty at the local level, as long as investments are targeted to activities that are loosely related with the technological domains in which the region is specialized. Foreign-knowledge inputs too closely related to the local technological base, could fail to provide new components through knowledge transfers, necessary to generate novel recombination, and maybe even reinforce existing path-dependence.

By contrast, relatedness-density around OFDIs is negatively and significantly associated with local recombinant novelty in the home region. The negative coefficients are in line with the findings for the competing hypotheses on the role of OFDIs, and seem to support hypothesis 4b, and in contrast with hypothesis 4a. The distance between OFDIs and the specializations in the *home* knowledge base is negatively associated with novelty. The higher the proximity between OFDI firms and the core of the specializations of the focal region, the more negative the association with novelty, indicating that hollowing out by very central firms could bear worse impacts than more peripheral ones. As mentioned in Section 2.3.1, this measure of outward relatedness only concerns the proximity of the home knowledge base around OFDIs, and does not consider the characteristics of the host economy.

**Table 2.3:** OLS two-way FE

| | Novelty | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Reldensin | -0.3899*** | -0.3853*** | -0.3713*** | -0.3394*** | -0.3412*** | | | | | |
| | (0.1012) | (0.1014) | (0.1004) | (0.0998) | (0.0999) | | | | | |
| Reldensout | | | | | | -0.4057*** | -0.3987*** | -0.3846*** | -0.3762*** | -0.3775*** |
| | | | | | | (0.1091) | (0.1087) | (0.1077) | (0.1078) | (0.1079) |
| PopDensity | 0.9408*** | 0.9110*** | 0.9312*** | 1.463*** | 1.473*** | 0.8506*** | 0.8238*** | 0.8466*** | 1.387*** | 1.397*** |
| | (0.2861) | (0.2861) | (0.2859) | (0.2957) | (0.2972) | (0.2873) | (0.2874) | (0.2871) | (0.2973) | (0.2987) |
| KnowledgeStock | | 0.0433 | 0.0479 | 0.0032 | 0.0030 | | 0.0409 | 0.0454 | 0.0002 | 0.0000 |
| | | (0.0330) | (0.0326) | (0.0319) | (0.0319) | | (0.0328) | (0.0324) | (0.0317) | (0.0317) |
| Variety | | | 0.7235** | 0.5066 | 0.5074 | | | 0.6851** | 0.4645 | 0.4653 |
| | | | (0.3165) | (0.3099) | (0.3099) | | | (0.3146) | (0.3074) | (0.3074) |
| GDPpercapita | | | | 0.9528*** | 0.9532*** | | | | 0.9605*** | 0.9609*** |
| | | | | (0.1408) | (0.1408) | | | | (0.1408) | (0.1408) |
| RegSpecialization | | | | | 0.1438*** | | | | | 0.1379*** |
| | | | | | (0.0256) | | | | | (0.0250) |
| Observations | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 |
| $R^2$ | 0.98217 | 0.98221 | 0.98227 | 0.98281 | 0.98282 | 0.98224 | 0.98227 | 0.98233 | 0.98288 | 0.98288 |
| Within $R^2$ | 0.00978 | 0.01175 | 0.01548 | 0.04540 | 0.04551 | 0.01363 | 0.01538 | 0.01872 | 0.04915 | 0.04925 |
| region FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Clustered (region) standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dep var: Number of novel patents in NUTS3 regions. *Reldensin* is the relatedness density of regional knowledge base around inward FDIs in the region. *Reldensout* is the relatedness density of regional knowledge base around outward FDI projects from the region. Explanatory variables are lagged by five years. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models are estimated through two-way fixed effects OLS estimators. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the NUTS3 level.

Overall, our baseline estimates are in line with the idea that $IFDI$ can stimulate recombinant novelty, through knowledge transfers increasing local spillovers, or through competitive effects within the home economy, with this impact increasing with the technological distance between FDIs and the local knowledge base. In contrast, we find evidence suggesting a potential hollowing-out of local resources through $OFDI$, with regions loosing the potential for generating recombinant novelty, loss magnified by the relatedness of OFDIs to the technological structure of the home region.

### 2.4.2 Heterogeneity

We add evidence on the spatial dimension of knowledge spillovers, which can give rise to Jacobian and Marshallian externalities at the local level, spurred by the agglomerations of FDIs (Burger and Meijers 2016) or in neighbourhood regions, contributing to local recombinant search process (Boschma, Martín et al. 2017). Our focus on NUTS3 regions is rooted in the need to understand FDIs' effects with greater nuance, and in its geographical heterogeneity, as knowledge transfers do not happen in a "territorial vacuum", but in highly contextualized local dynamics (Crescenzi and Iammarino 2018). While many studies have been performed either at a national level or for sets of regions of a single EU country, we add evidence at a fine-grained geographical level, exploring the core-periphery structure of regional innovation and FDI patterns within Eastern and Western European regions. CH2/figures 2.1a and 2.1b show the average (log) of greenfield $IFDI$ and $OFDI$. As far as inward FDIs are concerned, we observe that their spatial distribution partly overlaps with the traditional European core-periphery economic geography. The bulk of outward FDIs is mainly originating from the European core regions, which include central-northern regions and their global capital cities.

We explore geographical heterogeneity, breaking down FDIs according to the origin and destination of the flows. We build variables, for both inward and outward FDIs, concerning European regions only ($EUFDI$), or rest-of-the-world only ($ROWFDI$). This latter measure counts IFDIs coming from the rest of the world into European regions ($ROWIFDI$) or OFDIs leaving the continent ($ROWOFDI$). In addition, we build variables distinguishing between EU15 and nonEU15.[4] nonEU15 countries in the sample, and build variables with the same logic.[5]

Given the relevance of sectoral, functional, and value-chain aspects in the study of FDIs and innovation (for discussions, see Iammarino (2018) and Ascani, Bettarelli et al. (2020)), we also consider the knowledge-intensity of FDIs as a driver of recombinant novelty. We exploit the functional classification provided by fDi Markets, allowing us to distinguish knowledge-intensive projects involving R&D efforts, which as show in the previous section might

[4]While naming the variable EU15, we exclude the UK from our sample, counting into this group: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden and Norway

[5]The countries included are Bulgaria, Croatia, Cyprus, Czech Republic, Estonia, Hungary, Latvia, Liechtenstein, Malta, Poland, Romania, Slovakia. We exclude from the analysis, due to the lack of time-consistent regional data for control variables: Lithuania, Slovenia, Montenegro, Iceland, North Macedonia and Bosnia-Herzegovina.

be a source of heterogeneity in their relationship with recombinant novelty, for both inward and outward flows. Finally, we test a complementary measure for R&D FDIs, counting only FDIs which are not classified as R&D projects. As done for the previous measures, we build several variables for nonR&D FDIs according to their geographical variation (EU and ROW; EU15 and nonEU15). We provide complete list and description of the set of FDI variables that we are testing in Appendix Table A.1.

In Table 2.4 In columns 1-6, we focus on the same sample of European NUTS3 regions, and show first the same coefficients for FDIs coming from anywhere (within and outside the EU), in the first two columns. In turn, we break down FDIs by origin and destination, as those coming from (or going to) European countries ($EUIFDI$, $EUOFDI$, columns 3 and 4), and those coming from or going to the rest of the world ($ROWIFDI$, $ROWOFDI$, columns 5 and 6). The positive association for IFDIS remains positive and significant for within-EU IFDIs, while only significant at the 10% confidence for FDIs coming from outside-Europe regions ($ROWIFDI$). OFDIs, instead, seem to be more significantly associated to novelty when investments are flowing outside of the EU, with coefficients remaining negative in both cases.

Zooming into European flows, we further subset the analysis to European flows directed to (or coming from) EU15 vis-à-vis non-EU15 countries. In columns 7-10, we present estimates splitting the sample for EU15 only countries, and testing separately investments in (from) other EU15 countries or non-EU15 countries. The effects seem to be driven by the regions within EU15 countries, again in the line of hypothesis 1 and 2b, with a positive coefficient for $EU15IFDI$ and negative for $EU15OFDI$. In contrast, when considering only non-EU15 regions as a sample for the regressions (columns 11-14) the effects are insignificant, except for the positive coefficient, at the 10% significance level and with a larger magnitude, for investments coming from EU15 regions (column 11). Interestingly, the negative association with OFDIs, in this case, is always insignificant, suggesting that OFDIs at a lower level of technological capacities in non-EU15 regions, might not have negative effects on recombinant novelty. Additionally, we speculate that these results might be speaking to different dynamics in terms of competitive effects. In EU15 countries, outward flows from firms could result in a loss of knowledge elements to recombine, the opposite is true for non-EU15 regions. While the former have longer been exposed to the Single Market, the latter might not be loosing knowledge elements due to the hollowing-out, when MNEs gain access to EU15 knowledge pools.

**Table 2.4:** OLS two-way FE - FDI variables

| | All EU | | | | | | EU15 | | | | nonEU15 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| IFDI | 0.0434$^{***}$ | | | | | | | | | | | | | |
| | (0.0159) | | | | | | | | | | | | | |
| OFDI | | -0.0441$^{***}$ | | | | | | | | | | | | |
| | | (0.0142) | | | | | | | | | | | | |
| EU IFDI | | | 0.0583$^{***}$ | | | | | | | | | | | |
| | | | (0.0165) | | | | | | | | | | | |
| EU OFDI | | | | -0.0577$^{***}$ | | | | | | | | | | |
| | | | | (0.0146) | | | | | | | | | | |
| ROW IFDI | | | | | 0.0281$^{*}$ | | | | | | | | | |
| | | | | | (0.0158) | | | | | | | | | |
| ROW OFDI | | | | | | -0.0353$^{**}$ | | | | | | | | |
| | | | | | | (0.0176) | | | | | | | | |
| EU15 IFDI | | | | | | | 0.0365$^{**}$ | | | | 0.0921$^{*}$ | | | |
| | | | | | | | (0.0159) | | | | (0.0551) | | | |
| EU15 OFDI | | | | | | | | -0.0306$^{**}$ | | | | 0.0391 | | |
| | | | | | | | | (0.0134) | | | | (0.0587) | | |
| nonEU15 IFDI | | | | | | | | | 0.0266 | | | | 0.0396 | |
| | | | | | | | | | (0.0170) | | | | (0.0611) | |
| nonEU15 OFDI | | | | | | | | | | -0.0144 | | | | -0.0109 |
| | | | | | | | | | | (0.0154) | | | | (0.0639) |
| Full controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 6,426 | 6,426 | 6,426 | 6,426 | 1,526 | 1,526 | 1,526 | 1,526 |
| $R^2$ | 0.98283 | 0.98284 | 0.98286 | 0.98286 | 0.98277 | 0.98276 | 0.98402 | 0.98401 | 0.98401 | 0.98397 | 0.95094 | 0.95048 | 0.95072 | 0.95042 |
| Within $R^2$ | 0.04643 | 0.04715 | 0.04772 | 0.04779 | 0.04326 | 0.04217 | 0.01800 | 0.01728 | 0.01702 | 0.01507 | 0.08773 | 0.07912 | 0.08360 | 0.07798 |
| region FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Clustered (region) standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dep var: Number of novel patents in NUTS3 regions. $IFDI$ is the cumulative count of inward FDI projects in the region. $OFDI$ is the cumulative count of number of outward FDI projects from the region. The geographical breakdown of variables is computed as the flows, for the focal NUTS3 regions, only coming from (or going to) European countries ($EUFDI$), non-European countries ($ROWFDI$) or to respectively EU15 and nonEU15 countries, as detailed in section 2.3. Explanatory variables are lagged by five years. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models are estimated through two-way fixed effects OLS estimators. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the NUTS3 level.

The estimates presented so far considered any type of FDI investment, without distinguishing their nature, particularly in terms of knowledge-intensity. As mentioned above, the knowledge intensity of these projects might be covering for additional heterogeneity, given the different nature of R&D investments, which might exert a more positive effect especially for OFDIs, in contrast with the offshoring of production activities. Hence, we exploit information from fDiMarkets, and only consider those in R&D activities. In Table 2.5, we present similar estimates as in the previous Table, but focusing on investments concerning new R&D facilities. In column 1, the inward coefficient for total projects is not significant, while it remains negative and significant for outward total flows, in column 2. Once again, these effect seem to be driven by within-EU flows, while $ROWIFDI$ and $ROWOFDI$ are not significant. Looking at the within EU15 dynamics, the coefficient for $EU15IFDI$ (IFDIs flowing into EU15 regions, from EU15 regions - column 7) remains highly significant and positive, with a coefficient larger in magnitude. In comparison with total investments presented in the previous Table, $OFDI$ are never significant when considering within-EU FDIs in R&D. In comparison with total investments, knowledge-intensive OFDIs are not showing potential for a hollowing-out of local knowledge resources. While positive in aggregate, the breakdown does not show negative effects of OFDIs.

**Table 2.5:** OLS two-way FE - FDI variables for Research and Development Projects

| | | | All EU | | | | | EU15 | | | | | nonEU15 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| IFDI | 0.0244 | | | | | | | | | | | | | |
| | (0.0249) | | | | | | | | | | | | | |
| OFDI | | -0.0410** | | | | | | | | | | | | |
| | | (0.0172) | | | | | | | | | | | | |
| EU IFDI | | | 0.1108*** | | | | | | | | | | | |
| | | | (0.0302) | | | | | | | | | | | |
| EU OFDI | | | | -0.0683*** | | | | | | | | | | |
| | | | | (0.0202) | | | | | | | | | | |
| ROW IFDI | | | | | -0.0292 | | | | | | | | | |
| | | | | | (0.0254) | | | | | | | | | |
| ROW OFDI | | | | | | 0.0239 | | | | | | | | |
| | | | | | | (0.0381) | | | | | | | | |
| EU15 IFDI | | | | | | | 0.1136*** | | | | 0.0565 | | | |
| | | | | | | | (0.0328) | | | | (0.0614) | | | |
| EU15 OFDI | | | | | | | | 0.0039 | | | | -0.3717 | | |
| | | | | | | | | (0.0220) | | | | (0.3451) | | |
| nonEU15 IFDI | | | | | | | | | 0.0222 | | | | -0.2398* | |
| | | | | | | | | | (0.0466) | | | | (0.1307) | |
| nonEU15 OFDI | | | | | | | | | | 0.0011 | | | | 0.3025 |
| | | | | | | | | | | (0.0285) | | | | (0.6071) |
| Full controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 6,426 | 6,426 | 6,426 | 6,426 | 1,526 | 1,526 | 1,526 | 1,526 |
| $R^2$ | 0.98276 | 0.98276 | 0.98285 | 0.98277 | 0.98276 | 0.98275 | 0.98407 | 0.98397 | 0.98397 | 0.98397 | 0.95049 | 0.95046 | 0.95054 | 0.95043 |
| Within $R^2$ | 0.04243 | 0.04267 | 0.04768 | 0.04324 | 0.04217 | 0.04164 | 0.02070 | 0.01470 | 0.01478 | 0.01468 | 0.07920 | 0.07866 | 0.08016 | 0.07820 |
| region FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Clustered (region) standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dep var: Number of novel patents in NUTS3 regions. $IFDI$ is the cumulative count of inward FDI projects in the region. $OFDI$ is the cumulative count of number of outward FDI projects from the region. The geographical breakdown of variables is computed as the flows, for the focal NUTS3 regions, only coming from (or going to) European countries ($EUFDI$), non-European countries ($ROWFDI$) or to respectively EU15 and nonEU15 countries, as detailed in section 2.3. All FDI variables are only counted if labeled as R&D projects in the FDIMarkets database. Explanatory variables are lagged by five years. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models are estimated through two-way fixed effects OLS estimators. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the NUTS3 level.

Three possible factors might be at play in explaining the heterogeneous results for sub-samples of regions within the EU. First, the numerosity of the projects at such a fine-grained geographical level could represent an issue for in estimating this relationship, especially when considering the non-EU15 sample. Second, the negative association of total OFDIs with local novelty could be driven by the offshoring of resources in non-knowledge intensive projects, while MNEs internationalising in R&D project are not depleting regions of knowledge resources. Third, spillover effects at such a fine-grained geographical level could help explain the direct and indirect effects of FDIs on recombinant novelty. In order to understand further the functional aspects of FDIs, we run estimates only considering FDIs in non-R&D projects. In Table 2.6, we present symmetric estimates excluding R&D projects from the construction of FDI variables. The main results for global, overseas and European investments are confirmed as positive for inward FDIs. Interestingly, OFDIs directed outside of Europe (column 6) are negative and significant, suggesting that the hollowing-out hypothesis is indeed subject to a large degree of heterogeneity in terms of the knowledge-intensity of the investments, with R&D OFDIs going overseas not necessarily harming local recombinant innovation, in comparison with non-R&D flows.

**Table 2.6:** OLS two-way FE - FDI variables excluding Research and Development Projects

| | All EU | | | | | | EU15 | | | | nonEU15 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| IFDI | 0.0474*** | | | | | | | | | | | | | |
| | (0.0160) | | | | | | | | | | | | | |
| OFDI | | -0.0453*** | | | | | | | | | | | | |
| | | (0.0142) | | | | | | | | | | | | |
| EU IFDI | | | 0.0611*** | | | | | | | | | | | |
| | | | (0.0164) | | | | | | | | | | | |
| EU OFDI | | | | -0.0565*** | | | | | | | | | | |
| | | | | (0.0146) | | | | | | | | | | |
| ROW IFDI | | | | | 0.0357* | | | | | | | | | |
| | | | | | (0.0182) | | | | | | | | | |
| ROW OFDI | | | | | | -0.0365** | | | | | | | | |
| | | | | | | (0.0176) | | | | | | | | |
| EU15 IFDI | | | | | | | 0.0388** | | | | 0.0931* | | | |
| | | | | | | | (0.0157) | | | | (0.0550) | | | |
| EU15 OFDI | | | | | | | | -0.0101 | | | | 0.0496 | | |
| | | | | | | | | (0.0121) | | | | (0.0783) | | |
| nonEU15 IFDI | | | | | | | | | 0.0300* | | | | 0.0406 | |
| | | | | | | | | | (0.0174) | | | | (0.0619) | |
| nonEU15 OFDI | | | | | | | | | | -0.0137 | | | | -0.0106 |
| | | | | | | | | | | (0.0154) | | | | (0.0628) |
| Full controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 6,426 | 6,426 | 6,426 | 6,426 | 1,526 | 1,526 | 1,526 | 1,526 |
| $R^2$ | 0.98284 | 0.98284 | 0.98287 | 0.98285 | 0.98278 | 0.98276 | 0.98403 | 0.98399 | 0.98401 | 0.98397 | 0.95091 | 0.95047 | 0.95071 | 0.95042 |
| Within $R^2$ | 0.04700 | 0.04700 | 0.04832 | 0.04723 | 0.04362 | 0.04217 | 0.01843 | 0.01584 | 0.01706 | 0.01508 | 0.08715 | 0.07884 | 0.08343 | 0.07798 |
| region FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Clustered (region) standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dep var: Number of novel patents in NUTS3 regions. $IFDI$ is the cumulative count of inward FDI projects in the region. $OFDI$ is the cumulative count of number of outward FDI projects from the region. The geographical breakdown of variables is computed as the flows, for the focal NUTS3 regions, only coming from (or going to) European countries ($EUFDI$), non-European countries ($ROWFDI$) or to respectively EU15 and nonEU15 countries, as detailed in section 2.3. All FDI variables are only counted if not labeled as R&D projects in the FDIMarkets database. Explanatory variables are lagged by five years. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models are estimated through two-way fixed effects OLS estimators. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the NUTS3 level.

However, the coefficient for total $EUOFDI$ remains significantly negative (column 4), also in this case, again suggesting a potential competition effect within Europe. Interestingly, this coefficient is negative and significant for the overall sample, but not significant when breaking down the origin and destination of flows between EU15 and non-EU15 regions. The underlying mechanisms explaining these geographical differences, therefore, could be rooted in the core-periphery structure of both recombinant innovation and the geography of FDIs, adding to the functional aspects highlighted here. If spatial dependence is a factor at play, one could expect strong spatial spillover effects to be driving different effects between focal and neighbourhood regions across the EU.

### 2.4.3 Robustness checks and spatial analysis

We run several additional robustness checks to ensure the consistency of our baseline estimates, in particular for the construction of our FDI variables. In Table A.2 of the Appendix we run several specifications testing different measures and aggregations of the FDI and novelty variables, confirming our main results. In the first two columns, we run our baseline model, employing a measure for the net FDIs in the region, subtracting outflows from inflows. In columns 2-8, we run different specifications based on Capital Expenditures: first net Capex stocks, Capex stock depreciated at different rates, and the same for Capex stock normalized by unit of GDP. In column 9, we run a model without moving averages, aggregating novel patents as a rolling sum of 4 years. In column 10, we run our baseline model considering the number of patents per capita as a dependent variable. In Table A.3 of the Appendix, we run the same checks as the ones presented in the previous Table, except for the measures of net FDIs. Again, we confirm the baseline estimates and negative association between outward FDIs and recombinant novelty.

Furthermore, while a lag of 5-years is reasonable in our setting, we present results at different lags (1-8), for each column of Table A.4 of the Appendix. Inward FDIs seem to be positively associated to novelty, with significant effects starting to materialize after four years, in line with the idea that the impacts of foreign knowledge flows compounds over time. In Table A.5 of the Appendix, we run the same test for outward FDIs, finding similar results, although for a shorter time horizon.

We turn to addressing the issue of spatial dependence. We employ Spatial Durbin models (SDM), adding spatial lags of both the dependent and independent variables. We detail the methodology employed for running SDM models in the Appendix. In Appendix Table A.7, we present the marginal effects, broken down between direct and indirect effects, for both $IFDI$ and $OFDI$, again considering the full sample of NUTS3 regions (panel A), and splitting the sample in EU15 and non-EU15 (panel B). In addition to direct, indirect and total effects, we report also the coefficients for the first-order spatial autoregressive term, which is positive and significant across all specifications.

Once we account for spatial dependence of both FDIs and novelty, very interesting dynamics emerge. Indirect effects are negative and significant across our specifications and samples, revealing that spatial spillovers are an important factor at play. This result is indicative of the strong spatial concentration of FDIs, especially at the NUTS3 level. Panel B uncovers interesting spatial dynamics: the direct effects of $IFDIs$ are positive and significant, and for $OFDI$ are only negative when flowing to other EU15 countries, suggesting that the hollowing out of resources towards similarly developed countries, while insignificant in the case of nonEU15 countries. A very different picture emerges for the nonEU15 sample: direct and indirect effects, of both $IFDI$ and $OFDI$ are always positive and significant, suggesting a positive role for internationalization, where regions are less exposed to internationalization. The negative and significant coefficients for spatial spillovers, instead, suggest a staple importance for the agglomeration dynamics brought by increased connectivity, resulting in negative spillovers for the neighbourhood regions and in between-regions competition in benefiting from international investments (be it through forward or backward linkages).

Again, these results concern total FDIs, without distinguishing in their knowledge-intensity: in order to shed further light on the knowledge-spillover mechanisms, we repeat this exercises for both R&D FDIs and non-R&D FDIs. One would expect that once spatial spillovers are controlled for, the direct effects of knowledge-intensive FDIs could be more positively associated with recombinant novelty. In Table A.8, we show evidence in this direction.

Panel A shows how, for R&D intensive FDIs, direct effects are positive and significant, and the agglomeration effects are mostly responsible for negative total effects through indirect geographical spillovers. In the case of IFDIs in R&D projects, the geographical spillovers (indirect effects) are even turning positive when these flows stay within the EU. In Panel B a similar dynamic is present, with EU15 countries being able to better profit from different types of connectivity (both outside and inside EU15) through direct effects. Across all specifications, OFDIs' direct effects are never negative and significant. In particular, the direct effects seem to be more relevant for EU15 regions in the case on knowledge-intensive FDIs, with focal regions being more able to rip the benefits of both inward and outward FDIs. Interestingly, negative spillovers seem more relevant in the case of IFDIs, while insignificant for outward FDIs. For non-EU15 regions, the positive direct effects of IFDIs are driven, expectedly, by flows coming from Western Europe.

In Appendix Table A.9 we show instead the results for non-R&D projects. Generally, we confirm that the negative relationship with recombinant novelty is driven by the indirect effects of connectivity on neighbourhood regions, but in this case also direct effects can be negative (for example, in within-EU OFDIs). Comparing these estimates with those in the previous Table, we speculate that knowledge-seeking type of investments, even when flowing out of the regions, are not necessarily negative (or even positive) in terms of local novelty. Interestingly, in this case, both IFDIs and OFDIs in Eastern Europe are significantly positive. The evidence seems to suggest the presence of two-way knowledge transfers, and technological spillovers for R&D projects, while the hollowing-out effects can be at play in the case of non-R&D projects, and in particular in regions characterized by a higher degree of international exposure. Evidence on the knowledge-intensity of FDIs calls for careful policy considerations, in particular in terms of OFDIs, showing that the effects of internationalization on local novelty can largely differ depending on their technological content, as well as the specific regional context.

Importantly, our results on direct and indirect effects also speak to the recent literature on income agglomeration through FDIs in the US (Bathelt and Buchholz 2019; Buchholz et al. 2020; Bathelt, Buchholz and J. A. Cantwell 2023). In particular in terms of OFDIs, the benefits of external connectivity seem to be spurring recombinant novelty in the home regions, with important negative indirect effects. The latter suggest a strong competition for the external resources brought by cross-border investments, and are very evident at the provincial level. While speculative, this could indicate that both IFDI and OFDI can increase the concentration of novelty, substantiating the hypothesis of increasing returns to knowledge enhanced by international connectivity, both in terms of IFDIs and OFDIs (Buchholz et al. 2020). The negative geographical spillovers, in fact, could be driven by the labour-market effects described in Section 2.2, and contributing to increase inter-regional inequality through an increased spatial concentration of knowledge spillovers, not only because of IFDIs but also through backward linkages and OFDIs. In this sense, comparing Tables A.8 and A.9 underscores the functional differences in this sense, showing that generally the negative direct effects of connectivity, particularly in terms of OFDIs is moderated by the knowledge-intensity of the investments.

In summary, our estimates suggest a clearly positive role of inward FDIs for local recombinant novelty, with underlying sources of geographic and technological heterogeneity. Strong spatial spillovers are underlying the variation in both IFDIs' and OFDIs' relationship with recombinant novelty, suggesting potentially positive direct effects offset by negative indirect ones through neighbourhood regions.

## 2.5 Conclusions

In this paper, we have investigated the role of greenfield FDIs for regions' capacity to introduce technological novelty. Our analysis relies on a conceptual framework blending different streams of literature, i.e. the recombinant knowledge approach, evolutionary economic geography ad regional connectivity theory. In particular, the theoretical and empirical debate on regional technological novelty is grounded on the Schumpeterian tenet according to which innovation stems from the agents' capacity to combine knowledge and ideas in new ways. The concept of regional recombinant capabilities extends to the regional domain the appreciation of individuals' ability to manage novel recombination through recombinant reuse or creation dynamics (Carnabuci and Operti 2013a; Orsatti, Quatraro and Scandura 2023). In this direction, novelty emerges as the outcome of recombinant creation dynamics, according to which innovating agents in local contexts introduce unprecedented combinations of knowledge and ideas. A crucial factor affecting these dynamics is the composition of local knowledge pools in terms of heterogeneity.

Regional connectivity (Iammarino 2018), i.e. the capacity of regional actors to establish connections with the rest of the world to ensure flows of assets, knowledge and capabilities, represents an additional, and yet very relevant, source of variety that is likely to affect regional recombinant capabilities, and in particular increasing the prospects for the introduction of novel combinations. Consequently, we have put forth four main hypotheses. The first one suggests a positive connection between inward greenfield foreign direct investments (FDIs) and the emergence of new combinations of knowledge in European Union (EU) regions. As widely recognized in the FDI literature, these investments tend to create pro-competitive effects and technological spillovers in the receiving region. The second hypothesis articulates an alternative explanation regarding the potential impact of outward FDIs on the economy of the home region. This competing explanation suggests that both positive and negative outcomes are likely and indeed supported by empirical studies. Therefore, we examine both arguments and propose two nested hypotheses: one suggests a positive relationship between outward greenfield FDIs and the novelty of EU regions, while the other suggests a negative relationship. This approach allows us to investigate how outward FDIs contribute to reconfiguring the local industrial structure in the home region. Finally, the third and fourth hypothesis builds on the argument that the effect of multinational enterprises' (MNEs) investments may depend on the extent to which these investments concentrate in similar or different activities from those of local firms. Based on this, we propose a negative association between the proximity of inward FDIs, and symmetrically two competing explanations, positive and negative, for the proximity of the local knowledge base to outward FDIs leaving the region.

The analysis has focused on the generation of fpatents combining at least one pair of technological classes that have never been combined (Verhoeven et al. 2016). We have focused on a panel of 1096 NUTS3 European regions observed between 2003 and 2017, and implemented panel data and spatial econometrics models to test the relationship between FDIs and regional novelty. In particular, in line with the extant literature, we find that IFDIs are positively and significantly associated with the innovation efforts of regions, suggesting the presence of forward linkages fostering the recombination of distinct knowledge components. While additional micro evidence at the firm level may shed light on the underlying mechanisms, our results support the hypothesis that MNEs' cross-border investments activate knowledge flows connecting heterogeneous and distant knowledge bases.

The evidence on the role of OFDI is instead less straightforward, yet unveiling interesting patterns. In particular, we find an overall negative relationship between outward FDI investments and local recombinant novelty, which would support the hypothesis of a hollowing-out of local resources and knowledge in favour of foreign regions. However, the negative coefficients are also largely driven by local negative spillovers, and direct effects are either positive or insignificant when considering investments in R&D activities. This suggests that knowledge-seeking connectivity might be neutral or even beneficial for the sender region. Relatedness of both inward and outward FDIs shows a negative association with local recombinant knowledge, in line with the evolutionary tenet stressing the importance of variety, and in particular of unrelated variety, for the change in the structure of local economic and technological activities.

Finally, spatial analysis highlights that IFDIs are generally postively associated to local recombinant novelty, especially in knowledge-intensive sectors, with spatial spillovers often being negative, indicating competition between regions and increasing territorial concentration of knowledge returns. Evidence for OFDIs points to a similar direction. While potentially hollowing-out local resources, these also largely depend on negative spatial spillovers, relevant given the strong concentration of OFDIs. Moreover, knowledge-intensive FDIs tend to foster more favorable direct (and even indirect) effects on recombination in neighbourhood regions, whereas non-R&D FDIs are more likely to produce negative impacts through spillovers, showing how different types of connectivity greatly matter in understanding FDI-innovation nexus. Overall, we find that spatial dynamics play a staple role, with both IFDIs and OFDIs potentially contributing on the one hand to the positive acceleration of recombinant capabilities, but on the other to inter-regional competition and inequality.

Our analysis brings relevant policy implications. Previous literature has indeed shown that the capacity to generate technological novelty affects not only the economic performances of firms and regions but also the capacity to address important social challenges like climate change through eco-innovation. Understanding the drivers of technological novelty can hence provide useful indications to policymakers concerning measures possibly fostering path-breaking research and innovation strategies. In this direction, we are line with literature stressing the need to insert regional connectivity into the toolbox of regional development policy (Iammarino 2018). In particular, our results not only confirm the importance of incoming FDIs, but also point to the importance of FDIs outflows and to the need to design local and national strategies in this respect carefully. Policy should consider the motives and geographical aspects of different types of connectivity in designing policies highly embedded in a territorial and technological context. Our evidence shows that maximizing knowledge exchanges to generate local novelty requires a nuanced understanding of the mechanisms at play: namely in terms of the geographical and technological aspects of investments, which might increase inequality by increasing returns to knowledge agglomeration

(Iammarino 2018; Bathelt, Buchholz and J. A. Cantwell 2023).

As with any study, this one has some limitations that should be mentioned. The first aspect pertains to utilizing patent data to quantify regional innovation actors' technological efforts. While it is important to acknowledge that not all new technologies are patented and that patenting propensity is unevenly distributed across sectors, it is noteworthy that there exists a general scientific consensus regarding patents being a dependable indicator of the creation of novel technologies at the local level (Acs et al. 2002). In addition, our evidence only concerns greenfield FDIs and does not distinguish between modes of entry. Another limitation to consider is that the empirical framework employed does not enable us to establish definitive causal relationships. However, our findings offer statistically robust and intriguing associations among the primary variables examined.

Yet, our analysis provides insights for further research. For example, exploiting matched firm-level data to develop a finer-grained analysis of firm-level mechanisms underlying these associations would be useful. Future research could provide more in-depth knowledge of the asymmetrical effects of both incoming and outgoing FDIs, depending on the considered foreign regions and their technological structures. In addition, this paper focuses on EU regions: further research should provide evidence concerning recombinant novelty in different geographical locations, and in particular comparing advanced and developing economies. Finally, while our results on the role of spatial spillovers are suggestive, more research is needed in understanding both the local mechanisms driving negative indirect effects, and their potential consequences on the spatial concentration of recombinant novelty, potentially contributing to inter and intra-regional inequality.

# Chapter 3

# Green diversification, global knowledge sourcing and local skill composition: evidence from the US

**Chapter co-authored with Fabrizio Fusillo, Gianluca Orsatti, Francesco Quatraro and Alessandra Scandura**

**Abstract**

This work investigates the role of green foreign direct investments (FDIs) and local skill composition for regional technological diversification in green domains. We conduct the analysis on 287 US Metropolitan Statistical Areas (MSAs) observed from 2003 to 2018. Our results show that MSAs with higher volumes of green FDIs and higher intensity of abstract skills are more likely to diversify in green technological domains. Moreover, we find that the local endowment of abstract and routine skills moderates the impact of green FDIs, activating compensation and reinforcing mechanisms, respectively. The findings of this work provide novel insights for the academic debate on the determinants of green technological diversification and for the design of an effective policy toolbox to sustain the regional green transition.

## 3.1 Introduction

The 2023 Adaptation Gap Report released by the United Nations Environment Programme (UNEP) stresses that the progress of advanced countries on mitigation and adaptation has been slowing down while the dramatic effects of climate change are becoming more and more severe. One of the main reasons behind such a situation lies in the reduction of the resources committed to planning and implementing concrete actions to cope with environmental degradation and favour the green transition (UNEP 2023).

Given this evidence, understanding the factors that improve the environmental performance of human actions and their drivers remains of paramount importance. In this respect, policy and academic debates have long stressed the relevance of investments in innovation and new technologies as a key enabling condition. The concept of eco-innovation has gained momentum in the last decade. It has paved the way for a flourishing stream of literature investigating the economic drivers and the impact of the generation and adoption of innovation and new technologies aiming at reducing the exploitation of natural resources or the environmental damage associated with economic activities (Kemp 2010; Barbieri, Ghisetti et al. 2016).

More recently, the spatial dynamics underlying the generation of green technological knowledge have been stressed by an increasing number of empirical studies in the geography of innovation field, blending the analysis of eco-innovation dynamics with the regional branching approach. Strengthening the capacity to develop new regional comparative advantage in the supply of clean technologies is indeed an inescapable condition to develop smart specialization strategies for the green transition (Montresor and Quatraro 2020; Orsatti, Quatraro and Scandura 2023; Santoalha et al. 2021; Cicerone et al. 2023).

Based on the well-established tenet that green technologies are likely to stem out of recombination and hybridization of heterogeneous knowledge inputs (Zeppini and Bergh 2011; Zeppini 2015; Orsatti, Quatraro and Pezzoni 2020a), extant literature has shown that technological complementarities and recombinant creation capabilities facilitate regional green technological diversification. On the one hand, Montresor and Quatraro (2020) and Cicerone et al. (2023) stress the importance of local specializations in digital technologies or key enabling technologies, given their general purpose nature and horizontal applicability. On the other hand, Orsatti, Quatraro and Scandura (2023) and Quatraro and Scandura (2019) focus on the interplay between local recombinant capabilities and agency, stressing the relevance of the distinctive capacity to generate atypical knowledge combinations and the presence of agents of structural change such as academic inventors.

The present work intends to contribute to this strand of literature by empirically exploring the extent to which environmental-related foreign direct investments (FDIs) and human capital enable or hinder green technological diversification in the local economies of the United States (US). First, we hypothesize that regions' global knowledge sourcing, measured by the local presence of non-local agents – specifically multinational enterprises (MNEs) through their FDIs – provides access to heterogeneous and unrelated knowledge inputs and hence positively impacts entry into new green technological specializations. Second, given the inherent complexity of green technological knowledge and the consequent relevance of recombinant creation capabilities, we posit that the prevalence of exploration-oriented skills in local contexts is associated with a higher probability that regions will develop new green technological specializations. Thirdly, we investigate the interplay between the local skills endowment and external knowledge sourcing. This interplay is rooted in the tension between potential and realised local absorptive capacity. The former requires a higher ability to recognize, acquire, and assimilate useful external knowledge, while the latter requires the ability to transform and apply it effectively. If local skills suit the first requirement, we posit that they likely compensate for external knowledge in raising the probability of entering new green technological specializations; conversely, if local skills suit the second requirement, we contend that they reinforce the role of external knowledge in regional green technological diversification.

We carry out the empirical analysis on a balanced panel of 287 US Metropolitan Statistical Areas (MSAs) observed from 2004-2018. Our results support the hypotheses that both the local presence of inward environmental-related FDIs and the prevalence of exploration-oriented skills are positively associated with higher likelihoods of regional diversification in green technological domains. Moreover, our results suggest that there is more a compensatory than a reinforcing interplay between these two forces, suggesting that, considering the tension between knowledge sourcing and exploitation, skills more related to the former dominate in the sample analyzed, leading to higher potential than realised absorptive capacity.

This work adds to the literature in many respects. Firstly, theoretical reflections within the evolutionary economic geography literature have stressed the need to link agency to regional diversification, especially when the understanding of structural change dynamics is at stake. Yet, efforts in this direction are still underdeveloped. Some contributions have focused on the role of non-local agents like migrant inventors and foreign firms, with no focus on green technological diversification (Boschma, Coenen et al. 2018; Neffke et al. 2018b; Miguelez and Morrison 2023). Other studies have focused on FDIs in EU green specialization patterns (Castellani, Marin et al. 2022b) and on the structure of co-inventorship networks on country-level diversification patterns (Corrocher et al. 2024). Our study provides a step forward in that we articulate the analysis of the role of green inward FDIs by focusing on the region level of analysis, looking at the entry in new green specializations for regions that were not previously specialized in those domains. Further, our focus on the determinants of regional green technological diversification in the US enriches existing empirical evidence that is mainly centered on European countries. Secondly, the inclusion of the occupational structure as a proxy of the skill endowment of the local workforce explicitly brings to the fore the dynamics of know-how and learning that can either enable or thwart the development of a new technological trajectory. While recent studies have investigated the nexus between changes in the local skill structure and the green transition, they have mainly focused on the labor market impact of the greening of the economy in terms of both job creation and destruction and changes in the task content of occupations (Vona, Marin, Consoli and Popp 2018; Consoli, Marin et al. 2016; Vona, Marin and Consoli 2019). Yet, little emphasis has been put so far on the role of local human capital endowment in green technological diversification dynamics (Orsatti, Perruchas et al. 2020).

The remainder of the paper is organized as follows. Section 2 articulates the theoretical framework and puts forth the hypotheses. Section 3 presents the data and the empirical strategy. Section 4 discusses the results of the econometric analysis. Section 5 concludes and discusses the policy implications and the avenues for further research.

## 3.2 Background and hypotheses

### 3.2.1 FDIs and green technological diversification

The flourishing literature that analyses the determinants of eco-innovation underscores that green inventions show greater complexity (in terms of technological scope) than non-green counterparts (Marchi 2012; Ghisetti, Marzucchi et al. 2015) and that they are more likely to emerge out of novel recombination of existing knowledge components (Barbieri, Marzucchi et al. 2020; Orsatti, Quatraro and Pezzoni 2020a). Accordingly, the generation of green technological knowledge is conceived as the outcome of a set of complex processes involving the hybridization of highly diversified and loosely related sets of competencies (Zeppini and Bergh 2011; Fusillo, Quatraro et al. 2022; Fusillo 2023).

More recently, the evolutionary economic geography literature has contributed to extending such arguments to the geographical context, investigating the dynamics of the geography of green technologies. As a matter of fact, sectoral and technological diversification in eco-innovation is considered a priority by national and local governments in various geographical contexts, across both developed and developing countries. Yet, our understanding remains limited concerning the ability of regions to engage in eco-innovation and diversify their portfolio of green technologies (Losacker, Hansmeier et al. 2023). Extant research has shown that regions tend to diversify in green domains through a process of regional branching, thus relying both on green and non-green pre-existing capabilities (Tanner 2014; Tanner 2016; Montresor and Quatraro 2020). These findings align with the fundamental tenet of the relatedness approach, according to which regions primarily diversify by developing new technologies characterised by higher proximity to their pre-existing ones (Boschma 2017b). However, similarly to the general case of technological diversification (in any sector), different factors have been found to play a crucial role, besides as well as together with relatedness, for green-tech branching. In particular, while most of the available studies have focused on local factors, only limited attention has been paid to the factors that affect the green-tech diversification of regions by acting across their geographical boundaries, namely, extra-regional factors.

Recent contributions highlight that the development of eco-innovation at the local level and the consequent regional specialization in green sectors are highly dependent on the availability of heterogeneous knowledge sources. In particular, developing green technologies and, hence, technological specializations, require global connections (Castellani, Marin et al. 2022b; Corrocher et al. 2024). Global connectedness is crucial to accessing foreign knowledge and building, as well as strengthening, the endogenous green domestic capabilities (Amendolagine, Lema et al. 2021; Castellani, Marin et al. 2022b; De Marchi et al. 2022). MNEs and FDIs play a pivotal role in this respect. The core tenet brought forward by the international business literature on MNEs is that they are characterised by a distinctive ability to transfer, integrate and exploit knowledge from their geographically dispersed network. In so doing, MNEs act as channels for global pipelines of knowledge, which intersect with the so-called local "buzz" of domestic knowledge, thereby enhancing regional capabilities (Bathelt and P. Li 2020; Bathelt, Malmberg et al. 2004b).

Given the recognised properties of FDIs in shaping the knowledge and assets of the recipient locations, they certainly represent one of the most relevant channels of global connectedness, through which external knowledge gets into local contexts and interacts with local knowledge (Marino and Quatraro 2022). Through their investments, MNEs hence act as agents of structural change. These arguments have been recently applied to the eco-innovation field, thus identifying FDIs as a key source of green knowledge at the firm and region level. Analysing green inward FDIs in renewable energy sectors worldwide, Amendolagine, Lema et al. 2021 find that they enhance the overall orientation to sustainability of MNEs, strengthening their innovation activities related to green technologies. Investigating green innovation involving MNEs and their subsidiaries across 14 EU countries, De Marchi et al. (2022) show that MNEs' affiliates display superior performance in green innovation than domestic companies. Similarly, Amendolagine, Hansen et al. (2023) find that subsidiaries of MNEs created through green inward FDIs generate more green patents than locally owned firms thanks to fruitful bi-directional knowledge flows based on continuous interactions and learning processes. At the region level, Castellani, Marin et al. (2022b) found that inward innovative FDIs occurring in green industries across the EU positively influence the occurrence of a region's specialization in green technologies. The empirical evidence supports the argument that the transfer of (green) knowledge occurring via MNEs' dense global relationships has a crucial role in the sustainability transition.

Interestingly, although it is not straightforward to define *green* FDIs when referring to inward FDIs occurring in green-related sectors, the above-reviewed literature seems to clearly point to the argument that eco-innovation at the firm level is more likely to be induced by MNEs engaged in green R&D and inward FDIs characterised by environmental content. Yet, what remains unexplored is whether FDIs in green sectors influence the extent of regional technological diversification into green domains. We argue that inward FDIs are a source of green knowledge necessary for regions not only for specializing in green technologies, but also to expand the set of green technological domains they diversify into. Throughout FDIs and MNEs subsidiaries, host regions can access non-local (or global) green knowledge beyond their geographical boundaries, which is new with respect to the existing local one and can be possibly transferred and exploited for the sake of their green-tech transition (Neffke et al. 2018b; Boschma 2022b). Specifically, throughout a process of hybridization of local and global green knowledge, the combination of heterogeneous and loosely related knowledge inputs likely increases the probability of developing a specialization in a green domain that is new with respect to the pre-existing ones, hence spurring local green technological diversification.

Against this background, we posit the following hypothesis:

*Hypothesis 1:* Inward green FDIs are positively associated with regional technological diversification in green domains.

### 3.2.2 Recombinant capabilities and the local occupational structure

In the previous Section, it has been stressed that green technologies' reliance on the hybridization of heterogeneous and loosely related knowledge inputs implies that knowledge sourcing across geographical boundaries increases the probability of developing a new green specialization in areas that did not possess a previous comparative advantage in that specific domain.

However, the local availability of diverse knowledge components is not a sufficient condition for the success of green technological efforts. The presence of adequate capabilities is also crucial to ensure that heterogeneous knowledge inputs are effectively combined together for the production of green technological knowledge. An increasing number of studies have indeed framed the analysis of the antecedents of green technological knowledge leveraging on the distinction between recombinant reuse and creation capabilities (Carnabuci and Operti 2013b).

While recombinant reuse involves the refinement and improvement of known technological combinations, recombinant creation capabilities concern experimentation with unexplored combinations. Empirical studies exploiting patent-based indicators have shown that the command of recombinant creation mechanisms is crucial in generating green patents (Orsatti, Quatraro and Pezzoni 2020a; Quatraro and Scandura 2019) and in the dynamics of regional green diversification in Europe (Orsatti, Quatraro and Scandura 2023).

In this context, existing literature has emphasized that regional recombinant capabilities are influenced by human capital endowment. Specifically, it has been found that differences in green patenting across territories are associated with the prevalence of exploration-oriented skills, measured by the share of abstract skills in the local occupational structure. Within the task-based approach (Acemoglu and D. Autor 2011), abstract skills feature non-routine cognitive occupations that range from corporate managers to science and technology professionals. These latter, hence, involve occupations and skills that leverage recombinant creation capabilities to generate new scientific and technological knowledge and inventions (Orsatti, Perruchas et al. 2020).

Accordingly, the composition of the local skill set plays a significant role in influencing regional variations in the ability to support the generation of green technologies. The prominence of abstract skills holds crucial importance in this context, as they correlate with cognitive capabilities to integrate concepts and resources from diverse domains into fresh and unexplored avenues. By the same token, regional technological diversification in the green domain can be favoured by the configuration of local occupational structures characterized by a relatively high share of abstract skills.

In view of the above discussion, we spell out the following hypothesis:

*Hypothesis 2:* The local prevalence of abstract skills is positively associated with regional technological diversification in green domains.

### 3.2.3 The interplay between external knowledge and the local occupational structure

The skill composition of the local occupational structure can explain regional differences in terms of the capacity to develop new green technological specializations. High-quality human capital endowed with exploration-oriented skills can better bear boundary-spanning R&D and innovation efforts. Yet, not only scientific boundaries matter in this context.

A stylized fact within the geography of innovation is that scientific and technological capabilities are place-specific. This is the main reason why the injection of knowledge and capabilities from foreign places is expected to positively influence the local diversity and structure of the knowledge base, increasing the degree of unrelated diversification (Neffke et al. 2018b). As recalled above, this likely represents a key distinctive advantage for regions willing to develop new green technological specializations.

These arguments entail that inward green FDIs and the local occupational structure may interact in two distinct and opposite directions that can be conceptualized by elaborating upon the absorptive capacity argument. According to Cohen and Levinthal (1989) and Cohen and Levinthal (1990), absorptive capacity stands for agents' ability to spot and assimilate valuable external knowledge, and to exploit it commercially. An extensive stream of literature has documented the importance of this capacity for firms' innovative performances, focusing on learning dynamics emerging out of accumulated R&D investments, or on the quality of human capital within firms' boundaries (Gambardella 1992; Romijn and Albaladejo 2002; Lund Vinding 2006).

Further articulations of the absorptive capacity argument have proposed that absorptive skills are heterogeneous and may involve different tasks depending on whether the activity concerns the sourcing of knowledge or its exploitation. In the first case, *potential absorptive capacity* indicates the ability to explore, recognise, acquire and assimilate useful external knowledge. In contrast, *realised absorptive capacity* stands for the ability to transform and apply acquired knowledge effectively within organisations (Zahra and George 2002; Mason et al. 2020).

Accordingly, the interplay between the local occupational structure and global knowledge sourcing through FDIs can take two opposite forms, depending on the kind of absorptive capacity that prevails in the local occupational structure. On the one hand, given their characteristics, local qualified science and technology workers in abstract-skills intense occupations can be considered as the agents of potential absorptive capacity. Accordingly, they will tend to act as local brokers of global knowledge, specializing in the screening across geographically dispersed knowledge sources (Malecki 2002; Parjanen et al. 2011). In this case, high local shares of abstract-skills-intense occupations would compensate for weak flows of incoming green FDIs.

On the other hand, locally qualified workers in routine-intensive occupations can be considered agents of realised absorptive capacity, showing distinctive capabilities in translating globally sourced knowledge into new technologies that are eventually marketable. This capacity is more likely to feature professionals who, in some cases, can work at the crossroads between high and intermediate-level occupations (Mason et al. 2020; Girma 2005; Consoli, Fusillo et al. 2021). In this case, local high shares of routine-skills-intense occupations would complement incoming flows of green FDIs.

This discussion leads to the following hypotheses:

*Hypothesis 3a:* The local prevalence of abstract skills compensates for the effects of inwards FDIs on regional green diversification.

*Hypothesis 3b:* The local prevalence of routine skills augments the effects of inwards FDIs on regional green diversification

## 3.3 Empirical framework

### 3.3.1 Data and variables

To test our hypotheses, we build an original dataset on regional green-tech diversification, green FDIs and abstract skills for a balanced panel of 287 US Metropolitan Statistical Areas (MSAs) observed over the period 2003-2018.[1]

[1] In order to ensure comparability and consistency of territorial units over time and across different data sources, we follow the procedure described in Consoli, Fusillo et al. (2021) that allows the unique

We investigate the technological diversification of MSAs in the green domain by looking at the green technologies that enter the regional knowledge base. Following consolidated literature, we calculate *Entry* as a dummy variable taking value one if MSA $r$ acquires a Revealed Technological Advantage (RTA) – i.e., a technological specialization – in technology $i$, at time $t$ provided that it was not observed at time $t-1$. In order to measure the acquisition of a new green specialization, our main dependent variable $EntryGT_{i,r,t}$ is expressed as a dummy variable equal to one if technology $i$ is identified as a green technology. Formally:

$$EntryGT_{i,r,t} = 1 \quad \text{if } RTA_{i,r,t} \geq 1 \text{ and } 0 \leq RTA_{i,r,t-1} < 1 \tag{3.1}$$

where $RTA_{i,r,t}$ is defined as follows:

$$RTA_{i,r,t} = \frac{\frac{p_{i,r,t}}{\sum_i p_{i,r,t}}}{\frac{\sum_{r,t} p_{i,r,t}}{\sum_r \sum_i p_{i,r,t}}} \tag{3.2}$$

with $p_{i,r,t}$ is the number of patents in technology $i$, in MSA $r$, at time $t$.

The dependent variable is constructed using the PatentsView database, which contains detailed information on patent applications filed at the United States Patent and Trademark Office (USPTO).[2] In order to identify green technologies, we follow the ENV-TECH definition provided by the OECD (Haščič and Migotto 2015b) and group CPC codes for green technology

identification of MSAs over changing county composition.

[2]PatentsView database is available at: https://patentsview.org/

classes at the 4-digits level.[3] Patents have been assigned to MSAs based on the inventor's location as provided by PatentsView. Patents with multiple inventors residing in different MSAs have been assigned to each MSA.

The first key regressor, aiming at testing H1, measures the capacity of regions to attract FDIs in the green domain. To do so, we collect data on greenfield FDI from the fDi Markets database and work out the total amount of green-related inward FDI capital expenditure in each MSA (GreenCapex).[4] Greenfield FDIs, as opposed to Mergers and Acquisitions (M&A), are investments entailing the opening of a new firm, establishment, or factory in a foreign location. Each project recorded in fDi markets databese, contains information about the location of the investing firm, and the city of destination of the investment. We assign, by geolocating the city of destination, each investment project to a Metropolitan Statistical Area. In order to define whether a given investment is related to environmental technologies, we follow the methodology proposed by Castellani, Marin et al. (2022b). Accordingly, each investment in the fDi Markets database can be assigned to a NAICS sector. We start by assigning patents from PatentsView to NAICS sectors, based on the crosswalks provided in Lybbert and Zolas (2014a). In turn, we calculate the RTA in green technologies for each of these sectors,

[3]The OECD's classification for environmental technologies includes the following classes of patents: Environmental Management, Water-related Adaptation Technologies, Climate Change Mitigation Technologies Related To Trasportation, Climate Change Mitigation Technologies Related To Buildings, Climate Change Mitigation Technologies In The Production Or Processing Of Goods, Climate Change Mitigation Technologies Related To Energy, Climate Change Mitigation Technologies Related To Wastewater Treatment Or Waste Management, Capture, Storage, Sequestration Or Disposal Of Greenhouse Gasses

[4]fDi Markets is a database maintained by the Financial Times Intelligence Unit that tracks foreign-direct investments in greenfield projects. The database records flows of cross-border greenfield investments since 2003.

based on the patenting activity within the United States. Exploiting the sectoral specialization in green technologies, we link FDIs to green sectors and calculate regional measures of FDI inflows of green-related investments.We build two measures of green FDIs at the MSA level. First, we build a variable based on the amount of capital expenditures for inward FDIs in green sectors for each MSA. We build our regressors of interest as a five-year moving average in order to account for the volatility in FDI inflows[5]. Second, we build a dummy variable equal to one if the moving average of green FDIs in an MSA at time $t$ is greater than zero. In this way, we will be able to appreciate the effect of FDIs (green and brown) at the intensive and the extensive margins, respectively.[6].

The second key explanatory variable, is aimed at capturing the local endowment of abstract-skilled labor force. Accordingly, we employ and adapt the task-based framework initially proposed by D. H. Autor, Levy et al. (2003), along with its geographical adaptation (D. H. Autor and Dorn 2013). The rationale behind this framework lies in considering occupations as aggregations of tasks matched with the skills required to perform such tasks (D. H. Autor, Levy et al. 2003), allowing to delineate occupational structures based on individual attributes rather than educational proxies (Consoli and Rentocchini 2015; Vona and Consoli 2015). Following the task-based frame-

[5]In robustness checks, we ensure the consistency of our results by accounting for the volatility of FDI flows in different ways. These include calculating stocks using the permanent inventory method at different discount rates.

[6]The volume of capital expenditures for each FDI flow is estimated by FT Intelligence (see https://www.fdimarkets.com/faqs) In order to account for potential measurement problems, in the robustness checks section, we show the results of the estimations, including the share of green over total FDIs, expressed as a count of projects and as capital expenditures, assuming that the potential measurement error does not vary according to our classification of green and non-green FDIs.

work, abstract-skilled jobs encompass tasks demanding creativity, intuition, problem-solving and persuasion, typically found in professional, managerial, technical and creative occupations such as law, medicine, science, engineering, marketing and design. Hence, these abstract tasks are generally carried out by highly educated individuals with strong analytical skills.

In order to develop an indicator of the abstract task intensity in the MSAs' labor force, we follow previous literature (Fusillo, Consoli et al. 2022; Consoli, Fusillo et al. 2021) and rely on the crosswalk provided by Acemoglu and D. Autor (2011), which directly links occupations from two-digit Standard Occupational Classification codes to their respective task intensities. Collecting employment data by occupation from the Occupational Employment Statistics (OES) program developed by the US Bureau of Labor Statistics (BLS), we compute the abstract task employment share ($ASH$) for each MSA as follows:

$$ASH_{r,t} = \left( \sum_{j=1}^{J} L_{j,r,t} 1 \left[ ATI_j \right] \right) \left( \sum_{j=1}^{J} L_{j,r,t} \right)^{-1} \tag{3.3}$$

where $ASH_{r,t}$ represents the abstract employment share in MSA $r$ at time $t$; $L_{jit}$ is the employment in occupation $j$ in MSA $r$ at time $t$; $ATI_j$ is an indicator function taking value 1 if the corresponding occupation is abstract task intense. Our measure of abstract skills thus consists in the share of employment in abstract-intensive jobs. In the empirical estimations, we express $ASH_{r,t}$ as a dummy variable equal to one if the share of abstract skills in

MSA $r$ at time $t$ is above the median share of abstract skills across all MSAs in the sample, and 0 otherwise.

In order to test H3b, we develop an indicator of routine task intensity in the MSAs' labor force. According to the task-based framework, routine jobs belong to the mid-skill spectrum and entail performing repetitive manual or cognitive (i.e., blue-collar) work tasks. The former generally concerns 'Office and Administrative support' occupations (i.e., clerks) while the latter concerns 'production', 'maintenance and repair' occupations (i.e., performed by blue-collars). In line with the $ASH$ indicator, we compute the routine task employment share ($RSH$) for MSA $r$ at time $t$ as follows:

$$RSH_{r,t} = \left( \sum_{j=1}^{J} L_{j,r,t} 1 \left[ RTI_j \right] \right) \left( \sum_{j=1}^{J} L_{j,r,t} \right)^{-1} \tag{3.4}$$

where $RTI_j$ is an indicator function taking value 1 if the corresponding occupation is routine task intense. Hence, the routine skill indicator measures the share of employment in routine-intensive jobs. For the sake of consistency, in the empirical estimations, $RSH_{r,t}$ is also expressed as a dummy variable equal to one if the share of routine skilled workers in MSA $r$ at time $t$ is above the median across MSAs, and 0 otherwise.

In line with consolidated literature on regional technological diversification and its application to green diversification, we build a control variable for the path-dependent nature of the entry in new specializations (Hidalgo 2021), i.e., the relatedness density of new (green) technological specializations to the pre-existing specializations in the regional knowledge space. The construction of the technological relatedness index entails a number of steps. Firstly, we define a measure of technological proximity following the symmetric measure proposed by Eck and Waltman (2009). Based on the co-occurrence of technology classes within patent documents filed at USPTO, technological proximity is defined as follows:

$$\phi_{i,j} = \frac{mc_{ij}}{s_i s_j} \tag{3.5}$$

where $m$ is the total number of patents, $c_{ij}$ is the number of co-occurrences between the two technologies, and $s_i$, $s_j$ are, respectively, the total amount of occurrences of the two technologies. Hence, $\phi_{i,j}$ is a symmetric technology-by-technology proximity matrix. Secondly, we measure the relatedness density of technology $i$, in region $r$, at time $t$, as the average proximity between the specializations that the region has at time $t-1$, and the focal technology $i$:

$$TechRel_{i,r,t} = \frac{\sum_{j \in r} \phi_{ij} RTA_{j,r,t-1}}{\sum_{j \in r} RTA_{j,r,t-1}} \tag{3.6}$$

The final dataset consists of 287 time-consistent MSAs and 666 technologies (of which 74 are identified as green technologies) observed over the period 2003-2018.

Finally, we build a number of control variables varying by MSA and year. First, we create a measure of FDIs in non-green sectors, called $BrownCapex$, calculated as its counterpart $GreenCapex$, which will be used in some of the estimated models. Second, we work out the level of economic development of the MSAs, measured as per capita Gross Domestic Product ($GDPpc$). To build this variable, we collect data on GDP by MSA from the US Bureau of Economic Analysis and population data from the Bureau of Labour Statistics. Third, we control for the rate of employment growth ($EmpGrowth$) in the MSA. To do so, we collect employment data from the County Business Patterns and build the employment growth measure by adapting the indicator described in Haltiwanger et al. (2013) to the MSA level. Precisely, employment growth in MSA $r$ at time $t$ is defined as the difference between the MSA employment level at time $t$ and the MSA employment at time $t-1$, divided by the average employment level in the MSA over the two periods.[7] Fourth, we calculate, in line with Castellani, Marin et al. (2022b), the share of patents of the MSA within the total patents in the country ($ShPatents$). In addition, we calculate the share of establishments in green sectors ($shGreenEst$), where the definition of green sectors follows the one proposed to identify green FDIs, as described in Section 3.3.1. $shGreenEst$ is expressed as a dummy variable taking value one if the share of green establishments in a given MSA is above the national median, 0 otherwise. We also include in the list of control variables the extent of regional pre-existing green specializations through a variable called $GreenPrevRTA$, which counts the total number of specializations in green technologies for MSA $r$ at time $t-1$.

A synthetic description of the variables is provided in Table 3.1, while

[7]Let $E_{r,t}$ represent the level of employment in MSA $r$ at time $t$; formally, the employment growth rate is calculated as: $(E_{r,t} - E_{r,t-1})/X_{r,t}$, where $X_{r,t} = (E_{r,t} + E_{r,t-1}) * 0.5$. The measure is symmetric and centered around 0, and has the advantage of sharing log-difference measures properties while accommodating entry and exit (Haltiwanger et al. 2013)

their summary statistics are reported in Table 3.2.

**Table 3.1:** Variable description

| Variable | Description | Source |
|---|---|---|
| EntryGT | Entry in low-carbon technological specializations, with respect to t-1. | USPTO |
| GreenCapex | Capital expenditures for inward FDIs in green sectors (5 years moving average). | FDIMarkets (FT Intelligence) |
| (d) GreenCapex | Dichotomous indicator equal to one if the moving average of green Capex is greater than 0, and 0 otherwise | FDIMarkets (FT Intelligence) |
| BrownCapex | Capital expenditures for inward FDIs in non-green sectors (5 years moving average). | FDIMarkets (FT Intelligence) |
| (d) BrownCapex | Dichotomous indicator equal to one if the moving average of green Capex is greater than 0, and 0 otherwise | FDIMarkets (FT Intelligence) |
| shGreenCapex | Share of capital expenditures inward FDIs in green-intensive sectors over total inward FDIs (5 years moving average). | FDIMarkets (FT Intelligence) |
| ASH | Employment share of abstract-skilled workers | Bureau of Labor Statistics |
| (d) ASH | dichotomous indicator equal to one if the employment share of abstract skilled workers is above the national median, and 0 otherwise | Bureau of Labor Statistics |
| RSH | Employment share of routine-skilled workers | Bureau of Labor Statistics |
| (d) RSH | Dichotomous indicator equal to one if the employment share of routine skilled workers is above the national median, and 0 otherwise | Bureau of Labor Statistics |
| EmpGrowth | Employment growth | County Business Patterns |
| TechRel | Relatedness density of the new specializations with respect to the pre-existing ones. | USPTO |
| GDPpc | Gross Domestic Product per capita. | US Bureau of Economic Analysis, US Bureau of Labour Statistics |
| ShPatents | Total number of patents, as a share of the total patents in the country. | USPTO |
| shGreenEst | Share of establishments in green-intensive sectors over total establishments. | County Business Patterns |
| (d) shGreenEst | Dichotomous indicator equal to one if the share of establishments in green sectors is above the national median, and 0 otherwise | Bureau of Labor Statistics |
| GreenPrevRTA | Count of existing green specializations. | USPTO |

### 3.3.2 Methodology

To investigate the relationship between green-tech diversification, green FDIs, and abstract skills (H1 and H2), we estimate the following model:

$$\begin{aligned} EntryGT_{i,r,t} = \beta_1 GreenCapex_{r,t-1} + \beta_2 ASH_{r,t-1} + \beta_3 RSH_{r,t-1} \\ \beta_4 TechRel_{i,r,t} + \psi \boldsymbol{X'}_{r,t-1} + \delta_{i,t} + \gamma_s + \epsilon_{i,r,t} \end{aligned} \tag{3.7}$$

where $i$, $r$, and $t$ index, respectively, technology, MSA, and year. *EntryGT*

**Table 3.2:** Summary Statistics

| | N | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|
| EntryGT | 2,472,858 | 0.010 | 0.100 | 0.000 | 1.000 |
| GreenCapex | 2,472,858 | 55.680 | 218.190 | 0.000 | 3,760.140 |
| BrownCapex | 2,472,858 | 93.190 | 321.570 | 0.000 | 6,107.350 |
| (d) GreenCapex | 2,472,858 | 0.600 | 0.490 | 0.000 | 1.000 |
| (d) BrownCapex | 2,472,858 | 0.760 | 0.430 | 0.000 | 1.000 |
| shGreenCapex | 2,472,858 | 0.170 | 0.190 | 0.000 | 1.000 |
| GreenFDI | 2,472,858 | 0.900 | 2.380 | 0.000 | 27.200 |
| BrownFDI | 2,472,858 | 3.020 | 12.420 | 0.000 | 192.800 |
| ASH | 2,472,858 | 0.200 | 0.040 | 0.100 | 0.400 |
| RSH | 2,472,858 | 0.420 | 0.050 | 0.280 | 0.680 |
| TechRel | 2,472,858 | 10.940 | 12.800 | 0.000 | 100.000 |
| GDPpc | 2,472,858 | 47,174.430 | 12,821.290 | 20,320.000 | 171,389.060 |
| EmpGrowth | 2,472,858 | 0.007 | 0.032 | −0.270 | 0.206 |
| ShPatents | 2,472,858 | 0.000 | 0.010 | 0.000 | 0.120 |
| shGreenEst | 2,472,858 | 0.010 | 0.000 | 0.000 | 0.020 |
| GTprevRTA | 2,472,858 | 12.790 | 8.750 | 0.000 | 43.000 |

is the probability of observing a new specialization of MSA $r$ in technology $i$ at time $t$, conditional on $i$ being green. $GreenCapex$ and $ASH$ are the explanatory variables of interest, lagged by one year. The former is, alternatively, a dummy equal to 1 for the presence of inward FDIs in MSA $r$ or the share of green FDIs over total FDIs;[8] the latter is a dummy equal to 1 if the intensity of abstract-related skills in MSA $r$ is above the national median. $TechRel$ accounts for the role of technological capabilities at $t-1$, a key driver of regional green diversification. $\boldsymbol{X'}$ is a vector of time-varying control variables at the MSA level, 1 year lagged, that may affect the probability of green-tech diversification. $\delta_{i,t}$ are technology-year fixed effects, included to control for common technological trends likely to affect the probability of specializing in GTs in specific domains[9]; furthermore, we also include State

[8]When we use the share of green FDIs, we correct for the zero inflation in FDIs by including an indicator variable for the absence of inward FDIs in the MSA, following (Aghion, Akcigit et al. 2019).

[9]Technology codes for the fixed effects are build on 1-digit CPC codes, rather than 4-digits, in order

fixed-effects ($\gamma_s$) that allow to exploit within-State MSA variation while controlling for State-level characteristics[10]; $\epsilon_{i,r,t}$ is the idiosyncratic error term.

Finally, we test H3a and H3b augmenting model (3.7) firstly with the interaction term between $GreenCapex$ and $ASH$, secondly, the interaction between $GreenCapex$ and $RSH$, as follows:

$$\begin{aligned} EntryGT_{i,r,t} = {} & \beta_1 GreenCapex_{r,t-1} + \beta_2 ASH_{r,t-1} + \beta_3 RSH_{r,t-1} \\ & \beta_4 GreenCapex_{r,t-1} \times ASH_{r,t-1} + \\ & \beta_5 TechRel_{i,r,t} + \psi\chi_{r,t-1} + \gamma_s + \delta_{i,t} + \epsilon_{i,r,t} \end{aligned} \tag{3.8}$$

$$\begin{aligned} EntryGT_{i,r,t} = {} & \beta_1 GreenCapex_{r,t-1} + \beta_2 ASH_{r,t-1} + \beta_3 RSH_{r,t-1} \\ & \beta_4 GreenCapex_{r,t-1} \times RSH_{r,t-1} + \\ & \beta_5 TechRel_{i,r,t} + \psi\chi_{r,t-1} + \gamma_s + \delta_{i,t} + \epsilon_{i,r,t} \end{aligned} \tag{3.9}$$

We estimate Equation (3.7) and Equation (3.8) through fixed effects Logit estimators. In all specifications, continuous variables are log-transformed applying the inverse hyperbolic sine function, and standard errors are clustered at the MSA level to account for heteroskedasticity.

to allow for within-code variation, and maintain a larger sample. In robustness checks, we test for 4-digit by year fixed effects.

[10] State-level policies and taxation incentives for foreign companies, for example, might affect location choice for inward FDIs.

## 3.4 Results

### 3.4.1 Green FDI, skills and green diversification

We estimate several specifications of Equation (3.7) to test our working hypotheses. Table 3.3 reports the results concerning the impact of FDIs at the intensive margin, i.e. when $GreenCapex$ and $BrownCapex$ are calculated as the 5-year moving average of capital expenditure. Column (1) is the most conservative model in which we only include $BrownCapex$, year$\times$technology, and State fixed effects as control variables. We find positive and significant coefficients for both $ASH$ and $GreenCapex$ (point estimates, respectively, .04 and .27). This result supports our hypotheses H1 and H2. Then, from column (2) to column (4), we gradually augment the initial specification with the full set of control variables. In all specifications, we find positive and significant coefficients for both $ASH$ and $GreenCapex$ (point estimates ranging between .02 and .04 for $GreenCapex$, and between .16 and .2 for $ASH$), in line with the baseline results. For what concerns $ASH$, results show that the odds of acquiring a new specialization in green domains for MSAs endowed with above-national median levels of abstract-oriented occupational skills are higher than those of areas with below median intensity of abstract skills (with odds ranging between 17% and 22%). This result suggests that high levels of abstract-intense jobs in the local skill composition are associated with greater capability to integrate diverse concepts and resources which, leveraging on recombinant creation capabilities, favour generating GTs and acquiring technological specialization in the green domain. Moreover, the presence of inward green FDIs – a proxy for the local capacity to attract external competencies and knowledge in green domains – is a lever for diversifying towards GTs. Indeed, MSAs capable of attracting 1% additional capital expenditures through green FDIs are around 2.2% more likely to acquire new green specializations. This confirms that through inward green FDIs, local areas can access external (possibly global) green knowledge, which, coupled with existing local competencies, can be successfully exploited for novel recombinations, hence allowing regions to expand the set of technological specialization toward the green realm and accelerating their green transition. These results corroborate working hypotheses H1 and H2 proposed in Section 3.2.

We now move to the interplay between green FDIs and the local skills' endowment. Columns (5) and (6) of Table 3.3 provide evidence related to H3a and H3b, respectively. We are interested in the coefficients of the interactions between the $GreenCapex$ and the skills-related variables.

On the one hand, H3a posits that workers in abstract-skills-intensive occupations can be considered agents of *potential absorptive capacity.* In this direction, we expect that the local intensity of those skills compensates for the impact of the inflow of external knowledge on the entry into new green technological specializations. Areas with an abundance of workers with exploration-oriented skills are less dependent on trade flows to source external knowledge. The coefficient of interest, $\beta_4$ in Equation 3.8, is negative and significant, as reported in column (5), supporting H3a.

On the other hand, the successful exploitation of external knowledge requires the local availability of skills and capabilities, allowing the adaptation of knowledge created elsewhere to local idiosyncratic conditions. Following the arguments underlying H3b, local qualified workers in routine-skills-intensive occupations can be considered as carriers of such *realised absorptive capacity*. In this direction, routine-skills can work as enablers of the positive impact of FDIs on green technological diversification. The coefficient of interest, $\beta_4$ in Equation 3.9, is positive and significant, as reported in column (6), supporting H3b.

With respect to the control variables, we estimate positive and statistically significant coefficients for the variable $TechRel$ in all specifications, as expected. Indeed, the literature has widely recognised technological relatedness as a key driver of regional technological diversification in general and green diversification in particular. We confirm this well-documented evidence in our sample. Moreover, we estimate positive and statistically significant coefficients for the variables $BrownCapex$, $shGreenEst$, and $GreenPrevRTA$. The variable $BrownCapex$ allows to control for the overall non-green capacity of attracting inward FDIs, hence incoming foreign direct investments that bring knowledge and expertise from outside the region to local innovative actors. GTs are complex technologies requiring the recombination of heterogeneous and distant knowledge. Since inward FDIs help regions widen and diversify their knowledge bases, this favors regional branching, hence the entry into new green specializations. The variable $shGreenEst$ captures the role of local industrial composition exposed to environmental innovation activities. The higher the number of firms operating in sectors exposed to green activities, the higher the likelihood that the local system adds new green specializations to its technological portfolio. Lastly, we include the variable $GreenPrevRTA$ to control for the regional cumulative process of green specialization. Together with $TechRel$ and $shGreenEst$, we include this variable to control for the effect of the accumulation of local innovation competencies in environmental-related technological fields. The higher this accumulation, the higher the likelihood of green diversification, at least in the short term and up to a certain threshold of saturation of technological opportunities.

The previous analysis shows that the MSA's capacity to attract FDIs, both green and non-green, is positively associated with the likelihood of entering a new green technological specialization. While these results provide evidence of the impact of FDIs at the intensive margins, the extensive margins are also important when investigating the impact of international trade dynamics. For this reason, in Table 3.4, we provide the results of the econometric estimations in which brown and green FDI variables have been dichotomized.

The Table's structure is the same as Table 3.3. Column (1) reports the results of the most conservative model, in which we include only our regressors

**Table 3.3:** Green Diversification, skill composition, and Green FDI Capex

| | entryGT | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| GreenCapex | 0.0436*** | 0.0267*** | 0.0366*** | 0.0222*** | 0.0406*** | 0.0058 |
| | (0.0090) | (0.0084) | (0.0080) | (0.0080) | (0.0108) | (0.0092) |
| BrownCapex | 0.0340*** | 0.0133 | 0.0321*** | 0.0197** | 0.0219*** | 0.0215** |
| | (0.0098) | (0.0089) | (0.0087) | (0.0083) | (0.0083) | (0.0083) |
| (d) ASH | 0.2753*** | 0.2016*** | 0.1835*** | 0.1576*** | 0.2262*** | 0.1534*** |
| | (0.0471) | (0.0426) | (0.0400) | (0.0361) | (0.0422) | (0.0357) |
| (d) RSH | 0.0298 | 0.0463 | -0.0107 | -0.0283 | -0.0343 | -0.1140** |
| | (0.0431) | (0.0396) | (0.0356) | (0.0339) | (0.0333) | (0.0444) |
| TechRel | | 0.0143*** | 0.0143*** | 0.0125*** | 0.0126*** | 0.0126*** |
| | | (0.0009) | (0.0008) | (0.0008) | (0.0008) | (0.0008) |
| GDPpc | | | 0.2879*** | 0.1621* | 0.1715* | 0.1748* |
| | | | (0.1009) | (0.0984) | (0.0972) | (0.0984) |
| EmpGrowth | | | -0.0453 | -0.1767 | -0.1754 | -0.1563 |
| | | | (0.3985) | (0.3873) | (0.3841) | (0.3870) |
| ShPatents | | | -0.1905*** | -0.1755*** | -0.1644*** | -0.1661*** |
| | | | (0.0296) | (0.0306) | (0.0296) | (0.0307) |
| (d) shGreenEst | | | | 0.2986*** | 0.2898*** | 0.2925*** |
| | | | | (0.0866) | (0.0848) | (0.0852) |
| GreenPrevRTA | | | | 0.6514*** | 0.6822*** | 0.6503*** |
| | | | | (0.1841) | (0.1817) | (0.1818) |
| GreenCapex × (d) ASH | | | | | -0.0354*** | |
| | | | | | (0.0118) | |
| GreenCapex × (d) RSH | | | | | | 0.0347*** |
| | | | | | | (0.0117) |
| Observations | 2,102,562 | 2,102,562 | 2,102,562 | 2,102,562 | 2,102,562 | 2,102,562 |
| Log-Likelihood | -108,356.0 | -108,022.8 | -107,854.2 | -107,798.5 | -107,786.0 | -107,785.0 |
| Adjusted Pseudo $R^2$ | 0.06667 | 0.06953 | 0.07095 | 0.07142 | 0.07151 | 0.07152 |
| Year*Tech fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| State fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Clustered (MSA) standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dep var: MSA entry in a green technological specialization. *GreenCapex* and *BrownCapex* are expressed as continuous variables. *ASH* and *RSH* are dichotomous variables equal to one if the share of, respectively, abstract and routine skills is above the national median. Explanatory variables are lagged by one year. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models are estimated through fixed effects logit estimators. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the MSA level. *p<0.1; **p<0.05; ***p<0.01

of interest, along with year×technology and State fixed effects. At the extensive margin, areas featured by non-zero $GreenCapex$ and $BrownCapex$ have higher odds of diversifying in green technological domains than areas with no FDIs. This result further supports H1. The coefficient of $(d)ASH$ is also positive and significant, suggesting that higher local levels of abstract skills increase the likelihood of entering new green technological specializations, in line with H2. In columns (2) to (4), we gradually include the full set of control variables. The results concerning our focal regressors do not change qualitatively, though the inclusion of control variables causes a modest reduction in the magnitude of the estimated coefficients. Finally, in columns (5) and (6), we report the results of the estimations of models 3.8 and 3.9, which are intended to test H3a and H3b. The results are consistent with those reported in Table 3.3. The coefficient of the interaction between $(d)ASH$ and $(d)GreenCapex$ is negative and significant, suggesting the existence of a compensation mechanism according to which local areas with high levels of abstract skilled workers are less dependent on FDIs in the dynamics of green technological diversification. In column (6), the coefficient of the interaction between $(d)GreenCapex$ and $(d)RSH$ is positive and significant, providing further support to H3b concerning the enabling role of routine skills in the impact of green FDIs on the probability of opening a new green technological specialization. Overall, the magnitude of the coefficients at the extensive margins is larger than that of the coefficients at the intensive margins.

**Table 3.4:** Green Diversification, skill composition, and Green FDI Capex (dummies)

| | entryGT | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| (d) GreenCapex | 0.2078*** | 0.1190*** | 0.1409*** | 0.0726*** | 0.1602*** | 0.0046 |
| | (0.0337) | (0.0301) | (0.0300) | (0.0278) | (0.0398) | (0.0302) |
| (d) BrownCapex | 0.2166*** | 0.1518*** | 0.1657*** | 0.1206*** | 0.1300*** | 0.1170*** |
| | (0.0425) | (0.0389) | (0.0385) | (0.0346) | (0.0345) | (0.0343) |
| (d) ASH | 0.3418*** | 0.2173*** | 0.2052*** | 0.1599*** | 0.2383*** | 0.1601*** |
| | (0.0423) | (0.0371) | (0.0369) | (0.0321) | (0.0354) | (0.0319) |
| (d) RSH | 0.0411 | 0.0431 | -0.0038 | -0.0329 | -0.0415 | -0.1265*** |
| | (0.0394) | (0.0351) | (0.0323) | (0.0306) | (0.0300) | (0.0374) |
| TechRel | | 0.0155*** | 0.0157*** | 0.0128*** | 0.0128*** | 0.0128*** |
| | | (0.0009) | (0.0009) | (0.0008) | (0.0008) | (0.0008) |
| GDPpc | | | 0.3906*** | 0.2068** | 0.2195** | 0.2134** |
| | | | (0.0977) | (0.0933) | (0.0927) | (0.0937) |
| EmpGrowth | | | 0.2052 | -0.0286 | 0.0423 | 0.0139 |
| | | | (0.3498) | (0.3340) | (0.3347) | (0.3360) |
| ShPatents | | | -0.1531*** | -0.1545*** | -0.1510*** | -0.1514*** |
| | | | (0.0262) | (0.0264) | (0.0258) | (0.0261) |
| (d) shGreenEst | | | | 0.3359*** | 0.3399*** | 0.3412*** |
| | | | | (0.0790) | (0.0773) | (0.0776) |
| GreenPrevRTA | | | | 0.8439*** | 0.8480*** | 0.8430*** |
| | | | | (0.1760) | (0.1741) | (0.1742) |
| (d) GreenCapex × (d) ASH | | | | | -0.1699*** | |
| | | | | | (0.0382) | |
| (d) GreenCapex × (d) RSH | | | | | | 0.1679*** |
| | | | | | | (0.0407) |
| Observations | 2,867,130 | 2,867,130 | 2,675,988 | 2,675,988 | 2,675,988 | 2,675,988 |
| Log-Likelihood | -144,517.0 | -143,957.5 | -134,730.9 | -134,610.0 | -134,590.5 | -134,591.2 |
| Adjusted Pseudo $R^2$ | 0.06665 | 0.07025 | 0.07150 | 0.07232 | 0.07245 | 0.07244 |
| Year*Tech fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| State fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Clustered (MSA) standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dep var: MSA entry in a green technological specialization. *GreenCapex* and *BrownCapex* are expressed as dichotomous variables equal to one if the stock of, respectively, green or non-green FDI is greater than 0. *ASH* and *RSH* are dichotomous variables equal to one if the share of, respectively, abstract and routine skills is above the national median. Explanatory variables are lagged by one year. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models are estimated through fixed effects logit estimators. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the MSA level. *p<0.1; **p<0.05; ***p<0.01

The estimates of the marginal effects of the two variables of interest – graphically reported in Figures 3.1 and 3.2 – unveil interesting dynamics. Figure 3.1 shows that for low and medium levels of green FDI intensity, the expected probability of acquiring a new green specialization is always higher in MSAs endowed with a high level of abstract skills if compared to MSAs for which these skills are relatively lacking, though the gap decreases as the intensity of green FDIs increases. For what concerns areas in the top 5% of the distribution of green FDI capital expenditure, we find that substantially no differences can be observed in the probability of green technological diversification based on the relative endowment of abstract skills. In other words, in line with the compensation hypothesis spelt out in H3a, the lower the intensity of green FDIs, the higher the impact of abstract skills (i.e., $ASH = 1$) on green technological diversification, as compared to areas that are poorly endowed with these skills (i.e., $ASH = 0$). The left and right panels of Figure 3.1 show the differential effects for areas where the endowment of routine-skilled workers is below and above the median, respectively. Overall, one cannot find significant differences between the two states.

In Figure 3.2, we show how the impact of the endowment of routine skilled workers on green technological diversification changes along the distribution of $GreenCapex$. The left panel concerns areas in which $ASH = 0$, while the right panel refers to areas in which $ASH = 1$. The diagrams show that above-the-median levels of routine skills yield differential positive effects as compared to below-the-median levels only for high levels of $GreenCapex$.

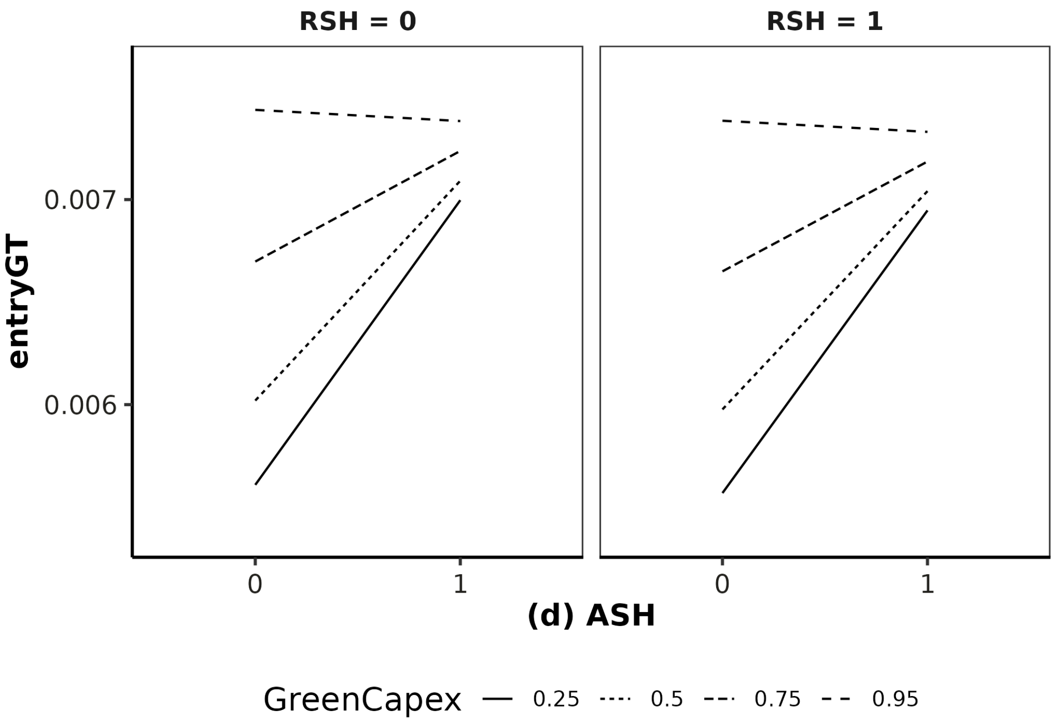



**Figure 3.1:** Marginal effects of the interaction between the share of abstract skills (dummy) and *GreenCapex* at different percentiles of the distribution.

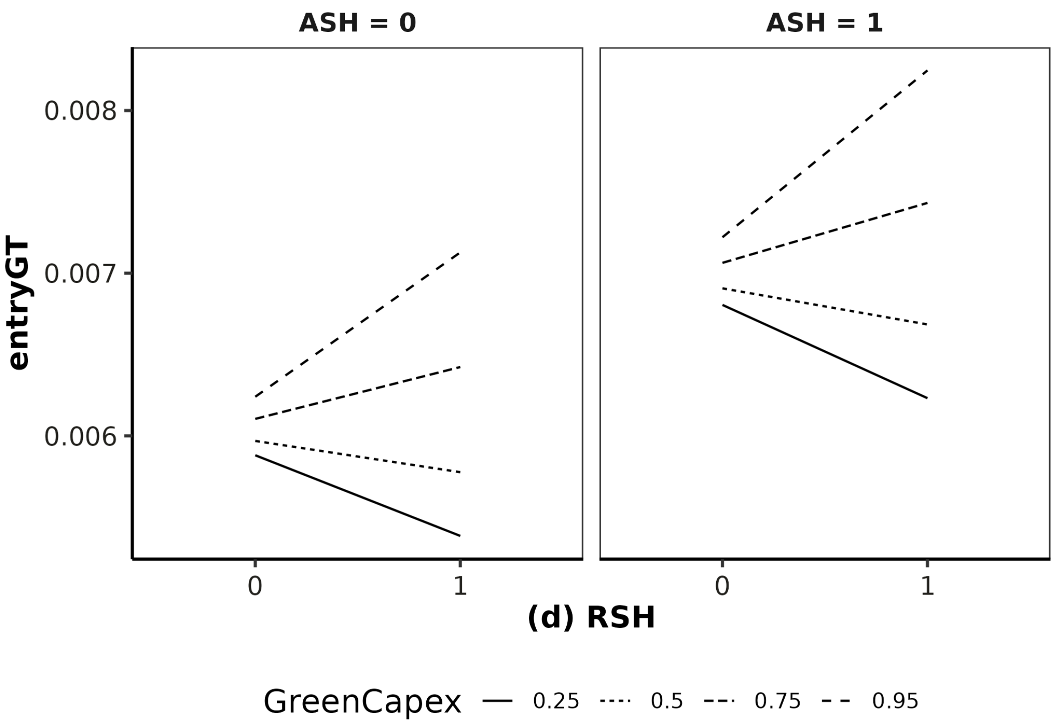



**Figure 3.2:** Marginal effects of the interaction between the share of routine skills (dummy) and *GreenCapex* at different percentiles of the distribution.

Moreover, the impact of $RSH = 1$ grows as $GreenCapex$ increases. This is in line with the augmenting effect hypothesized in H3b, i.e., the argument that routine skills provide local areas with the *realized absorptive capacity* necessary to translate external knowledge into actual innovation.

The evidence discussed so far provides empirical support to our hypotheses concerning the direct effects of green FDIs and abstract skills endowment on the probability of entering new green technological specializations, as well as the interplay between FDI inflows and local skills structure. In the next Section, we investigate sources of heterogeneity and discuss the results of robustness checks.

### 3.4.2 Heterogeneity and robustness

**Heterogeneity: upstream vs. downstream green FDIs**

The data provided by fDi Markets can be exploited to investigate sources of heterogeneity along the value chain. We map FDIs according to their value-chain position, derived from the classification of business functions proposed by Sturgeon (2008). This categorization allows us to map investments at different stages of the value chain consistently across both countries and sectors. We follow the approach proposed in Crescenzi, Pietrobelli et al. (2014) and distinguish investment projects by functions, such as *"Research and Development"*, or *"Retail"*. In turn, we draw upon Ascani, Crescenzi et al. (2016) and group investments into upstream, downstream, and production activities, as summarized in Table 3.5.

**Table 3.5:** Classification of business functions

| Classification | Activity |
|---|---|
| **Upstream** | Business Services |
| | Design, Development and Testing |
| | Education and Training |
| | Headquarters |
| | Research and Development |
| **Downstream** | Customer Contact Centre |
| | Logistics, Distribution and Transportation |
| | Maintenance and Servicing |
| | Recycling |
| | Sales, Marketing and Support |
| | Shared Services Centre |
| | Technical Support Centre |
| **Production** | Construction |
| | Electricity |
| | Extraction |
| | ICT and Internet Infrastructure |
| | Manufacturing |

Following this classification, we distinguish between FDIs concentrated in upstream and production (supply) functions on the one hand and those in downstream (demand) functions on the other one. Hence we run additional estimations of Equation (3.7), Equation 3.8 and Equation 3.9 by jointly including $GreenCapexSupply$ and $GreenCapexDemand$ instead of $GreenCapex$.

Table 3.6 shows the results of the estimations including FDIs-related variables as the 5-year moving average of capital expenditure. Coherently with the previous regression tables, column (1) presents the baseline version of the model. The breakdown of total green FDI expenditure reveals that the positive and significant effect is driven by $GreenCapexSupply$ rather than $GreenCapexDemand$. This latter variable features, indeed, a non-significant and negative coefficient. The coefficient of $(d)ASH$ is positive and significant. In columns (2) to (4), we gradually add the whole set of control variables to the estimations. The evidence is confirmed, as the coefficient of $GreenCapexSupply$ remains positive and significant, and so is the coefficient of $(d)ASH$. The coefficient of $GreenCapexDemand$ remains non-significant. These results suggest that the mechanisms underlying the impact of green FDIs on the probability of green technological diversification are driven by projects clustered in upstream activities, i.e. related to the location of functions mostly focusing on education, training, creativity, and knowledge generation.

Columns (5) and (6) investigate the interplay between $GreenCapexSupply$ and the skill structure of the local workforce (H3a and H3b). The coefficient of the interaction with $(d)ASH$ is negative and significant, while that of interaction with $(d)RSH$ is positive and significant. These results are consistent with those shown in the previous tables and further strengthen our hypotheses about the role of potential and realized absorptive capacity. Workers in abstract skill-intensive occupations deal with creativity and exploration of the knowledge space and hence compensate for the lack of adequate levels of inward FDI. Workers in routine skill-intensive occupations are connected to the production floor and develop tacit knowledge through localized learning dynamics, which is crucial to translating knowledge from outer spaces into

**Table 3.6:** Green Diversification, skill composition, and Supply vs. Demand Green FDI Capex

| | entryGT | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| GreenCapexSupply | 0.0428*** | 0.0302*** | 0.0340*** | 0.0213*** | 0.0386*** | 0.0040 |
| | (0.0092) | (0.0085) | (0.0083) | (0.0082) | (0.0109) | (0.0100) |
| GreenCapexDemand | -0.0036 | -0.0242 | 0.0084 | $-7.39 \times 10^{-5}$ | 0.0079 | 0.0051 |
| | (0.0149) | (0.0149) | (0.0125) | (0.0120) | (0.0114) | (0.0122) |
| BrownCapex | 0.0362*** | 0.0186** | 0.0321*** | 0.0204** | 0.0216*** | 0.0214** |
| | (0.0097) | (0.0088) | (0.0087) | (0.0084) | (0.0084) | (0.0084) |
| (d) ASH | 0.2768*** | 0.2064*** | 0.1824*** | 0.1573*** | 0.2223*** | 0.1534*** |
| | (0.0473) | (0.0423) | (0.0401) | (0.0361) | (0.0413) | (0.0358) |
| (d) RSH | 0.0301 | 0.0459 | -0.0103 | -0.0280 | -0.0350 | -0.1087** |
| | (0.0432) | (0.0393) | (0.0357) | (0.0339) | (0.0333) | (0.0434) |
| TechRel | | 0.0145*** | 0.0143*** | 0.0125*** | 0.0126*** | 0.0126*** |
| | | (0.0009) | (0.0008) | (0.0008) | (0.0008) | (0.0008) |
| GDPpc | | | 0.2883*** | 0.1619 | 0.1727* | 0.1749* |
| | | | (0.1012) | (0.0985) | (0.0977) | (0.0987) |
| EmpGrowth | | | -0.0545 | -0.1851 | -0.1839 | -0.1679 |
| | | | (0.3965) | (0.3850) | (0.3822) | (0.3849) |
| ShPatents | | | -0.1930*** | -0.1751*** | -0.1672*** | -0.1677*** |
| | | | (0.0304) | (0.0310) | (0.0303) | (0.0310) |
| (d) shGreenEst | | | | 0.3011*** | 0.2931*** | 0.2959*** |
| | | | | (0.0868) | (0.0855) | (0.0858) |
| GreenPrevRTA | | | | 0.6569*** | 0.6743*** | 0.6510*** |
| | | | | (0.1866) | (0.1847) | (0.1839) |
| GreenCapexSupply × (d) ASH | | | | | -0.0358*** | |
| | | | | | (0.0118) | |
| GreenCapexSupply × (d) RSH | | | | | | 0.0343*** |
| | | | | | | (0.0120) |
| Observations | 2,102,562 | 2,102,562 | 2,102,562 | 2,102,562 | 2,102,562 | 2,102,562 |
| Log-Likelihood | -108,359.3 | -108,019.3 | -107,855.7 | -107,799.3 | -107,787.0 | -107,786.3 |
| Adjusted Pseudo $R^2$ | 0.06663 | 0.06955 | 0.07093 | 0.07140 | 0.07150 | 0.07150 |
| Year*Tech fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| State fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Clustered (MSA) standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dep var: MSA entry in a green technological specialization. *GreenCapexSupply* is the amount of capital expenditure in green FDIs in supply (upstream and production) sectors. *GreenCapexdemand* is the amount of capital expenditures in green FDI in demand (downstream) sectors. *ASH* and *RSH* are dichotomous indicators equal to one if the share of, respectively, abstract skills or routine skills is above the national median. Explanatory variables are lagged by one year. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models are estimated through fixed effects logit estimators. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the MSA level. *p<0.1; **p<0.05; ***p<0.01

innovation fitting the idiosyncratic conditions of local production systems (Antonelli 2006).

The previous table shows the results concerning the impact of supply and demand green FDIs at the intensive margins. In Table 3.7, we show the results of the estimations at the extensive margins. We move from the baseline model in column (1) to gradually saturate the model in column (4). The results are consistent with those presented in table 3.6. It is also worth noting that at the extensive margins FDIs in downstream business functions yield a positive and significant impact on the probability of entering a new green technological specialization.

In columns (5) and (6), we find the estimations that include the interaction between the green FDI supply dummy and the skills-related variables. The results are in line with the previous estimations, confirming the compensation role of $(d)ASH$ and the boosting role of $(d)RSH$.

**Robustness checks**

As stressed in Section 3.3.1, the volume of capital expenditures for each FDI flow is estimated by FT Intelligence. To account for potential measurement problems, we run further estimations to check for the robustness of our results. First of all, we assume that if any bias is introduced by the estimation procedure, it is supposed to affect all the sampled FDI projects in the same

**Table 3.7:** Green Diversification, skill composition, and Supply vs. Demand Green FDI Capex (dummies)

| | entryGT | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| (d) GreenCapexSupply | 0.1662*** | 0.1069*** | 0.1257*** | 0.0710** | 0.1393*** | -0.0056 |
| | (0.0361) | (0.0338) | (0.0329) | (0.0326) | (0.0447) | (0.0363) |
| (d) GreenCapexDemand | 0.1435*** | 0.0590* | 0.1167*** | 0.0770*** | 0.0838*** | 0.0812*** |
| | (0.0348) | (0.0306) | (0.0282) | (0.0248) | (0.0244) | (0.0244) |
| (d) BrownCapex | 0.2220*** | 0.1733*** | 0.1796*** | 0.1415*** | 0.1430*** | 0.1374*** |
| | (0.0429) | (0.0403) | (0.0396) | (0.0367) | (0.0362) | (0.0363) |
| (d) ASH | 0.2612*** | 0.1784*** | 0.1711*** | 0.1441*** | 0.2245*** | 0.1376*** |
| | (0.0434) | (0.0400) | (0.0381) | (0.0349) | (0.0434) | (0.0343) |
| (d) RSH | 0.0223 | 0.0365 | -0.0105 | -0.0283 | -0.0310 | -0.1496*** |
| | (0.0417) | (0.0384) | (0.0349) | (0.0334) | (0.0329) | (0.0486) |
| TechRel | | 0.0141*** | 0.0145*** | 0.0125*** | 0.0126*** | 0.0126*** |
| | | (0.0009) | (0.0008) | (0.0008) | (0.0008) | (0.0008) |
| GDPpc | | | 0.2927*** | 0.1638* | 0.1749* | 0.1804* |
| | | | (0.0978) | (0.0955) | (0.0946) | (0.0960) |
| EmpGrowth | | | 0.0333 | -0.1246 | -0.1307 | -0.1537 |
| | | | (0.4016) | (0.3888) | (0.3849) | (0.3876) |
| ShPatents | | | -0.1696*** | -0.1637*** | -0.1584*** | -0.1582*** |
| | | | (0.0287) | (0.0295) | (0.0287) | (0.0290) |
| (d) shGreenEst | | | | 0.2855*** | 0.2885*** | 0.2911*** |
| | | | | (0.0834) | (0.0817) | (0.0817) |
| GreenPrevRTA | | | | 0.6731*** | 0.6958*** | 0.6727*** |
| | | | | (0.1791) | (0.1769) | (0.1769) |
| (d) GreenCapexSupply × (d) ASH | | | | | -0.1464*** | |
| | | | | | (0.0495) | |
| (d) GreenCapexSupply × (d) RSH | | | | | | 0.1889*** |
| | | | | | | (0.0516) |
| Observations | 2,102,562 | 2,102,562 | 2,102,562 | 2,102,562 | 2,102,562 | 2,102,562 |
| Log-Likelihood | -108,317.0 | -107,983.4 | -107,840.4 | -107,782.0 | -107,771.7 | -107,764.2 |
| Adjusted Pseudo $R^2$ | 0.06700 | 0.06986 | 0.07106 | 0.07155 | 0.07163 | 0.07169 |
| Year*Tech fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| State fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Clustered (MSA) standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dep var: MSA entry in a green technological specialization. *GreenCapexsupply* is the amount capital expenditures in green FDI in supply (upstream and production) sectors. *GreenCapexdemand* is the amount of capital expenditures in green FDI in demand (downstream) sectors. All FDI variables are expressed as a dummy variable equal to 1 if the amount of capital expenditures is higher than 0, and 0 otherwise. *ASH* and *RSH* are dichotomous indicators equal to one if the share of, respectively, abstract skills or routine skills is above the national median. Explanatory variables are lagged by one year. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models are estimated through fixed effects logit estimators. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the MSA level. *p<0.1; **p<0.05; ***p<0.01

way. Hence, calculating the share between green and total FDIs should clean any possible measurement issue. The results of these estimations are reported in the Appendix in Table B.1.

We find a positive and statistically significant coefficient for $shGreenCapex$ in all specifications, concluding that the relative weight of green FDIs over total FDIs is a positive driver of regional green diversification. Precisely, in the full model (column (4) in Table B.1 in the Appendix), a higher intensity of inward green FDIs is associated with a 17% increase in the likelihood of entering a new green technological specialization. This suggests that local areas that can orient the flow of FDIs towards green economic activities are also more likely to diversify in new green technological specializations. The positive and significant relationship between abstract skills intensity and $entryGT$ is also confirmed. In particular, MSAs endowed with a high intensity of abstract-oriented occupational skills are 17.3% more likely to acquire new green technological specializations (as in the full model reported in columns (4) of Table B.1). Also, the results concerning the interaction between green FDIs and the skills-related variables are coherent with those discussed above. H3a and H3b on the compensating effect of ASH and the boosting effect of RSH, respectively, are supported, too. As for the control variables, we confirm the main findings discussed above. In fact, in all specifications reported in Table B.1 we estimate positive and statistically significant coefficients for the control variables $TechRel$, $shGreenEst$, and $GreenPrevRTA$.

As a further robustness check, we measure green FDIs as the share between the number (count) of green and total FDI projects. The Appendix reports the results in Table B.2. Overall, the results are consistent with the evidence presented so far and support H1, H2, H3a, and H3b. It is worth noting that the magnitudes of the effects of $dhGreenFDI$ and $(d)ASH$ are also aligned in this case. As a final robustness check, Table B.3 in the Appendix reports the results of estimations including a stricter set of fixed effects: $Tech * Year$ fixed effects, where technologies are identified at the 4-digit level and $State * Year$ fixed effects. This causes a drop in the number of observations preserved by the estimation. Yet, the results are fully consistent with those reported and discussed in the previous sections.

## 3.5 Conclusions

This work investigates the role of global connectedness and skill composition for local technological diversification into green domains. Specifically, we hypothesize that inward green FDIs and exploration-oriented skills positively influence the probability that a given locality develops a new green technological specialization. Additionally, we explore the role of the interplay between FDIs and the local skills endowment in these dynamics, stressing the importance of distinguishing between potential *vis-à-vis* realised local absorptive capacity as drivers of compensating or boosting mechanisms. We conduct the empirical analysis on a dataset of 287 US MSAs observed over 2003-2018, employing fixed effects Logit estimators with MSA-clustered standard errors.

The results show that green inward FDIs, measured through capital expenditure, count of projects, shares, or dichotomous variables, are positively related to the probability for MSAs to develop new green-tech specializations, supporting H1 of this work. Likewise, the probability of acquiring a new green specialization is higher for MSAs endowed with above-national median levels of abstract-oriented occupational skills compared to the rest of MSAs, thus supporting H2. Lastly, on the one hand, we estimate a negative interaction between green FDIs and abstract skills, suggesting that a compensation mechanism is at stake between the two drivers. Coherently with H3a, estimated marginal effects show that a high endowment of exploration-oriented skills is crucial for regional green diversification, regardless of the relative weight of green FDIs. We interpret this result as follows: in local labour markets that are characterized by above median levels of abstract-skills intensity, knowledge-search, and exploration-oriented capabilities that speak to *potential absorptive capabilities* prevail. On the other hand, we find robust evidence of a positive interaction between green FDIs and the local endowment of routine skills. In line with H3b, this suggests that the presence of workers employed in production-floor functions is crucial for the development of tacit knowledge through localized learning, which favors the translation of external knowledge into actual innovation, i.e. *realised absorptive capacity.*

The findings of this work are relevant both for the academic discourse and for the policy debate around the drivers of the green transition. As for the former, while the positive association between green FDIs and eco-innovation is corroborated in the academic literature – particularly in the contributions concerned with MNEs eco-innovation performances (e.g. De Marchi et al. 2022; Amendolagine, Hansen et al. 2023) – our results add to the recently emerged stream of studies that investigate the role of non-local knowledge flows for local green technological trajectories (Castellani, Marin et al. 2022b; Corrocher et al. 2024). Specifically, our study shows that global knowledge flows are important not just for specializing in green technologies but also, and importantly, for the process of green technological branching. While specialization simply indicates the acquisition of new technologies, diversification implies that regions acquire novel green specializations, hence proving that a process of branching into the green domain is effectively taking place. We provide evidence that MNEs act as non-local agents of structural change via their green investments, allowing host regions to access global knowledge beyond their geographical boundaries (Neffke et al. 2018b; Boschma 2022b). Through a process of hybridization of local and global green knowledge, the combination of heterogeneous and distant knowledge components increases the probability of developing a new green specialization, which, in turn, will foster local green technological diversification.

Our study also adds to the recent innovation literature investigating the nexus between the change in the local skill structure and the green transition (Consoli, Marin et al. 2016; Vona, Marin, Consoli and Popp 2018; Vona, Marin and Consoli 2019). Yet, while extant research mostly focuses on the labor market impact of the greening of the economy, we delve into the role of local human capital endowment for green technological diversification dynamics, showing that the composition of the local skill set plays a crucial role in influencing the ability to support the generation of green technologies. Specifically, the prominence of abstract skills is of paramount importance as it correlates with the cognitive capabilities necessary to integrate concepts and resources from diverse domains into fresh and unexplored avenues.

The findings of this work are also relevant for policy-makers. On the one hand, diversification into green technologies is undoubtedly a priority for governments across different geographical contexts and at different geographical scales (Mazzucato and Perez 2015; Corrocher et al. 2024). However, diversifying in green technologies depends on several internal and external factors, most notably when referring to local technological patterns. Our evidence on US MSAs is useful to show that the attraction of MNEs' green investments can become a fundamental ingredient of smart specialization policy toolboxes aimed at increasing environmental sustainability. On the other hand, transitioning to green growth encompasses more than just developing and adopting new green technologies. Much of the necessary innovation involves organizational and institutional changes, breaking away from established norms and carrying inherent uncertainties about their effectiveness. Thus, fostering the development and adaptation of human capital becomes a key area for policy intervention. Active labor market policies are vital not only for swiftly reintegrating displaced workers but also for addressing or preventing skill gaps. To ensure a seamless adjustment of labor markets to these demands, concerted efforts are necessary to discern the direct (e.g., market demand) and indirect (e.g., regulatory) impacts of addressing climate change on existing job profiles and the required skill sets for emerging green sectors. Moreover, beyond the quantitative implications, public authorities should assist businesses in generating quality employment opportunities during their transitions to greener practices, thereby aiding local labor market adjustments. From a dynamic perspective, agile, adaptable, and targeted education and training systems are pivotal in laying the groundwork for an equitable shift towards a low-carbon economy. Given the territorial specificity of climate change, local labor market institutions will play a crucial role in balancing national or supranational regulations with incentives to promote sustainable business ventures.

As with any study, ours is not free from limitations, some of which relate to the empirical framework. The latter, as it stands, does not allow us to fully rule out the endogeneity concerns revolving around the potential bi-directional link between green diversification and FDIs. In fact, it is possible that MNEs target regions already specialized in green technologies and contribute to spurring green diversification. Taking into account such considerations is relevant for any conclusion about causality and should be the focus of future research. Similarly, future research avenues aimed at disentangling causality relationships should open the black box of FDIs. In particular, given the well-known heterogeneity of FDIs (Ascani, Balland et al. 2020b), it is important to account for their diverse nature, notably with respect to the stage of the value chain they refer to (e.g. upstream versus downstream). We take a first step in this direction by showing that the mechanisms underlying the role of green FDIs for green technological diversification are likely driven by investment projects clustered in upstream activities, i.e. related to the location of functions mostly focusing on education, training, creativity and knowledge generation.

# Chapter 4

# Climate Policy Uncertainty and Directed Technical Change: evidence from European firms

**Abstract**

In this paper, I derive novel indexes of Climate Policy Uncertainty for four European countries. Exploiting a new dataset of web-scraped newspaper archives and text-as-data techniques, I explore the role of policy stance underlying aggregate indices of CPU, deriving sub-indexes for uncertainty suggesting increasing or decreasing future stringency. Building on the Directed Technical Change literature, I test empirically the relationship between CPU sub-indexes and environmentally sensitive technologies, in a panel of European firms between 1990 and 2020. I find a significant relationship between the direction of firms' technological efforts, proxied by patents, and that of policy uncertainty. The results suggest that policy uncertainty is a relevant factor in affecting the direction of technical change, bearing important implications in terms of both climate and green industrial policy making.

## 4.1 Introduction

Addressing the climate crisis, and achieving the Paris Agreement targets of containing average emissions' increase well below 2°C, requires radical decarbonization of the world economy. The scale and speed of change necessary for achieving climate targets entails major coordination efforts, in which the role of governments in steering market forces is fundamental. The scale of these efforts, in fact, has been described as an industrial revolution against a deadline (Schmitz et al. 2013; Lütkenhorst et al. 2014). The development of carbon-neutral technologies in all sectors of the economy, from transport to energy production, plays an essential part in the tension between decarbonization and economic growth (IEA 2020).

Within the green growth paradigm, the development and production of sustainable technologies open new economic opportunities, while contributing to decouple growth from polluting emissions. The costs required for achieving climate targets and stimulate green innovation must be met timely, in order to avoid further climate damage. Green growth represents an opportunity for the economic success of countries and regions, but is also the source of deep tensions. At its core, climate policy aims at pricing environmental externalities and steering market prices towards making green products and technologies relatively more convenient than polluting ones (Gugler et al. 2024). Transitioning away from a fossil-based model of economic growth is met by resistance of stakeholders of sunset industries. Transitioning away from polluting products and technologies could result in unjust outcomes, by favoring specific economic players, income groups, or territories, and creating new forms of climate-related inequalities (Pegels 2014; Rodríguez-Pose and Bartalucci 2023).

In this landscape, the support for climate and green industrial policies critically depends on the success of the policy mix in delivering a just and effective transition (Altenburg and Rodrik 2017). Climate and green industrial policies have taken a central role in the academic and public debates during the past decade (Wade 2014; Cherif and Hasanov 2019). Strong government intervention is essential in steering economies towards a sustainable growth path, and the implications of this necessity are at the core of the tension between the State and the market (Mazzucato 2011; Rodrik 2014).

The crucial role of green innovation in ensuring both emissions reduction and competitiveness is everyday more relevant (Fankhauser et al. 2013; Aghion, Ahuja et al. 2023). In this sense, directing technical change away from a high-carbon equilibrium towards a low-carbon one, requires a policy mix able to steer economic incentives for innovation in a cleaner direction (Acemoglu, Aghion et al. 2012). The need for strong climate policies has been stressed for decades, and progress has been made, but their implementation has been subject to periods of deceleration and doubt. In light of the green-tech race, and more in general the success of the transition, consensus and clarity around climate and industrial policy-making are of immense importance (Altenburg and Rodrik 2017). Uncertainty in climate policy making has recently been subject to the attention of scholars, as a potential factor slowing down investments and hindering the transition (Basaglia et al. 2021).

As defined by (Baker et al. 2016), Economic Policy Uncertainty (EPU) regards government actions, regulations, and policies that can influence economic and business decisions. Climate policy uncertainty (henceforth, CPU) is a specific subset, focusing on the ambiguity surrounding the design, implementation, or future trajectory of policies aimed at addressing climate change, and achieving the transition. This uncertainty includes unclear government attitudes towards climate regulations, in terms of emissions' targets, carbon pricing mechanisms, or international climate agreements. CPU could in turn affecting the behavior of economic agents, particularly delaying the transition to a low-carbon economy.

In this chapter, I contribute to the empirical literature on directed technical change (DTC), and on the behavior of economic agents facing climate-related uncertainty (Pindyck 2021), in different respects. First, I build on the empirical literature on policy uncertainty started by Baker et al. (2016), and construct new measures of CPU for France, Germany, Italy and Spain. Exploiting text-as-data techniques and Natural Language Processing (NLP) I derive novel sub-indexes of CPU, leaning towards increasing or decreasing stringency, in order to map the direction of uncertainty. Second, I employ semi-supervised machine learning on this data to make this exercise extensible flexibly to other data sources, and test the relevance of text-as-data techniques in policy-uncertainty and environmental economics applications (Dugoua et al. 2022).

Adding to previous empirical exercises in the literature, I explicitly adopt a DTC framework, and study the effects of CPU on the direction of technological change in firms, studying their low-carbon and polluting patenting activity. I run an empirical analysis on a panel of around 4800 European firms between 1990 and 2020 and argue that the direction of policy uncertainty (suggesting increasing or decreasing probability of future policy stringency) affects the belief revision of firms and in turn the direction of innovation.

The remainder of the paper is organized as follows. Section 4.2 presents the relevant literature context and develops hypotheses. Section 4.3 presents the data and the empirical strategy. Section 4.4 presents and interprets the results. Section 4.5 concludes and derives policy implications.

## 4.2 Literature background

### 4.2.1 Policy uncertainty and firms' behavior

In recent years, economists have stressed the importance of endogenous growth in the context of climate change. Starting from Acemoglu, Aghion et al. (2012), numerous studies have investigated the effects of climate and green industrial policies on the direction of technological change (Dechezleprêtre, Martin et al. 2019). A recent body of empirical evidence has confirmed the relevance of DTC frameworks, showing how innovation in climate-relevant technologies is sensitive to economic incentives, and how incentives can be altered by climate policy instruments. (Dechezleprêtre and Hémous 2022) and Hémous and Olsen (2021) provide recent reviews of the theoretical and empirical developments in this area. Climate policies aimed at affecting the relative price of green and dirty goods, such as carbon taxes or green R&D subsidies affect innovation outcomes, directing technical change away from polluting technologies towards low-carbon ones.

In this context, an emerging stream of literature has started investigating the role of policy uncertainty. The study of firms' behavior in response to uncertainty has a long tradition (Bernanke 1983; McDonald and Siegel 1986). More recently, novel forms of climate-related uncertainties are being increasingly recognized as factors affecting the incentives system faced by economic agents (Pindyck 2021). In addition, the increased availability of text data, in the past decade, has paved the way for a flourishing empirical literature on policy uncertainty, building on the initial idea for EPU proposed in Baker et al. (2016). Differently from other measures of market uncertainty, based often on the volatility of stock prices or on econometric measurement, they developed measures based the text of newspaper articles. Both this work and a large number of follow-up studies have shown the negative effects that EPU shocks exert on the economy during periods of high uncertainty about economic policy actions (for a comprehensive review about measurements and effects, see Cascaldi-Garcia et al. (2023)).

EPU indexes are built on a set of keywords able to capture events in which EPU has risen historically, making it possible to operationalize indicators of policy uncertainty across languages and time. Building on similar methodology, an emerging stream of literature has been developing similar indexes for CPU (Gavriilidis 2021; Basaglia et al. 2021; Noailly et al. 2022). Differently from EPU, CPU is built based on a different set of keywords for climate-related newspaper articles, rather than capturing a broad range of articles dealing with the economy (the precise construction of the index is detailed in Section 4.3). CPU aims at quantifying how uncertain the climate-policymaking process is, based on a set of nationally relevant newspapers, which might affect the behavior of economic agents.

At the firm level, the responses to uncertainty in adapting expectations are rooted in real-options theory (Dixit and Pindyck 1994). The effect of CPU on firms' behavior can be understood through two complementary conceptual mechanisms: real-options theory and anticipatory behavior. According to real-options theory, uncertainty about future policies increases the value of delaying investments, particularly when these investments involve high up-front costs or are irreversible (Bernanke 1983; Dixit and Pindyck 1994). For green technologies, which are often capital-intensive and highly dependent on regulatory clarity, this mechanism can be particularly relevant. Uncertainty regarding the timing and stringency of measures such as carbon taxes or green subsidies can lead firms to adopt a wait-and-see approach, postponing investments until greater clarity emerges. This delay can be further exacerbated by the path-dependency of green technologies, where early inertia in the development of polluting technologies can create additional barriers to shifting investment priorities.

On the other hand, anticipatory behavior can drive firms to act preemptively in response to policy uncertainty, accelerating investments to gain a strategic advantage in expected future markets. This mechanism may be particularly relevant for green technologies, given their reliance on government intervention to address market failures and their potential for long-term competitiveness in a transitioning economy (Acemoglu, Aghion et al. 2012) .

While this analysis is built on the notion that firms merely react to CPU, firms may also generate uncertainty by influencing government action by lobbying policy-makers and politicians. While this is highly plausible given the size and relevance of sectors and firms affected by the transition, this avenue of research is beyond the scope of this analysis, and will be further discussed as limitation in my empirical setup, being a source of possible endogeneity. For the scope of the present study, firms are reacting to increasing CPU by a wait-and-see mechanism or anticipatory behaviors, in terms of their technological direction.

Empirically, the net effect of concurrent mechanisms, in the context of climate related risks and uncertainties, is still unclear (Pindyck 2021). In particular, different studies find rather heterogeneous results, depending on the employed measures for policy uncertainty. In the next subsection, I review the extant empirical literature at the crossroads between policy uncertainty and environmental innovations, and develop the hypothesis.

### 4.2.2 Empirical evidence on uncertainty and green innovation

In Table 4.1, I review of recent studies linking policy uncertainty and firm outcomes, from an environmental and green innovation perspective. I consider two different measures for policy uncertainty. First, I review papers from the literature on EPU, including only studies dealing with environmentally-related outcomes, namely green investments and patenting, or greenhouse gases (GHG) emissions. Second, I include exercises employing CPU as the explanatory variable of interest.

**Table 4.1:** Literature Review: Policy uncertainty and innovation

| Paper | Sample | Measure | Countries | Outcome | Frequency | Effects |
|---|---|---|---|---|---|---|
| Bai et al. 2023 | firms, 2011-2020 | CPU | China | Green Patents | yearly | Positive |
| Basaglia et al. 2021 | firms, 1990-2019 | CPU | US | Stock returns, RD, patenting, employment | quarterly | Negative |
| Berestycki et al. 2022 | firms, 1990-2018 | CPU | 12 OECD countries | Investments | yearly | Negative |
| Bettarelli et al. 2023 | countries, 1976-2020 | EPU | 81 countries | Green patents | yearly | Negative |
| Bouri et al. 2022 | firms, 2000-2021 | CPU | US | Stock returns (green vs brown) | monthly | Positive |
| Cui et al. 2023 | firms, 2005-2019 | EPU | China | Green Patents | yearly | Negative |
| Dorsey 2019 | plant, 2002-2011 | CAIR (single policy) | US | Investments and emissions | yearly | Negative |
| Y. Feng and X. Ma 2024 | firms, 2011-2021 | PEU (text-based at firm level) | China | Green Patents | yearly | Negative |

| | | | | | | |
|---|---|---|---|---|---|---|
| G.-F. Feng and Zheng 2022 | countries, 2000-2022 | EPU | 22 countries | Renewable Energy patents | yearly | Positive |
| Gavriilidis 2021 | US, 2000-2021 | CPU | US | CO2 emissions | monthly | Negative |
| Hoang 2022 | firms, 2000-2019 | CPU | US | R&D expenditures | quarterly | Negative |
| Hu et al. 2023 | firms, cross-section | survey-based EnvPU | China | Green investments | yearly | Negative |
| W. Huang 2023 | firms, 1987-2019 | CPU | US | Green Patents | yearly | Negative |
| Khalil and Strobel 2023 | macro and firms, 2000-2019 | CPU | US | Market value, Investments | quarterly | Positive |
| Kyaw 2022 | firms, 2002-2020 | EPU | US | EnvInnovation score | yearly | Positive |
| Xiaoqing Li et al. 2021 | provinces, 2000-2017 | EPU | China | Green Patents | yearly | Negative |
| Noailly et al. 2022 | macro and firms, 1990-2019 | EnvPU | US | Green VC in startups | quarterly | Negative |
| Peng et al. 2023 | provinces, 2000-2017 | EPU | China | Green Patents | yearly | Positive |
| Ren, Shi et al. 2022 | firms, 2009-2020 | CPU | China | Total Factor Productivity | yearly | Negative |

| | | | | | | |
|---|---|---|---|---|---|---|
| M. Wang et al. 2023 | firms, 2000-2020 | CPU | US | CO2 Emissions and Green Patents | yearly | Positive |
| J. B. Wang 2022 | cities and firms, 2003-2019 | Local CPU (instrumented) | China | Green RD, Patents, Employment | yearly | Negative |
| Xu and Z. Yang 2023 | cities, 2005-2016 | EPU | China | Green Patents | yearly | Positive |
| Yu and Chen 2023 | firms, 2007-2020 | EPU | China | Green Patents | yearly | Negative |

Evidence about the effect of EPU and green innovation is far from conclusive. In a recent working paper, analyzing a large sample of countries and sectors, Bettarelli et al. (2023) suggest that an increase in EPU, measuring the general uncertainty about government's economic policy, depresses green patenting. Cui et al. (2023) and Niu et al. (2023) study the effect of EPU on firm-level green patenting in China, also finding a negative relationship. Xiaoqing Li et al. 2021, at the level of Chinese provinces, adds evidence in this direction. Yu and Chen (2023) and Hu et al. (2023) also report a negative association between EPU and green patenting in China, at the firm level. On the contrary, in the United States, Kyaw (2022) and M. Wang et al. (2023) find a positive effect on measures of green innovation, including investments, patents, and survey-based eco-innovation measures. Xu and Z. Yang (2023) and Peng et al. (2023) finds similar results at the provincial level in China. At the country level, G.-F. Feng and Zheng (2022) adds to this positive relationship.

EPU is a measure capturing general aspects of economic policy, including monetary policy shocks, terrorist attacks, trade shocks or electoral uncertainty. Environmentally-related technologies might be more sensitive to a general uncertainty shocks compared to other technologies (Bettarelli et al. 2023), because of the different nature of green technologies in terms of risk, complexity, or their need for government support.

Green technologies might be particularly sensitive to policy uncertainty due to their unique characteristics, entailing a higher complexity in their recombinant properties (Barbieri, Marzucchi et al. 2020; Fusillo 2023). Unlike other technologies, green innovations address externalities that markets fail to price adequately, requiring consistent government support through carbon pricing, renewable energy incentives, and green R&D subsidies (Acemoglu, Aghion et al. 2012). These technologies also involve higher financial and technological risks. Long development cycles and reliance on novel, cross-disciplinary knowledge create significant uncertainty for firms, especially when policy environments are unpredictable. Delays or reversals in key regulations, such as carbon taxes, can therefore devalue investments. Uncertainty could further amplify their exposure to fragmented or inconsistent international policies, disrupting innovation ecosystems and the spatial diffusion of these technologies (Losacker, Horbach et al. 2023).

Green investments and technologies are therefore arguably more sensitive to uncertainty, and in particular to that specifically bound to climate and environmental policies. An emerging stream of empirical studies, on which this analysis builds, quantifying this type of uncertainty based on the methodology put forward by (Baker et al. 2016). Gavriilidis (2021) measured CPU, based on a sample of nationally-relevant newspapers in the United States, finding a negative relationship with emissions' reduction in a sample of firms. Many studies relate to the effect of CPU rises in the United States, employing the index constructed by Gavriilidis (2021), and observing its relationship with firm level outcomes. Most of these studies focus on the effect of rising

CPU in US and Chinese firms. Since then, a number of different exercises have developed alternative CPU indexes.

Noailly et al. (2022) develops a similar index for environmental policy uncertainty (EnvPU) employing text-as-data techniques, and testing its effect on venture capital funding for US startups. Using quarterly data, they find that an increase in Environmental Policy Uncertainty is associated to lower amounts of capital raised by clean-tech startups. Other recent exercises have built CPU indexes for a larger number of countries. Basaglia et al. (2021) and Berestycki et al. (2022) are the most connected to this paper, measuring and studying the impact of CPU respectively in the United States (Basaglia et al. 2021), and on OECD countries, exploiting a global firm-level datasets (Berestycki et al. 2022). Both studies find a reduction in investments and firm level performances, and Basaglia et al. (2021) also explicitly measures the direction of uncertainty that underlies variation in CPU in English-speaking countries. They use a keywords-based approach to distinguish newspaper articles pointing towards more or less stringent regulation. By interacting emissions intensities as a form of exposure to climate policies, they show how US-firms are more sensitive to variation in CPU that is pointing towards more rigid regulation. While they test results for R&D expenditures, share price volatility, and other outcomes in the US, they do not look explicitly into the direction of effects in terms of green-vs-dirty patenting.

Again for the context of the United States, J. B. Wang (2022) measures

CPU differently, exploiting the volatility in votes regarding climate legislation, finding that firms adopt an anticipatory behavior with respect to innovation and adoption of climate technologies. Adding to the US evidence, Hoang (2022) distinguishes between low and high-emitting firms, finding a negative effect for the latter, suggesting that heavy-emitting firms might be adopting a wait-and-see investment strategy. Two studies explicitly look at the relative performance of green and dirty outcomes in response to climate policies. At the macro level, Khalil and Strobel (2023) employ both a general equilibrium models, and granular firm level data for the US. They find evidence of capital reallocation towards cleaner assets with respect to more polluting ones, while lowering investments in carbon intensive industries and increasing it in "greener" firms. These findings are in line with Bouri et al. (2022), finding a positive role for US-CPU on the relative performance of green energy stocks vis-à-vis brown counterparts. In a study precedent to the empirical literature based on newspaper data, Dorsey (2019) exploits a quasi-experimental framework relating to a single climate policy measure. He finds that firms exposed to a higher level of CPU reduced investments and experience a lower reduction in emissions.

Outside of the US, a number of empirical studies have investigated the CPU and firms' performance in China. Ren, X. Zhang et al. (2022) find a negative effect of US-CPU on total factor productivity in a sample of Chinese firms, channeled through a reduction in R&D expenditures and cash flows, with results varying according to the institutional ownership of firms. Bai et al. (2023) test the US-based index developed in Gavriilidis (2021) on a sample of listed Chinese firms, finding a positive relationship with green patenting. Similarly, Ren, Shi et al. (2022) find a strong non-linear correlation between CPU and investments, negative in polluting industries and positive for green-related investments. Differently from other studies based on newspaper data, Hu et al. (2023) uses survey-based measures of policy uncertainty (policy content and policy enforcement), at the local level, and find a negative relationship with green patenting in Chinese firms. Other papers focus instead on Chinese CPU. Recently, Y.-R. Ma et al. (2023) employed deep learning techniques to build indexes with geographical variation of CPU in China. Y. Feng and X. Ma (2024) also use text analysis techniques on company reports in order to build a measure of environmental uncertainty perceived by the firms, finding that it might hinder green innovation. At the city-level, M. Wang et al. (2023) finds that a reduction in Chinese green R&D, patents and employment, due to uncertainty specifically constructed around the allocation system of subsidies allocated by the central government.

In summary, while in the case of the EPU index, capturing a more general aspect of policy uncertainty related to economic policies, uncertainty can be expected to be detrimental to any innovation process (Basaglia et al. 2021), depressing general investment activity, the empirical evidence seems to show more mixed results. However, in the case of CPU, the same effect cannot be expected a-priori, as the underlying signals in the climate-policymaking process could be differently affecting environmentally-opposing technologies.

### 4.2.3 CPU's direction and technological change

A number of gaps emerge from reviewing the literature on EPU and CPU's effects. First, in terms of technological dynamics, evidence beyond general investments is still scant. Studies focusing on green patenting do not explicitly adopt a directed technical change perspective, controlling for the strong path-dependencies characterizing environmentally-sensitive technologies (Acemoglu, Aghion et al. 2012; Aghion, Dechezleprêtre et al. 2016).

If CPU is differently affecting green and brown investments (or sectors) it is plausible to think that climate-sensitive technologies (green or dirty) would also be affected in different ways. Khalil and Strobel (2023), from a macro perspective in a general equilibrium framework, find evidence for a mechanism of capital reallocation, with investments shifting from brown towards cleaner sectors.

Bouri et al. (2022) add evidence on the positive effect of CPU on the relative performance of green vis-à-vis brown financial stocks. I add to this evidence focusing on technological dynamics. CPU could affect firms' behaviors in terms of future costs and values of the technologies. In an environmentally-positive direction, CPU could rise, for example due to discussion about the implementation of a carbon tax. This would directly affect the (expected) cost of capital for polluting technologies (Khalil and Strobel 2023), and indirectly the future value of clean-tech alternatives, causing a shift in investments efforts from dirty technologies to green technologies.

From this perspective, policy direction within uncertainty indexes becomes central in the expectation-revision of firms, driving investments towards two alternative technologies. As discussed in depth in the next section, CPU indices are built on a set of environmental, policy, and uncertainty keywords. They capture both directions of the climate policy-making process, aggregating both milestones and setbacks. As mentioned, this aspects could be crucial in terms of firms' expectations and technological trajectories. With the exception of Basaglia et al. (2021) and Berestycki et al. (2022), most of these studies do not unpack CPU indices, and consider aggregate CPU. Building on their work, I add nuance in terms of the direction of policy-uncertainty. If policy uncertainty points towards a more stringent environmental regulation, firms might be inclined to accelerate innovation in terms of environmental technologies, and divest from fossil-based technologies. A symmetric behavior could be expected in terms of polluting technologies. If CPU increases are driven by setbacks in the climate policy-making process, firms could continue, or accelerate, investments into polluting technologies. Therefore, rather than its aggregate level, a driving factor for innovation could be the underlying variation in "good" or "bad" news for the environment, with uncertainty indicating a higher probability of future regulation that could affect both the costs and returns from alternative technologies.

An increase in positive policy uncertainty would enter the production function of profit-maximizing agents as a (potential) extra cost for the production of environmentally-damaging technologies. A decrease in the probability of a carbon tax, for example, could represent a (relatively) higher expected cost for the development of green technologies vis-à-vis polluting ones.

Hence, uncertainty caused by setbacks in climate policy-making, could cause firms to continue investing in fossil-based technologies. Furthermore, given the strong path-dependency (Aghion, Dechezleprêtre et al. 2016), an increase (or a non-decrease) in future value of polluting technologies could be detrimental for development of low-carbon alternative, and incentivize agents to continue developing polluting technologies. In an economic equilibrium which is already favoring polluting technologies, uncertainty could therefore be a significant factor in steering change towards a cleaner path (Acemoglu, Aghion et al. 2012). By the same logic, an increase in the probability of green subsidies, could decrease the expected costs of firms in developing green technologies with respect to fossil-based ones, therefore incentivizing the former and discouraging the latter. In line with the evidence on green industrial policy (Pegels 2014), clarity and commitment of legislators around policies is crucial in this sense, as CPU could be an underlying factor altering expectations and investments into alternative technologies. In this chapter, I hypothesize, that different signals underlying CPU matter for the direction of technical change:

> *Hypothesis 1*: Climate Policy Uncertainty affects the direction of technological change in firms, depending on the underlying changes in the

probability of a more stringent environmental regulation.

To the best of my knowledge, no study has yet tested the relationship between (sub) indices of CPU and DTC by considering both low and high-carbon technologies. Furthermore, many of the cited studies using green patents as an outcome variable do not explicitly account for the path-dependency in climate-sensitive technologies, adopting a framework of directed technical change.

The motivation for this study, therefore, stems from the necessity to understand how support for green technologies demands stable, harmonized, and long-term policy frameworks. Without this consistency, the high risks and systemic requirements of green technologies will continue to limit innovation and delay the transition to a sustainable economy. This is particularly relevant in the context of the unpriced externalities and path-dependencies emerging when considering fossil technologies alongside green ones.

I add another two contributions with respect to the extant literature. First, I bring evidence for the European context, while most of the studies reviewed bring forward evidence regarding the US and China. Second, I contribute to the emerging literature applying text-as-data techniques in environmental economics and policy (Dugoua et al. 2022), adopting a novel approach to measure the policy stance of news articles.

## 4.3 Empirical Framework

### 4.3.1 Data

#### Newspapers' archives and CPU

In order to build European indices of CPU, I collect millions of full text articles by means of web scraping. Web scraping automates the collection of the content of web pages. I built scrapers for several newspapers archives across my sample countries (Germany, France, Italy and Spain). I focus on multiple archives for each country, as common in the policy uncertainty literature to smooth effects due to the structure of a single outlet.

I collect nationally-relevant archives, although the selection of sources by each country was limited by the availability of digitized news archives and the feasibility of the scraping process. I target outlets with different political leaning, in order to balance reporting biases, which is important for the reliability of this measures, although numerous normalization steps are performed in line with the literature. The selection of newspapers for this paper largely resembles that of similar exercises in the literature (Basaglia et al. 2021).

With the exception of Germany, for which I focus on the weekly outlets Der Spiegel and Die Zeit, all remaining news sources detailed in Table 4.2 have a daily frequency. For France, I collect data for Le Monde and Figaro. In Italy for La Stampa, La Repubblica, Il Foglio and il Sole 24 Ore. For

Spain, I collect data on El Pais and El Mundo newspaper archives. The resulting dataset allows me to exploit around 17 million full-text newspaper articles, spanning the period 1990-2020.

The data I collect via web scraping include the date on which the article was published, its title, and the body of text. No reliable information about the relevance of the article within that day's newspaper (or within the website) was available. I clean the collected data from duplicates (based on the webpage's URL, its unique identifier). Furthermore I remove near-duplicate texts belonging to different URLs, often resulting from the process of digitization of newspaper scans for articles belonging to the physical editions.

I do not distinguish, in this database, between digitized articles which originally appeared in the paper version, and digitally native articles that gained importance since the early 2000s. While this distinction could help understand the structure of newspapers archives, for most news sources it is not possible to identify the origin of the news article. Thus, I consider the archives available online as a single entity, blending digital and physical news. A more detailed exploration of newspaper data, or a structured digitization of raw scans[1] could have relevant implications for policy uncertainty indicators, but is beyond the scope of this paper. The span and richness of the dataset collected, allows me to build CPU indicators, for four European countries, and a longer time span than the one considered in previous exer-

[1]For a recent example see Dell et al. (2024)

cises.

Following the methodology proposed by Baker et al. (2016) and subsequent work, I match newspaper articles employing three sets of keywords, and consider an article expressing CPU if it matches all three sets. The first set contains climate-related words (e.g. climate, environment, CO2), for each language. The second set of keywords, instead, matches articles referring to policy issues (e.g. government, policy), including policy-specific terms where relevant (for example ETS - Emissions Trading System). I build on, and expand, the sets of climate and policy keywords adopted in previous exercises (Gavriilidis 2021; Basaglia et al. 2021; Berestycki et al. 2022).

The most important difference, compared to the extant literature, is in the set of keywords expressing uncertainty. The majority of exercises in climate policy uncertainty only match articles based on the keywords "uncertain" or "uncertainty". This selection of keywords has been subject to criticism. Tobback et al. (2018), in the case of EPU, employs a wider set of keywords more generally expressing uncertainty (e.g. doubt, maybe, perhaps). They borrow from the concept of modality in linguistics: modality relates to different ways of expressing degrees of doubts and certainties. Tobback et al. (2018) show that this approach is preferable to simple matches of the words "uncertain" and "uncertainty".

Adapting from this work, I compile a similar list of keywords representing uncertainty. Employing a larger set of keywords helps minimizing false negatives from the sample of articles, by matching a larger number of articles compared to the more restrictive approach. Given this much larger set includes very common keywords, I minimize false positives by only considering articles with a number of modality-expressing words in the top 15th percentile, replicating the approach proposed by Tobback et al. (2018). I match and calculate percentiles separately for each newspaper archive.

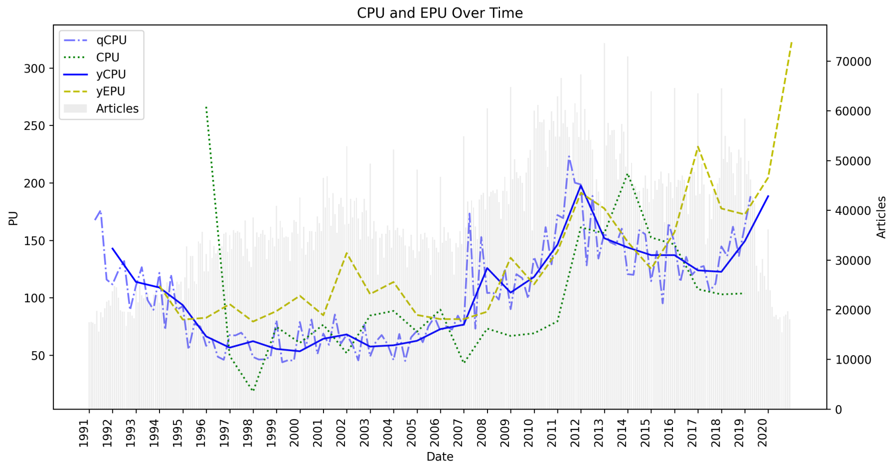
After identifying articles matching the three sets of keywords, I derive monthly time series, for each newspaper, dividing the monthly count of CPU articles by the total amount of articles. In turn, following standard practice in this literature, I normalize the time series by standard deviation, and multiply it for their mean. This normalization helps to remove newspaper-specific factors, due for example to the structure of the newspaper archive.[2] I take averages between newspapers (in the overlapping periods) and multiply the series by 100, making my measures comparable to other policy uncertainty time series.[3]

[2]I calculate standard deviations and means based on periods of consistent number of articles in the archives. I follow Baker et al. (2016), for each newspaper in common in the sample, in defining breaking periods for calculating standard deviations. I also calculate different standard deviations in periods where the total number of articles in the archive shows structural breaks. This might indicate a change in the format of the outlet or in the total amount of digitized news, and could add noise to the measurement.

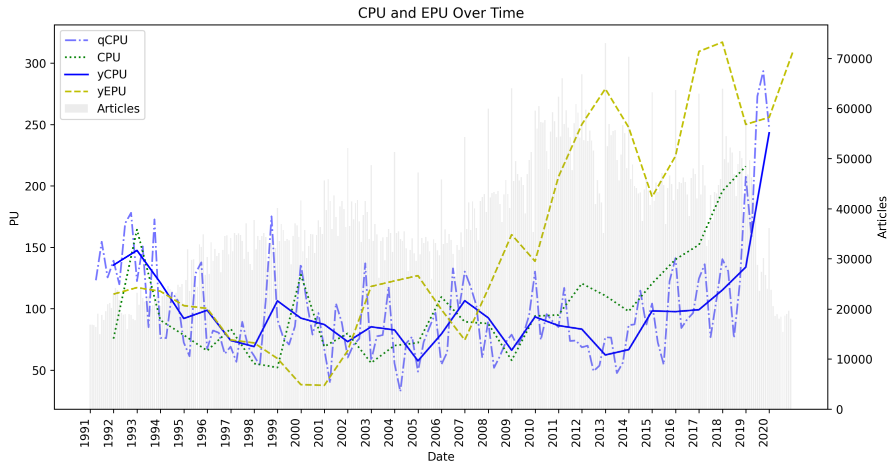
[3]Several series from different exercises are updated and available at: https://policyuncertainty.com

**Table 4.2:** Database of newspaper archives

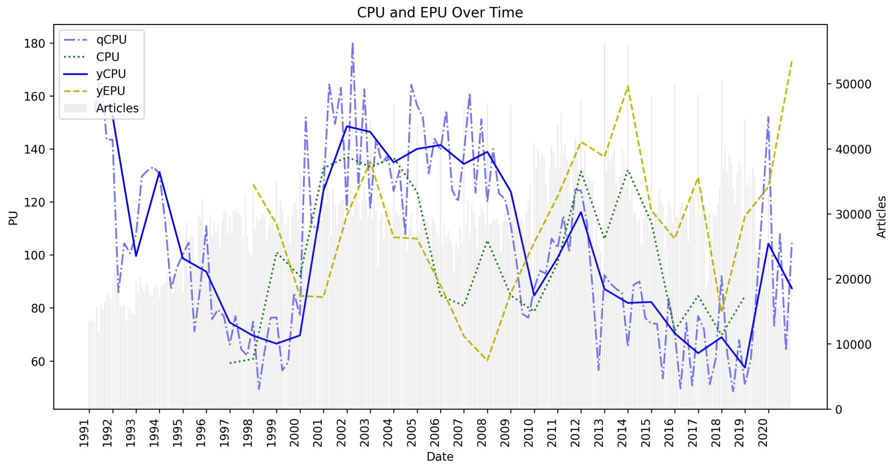
| **Country** | **Archive** | **Articles** |
|---|---|---|
| Germany | Der Spiegel | 307,103 |
| | Die Zeit | 220,497 |
| France | Figaro | 1,773,778 |
| | Le Monde | 1,491,681 |
| Italy | Il Foglio | 53,541 |
| | La Stampa | 5,813,893 |
| | La Repubblica | 4,528,482 |
| | Il Sole 24 Ore | 149,839 |
| Spain | El Mundo | 421,725 |
| | El Pais | 2,490,156 |
| Total | | 17,250,695 |



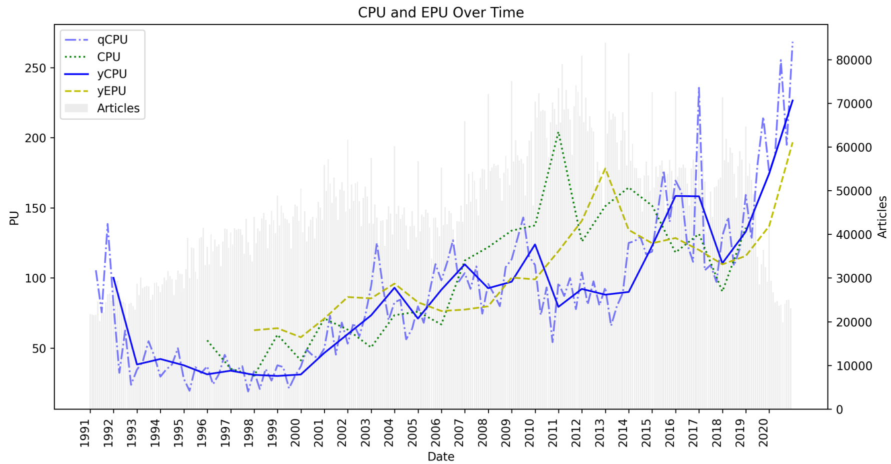

**(a)** Germany




**(b)** France




**(c)** Italy




**(d)** Spain

**Figure 4.1:** Policy uncertainty indexes for sample countries

In Figure 4.1, I plot these results for the four countries. In blue, I represent the (yearly and quarterly) averages for the CPU index built with the procedure illustrated above. I compare it with a similar index developed by the OECD (Berestycki et al. 2022), as well and the EPU series available from Baker et al. (2016). While it correlates quite highly, in yearly aggregation, with the OECD's CPU index, there are notable differences, most likely due to the different methodologies in keywords matching, and a different selection of newspaper archives. Importantly, the CPU index differs from the Economic Policy Uncertainty.

Interestingly, all indices seem to be spiking around 1992-1993 (years of the discussions around the Kyoto Protocol). Also, spikes in the index correspond to the passing of climate legislation in 2007-2008, when during France's EU presidency, the "Climate and Energy Package" was discussed and adopted, fixing climate targets for 2020. In Germany, the index spikes around 2011, during the discussions on Energiewende, the comprehensive climate-policy agenda for the energy transition, featuring a 60% Greenhouse Gases (GHG) reduction before mid-century. For France, in the early 2000s, some spikes relate to the Climate Act (2001), as well as to the 2003 heatwave (peaking also in Italy and Spain).

In Italy, the index first spikes in the middle of the 1990s, at the beginning of the debate on energy market liberalization, began with the 1996 EU directive, implemented in Italy with the 1999 Bersani Law. Another strong rise happens in Italy, in the early 2000s, during the implementation of the reforms and the privatization of energy markets. Another peak in 2011 refers to peaks in solar panel subsidies.

In more recent years, following 2017, CPU seems to be rising across all four countries. The rising trend starting from 2016, could be due to country-specific factors (the Gilet Jeunes protests following a fuel-tax rise in France) or global discussions. As mentioned, in aggregate, all these indices are summing up different types of events. Both President Trump's decision to exit the Paris Agreement (2017), and the Government's responses to the Climate Strikes began by Gretha Thunberg (2018), or the passing of environmental protection laws could be contributing to the rise of the index. [4]

As noted in Basaglia et al. (2021), while one could expect that a rise in general EPU to be slowing down firms' investments and in general economic activity, and even in green patenting (see Bettarelli et al. (2023) among others), it is not necessarily the same with CPU. As mentioned, the direction of CPU seems particularly relevant for the direction of innovation.

CPU might be pointing in two different directions: at a strengthening or a weakening of future climate stringency, for example suggesting further implementation, or a slowdown in the policymaking process. Previous attempts

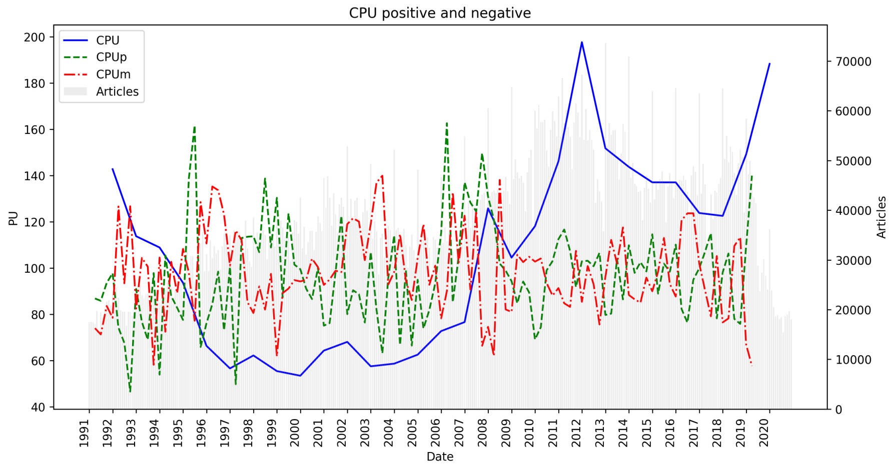
[4] In addition, some news articles refer to specific place-based policies (oil leakages, or polluting plants) which might not necessarily have national relevance. The distinction between local or nationally relevant events is an interesting avenue for further research, but beyond the scope of this paper.

(Berestycki et al. 2022; Basaglia et al. 2021) have mapped the direction of CPU, creating sub-indexes for CPU+ and CPU- (increasing or hindering climate policies) for English-speaking countries, based on sets of keywords capturing the direction of uncertainty. Previous measurement exercises in policy uncertainty have employed human labellers for the validation of these indexes, in particular to filter out false positives. I take a different approach, exploiting full-text data. Recent literature has shown the potential of Large Language Models (LLMs) for accurate annotation of textual data, even over-performing crowd labeling (Gilardi et al. 2023). LLMs, in fact, are opening the possibility to lower dramatically the cost of labeling while still achieving human-level accuracy on a variety of different tasks (for a recent review and application, see X. Wang et al. (2024)).

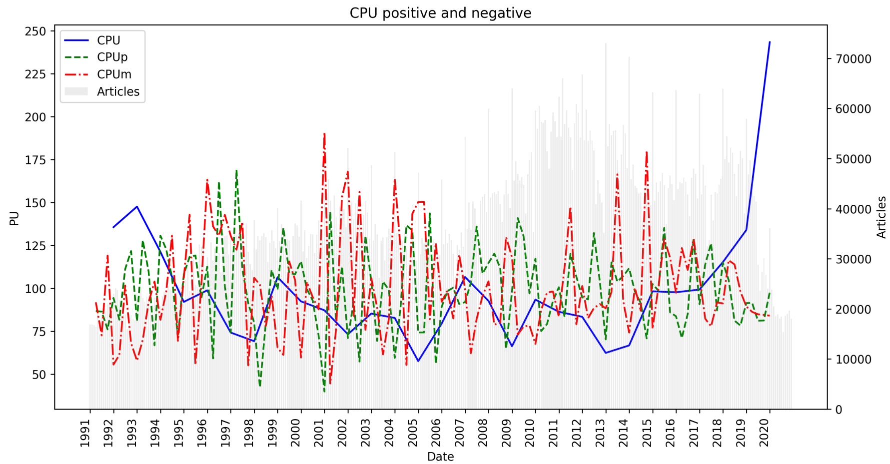
Building on this recent computational social science literature, I prompt the OpenAI API for labeling the universe of CPU news. I make use of the most recent ChatGPT-4o model.[5] I ask three questions during the labeling process. The first serves in filtering out further false positives resulting from the keywords matching: *"Is this news article about climate policy issues?"*. The second two questions are asked to collect information about the policy stance of news articles, and build indicators of positive or negative CPU. The first question is *"Does this piece of news imply a strengthening or weakening of climate policy?"*, and the second *"Are the consequences of this news positive or negative for the environment?"*. The first question is forced as a

[5]At the time of writing, ChatGPT-4o is the largest LLM model available by number of parameters, estimated at around 1.5 trillion.

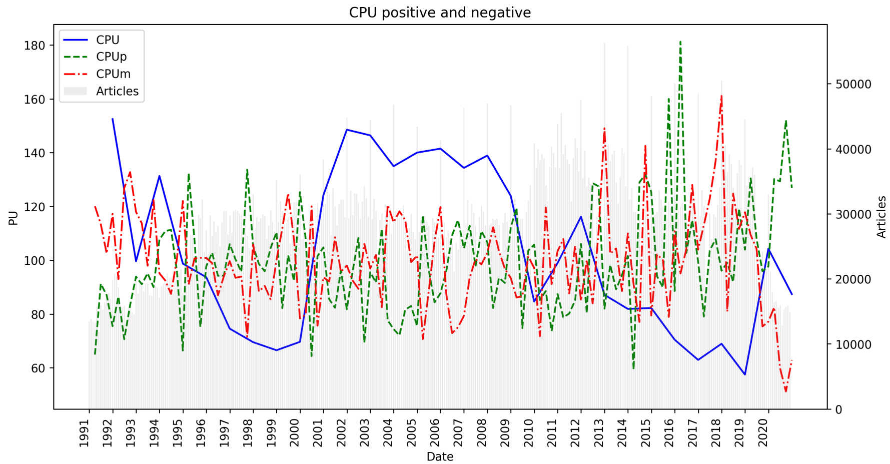
binary answer, while I leave the possibility, for the latter two questions, to be answered with negative, positive or neutral labels. I perform validation and experimentation of the results of the prompting on a random sample of CPU articles, in similar fashion to the procedure explained in Berestycki et al. (2022) for keywords selection.



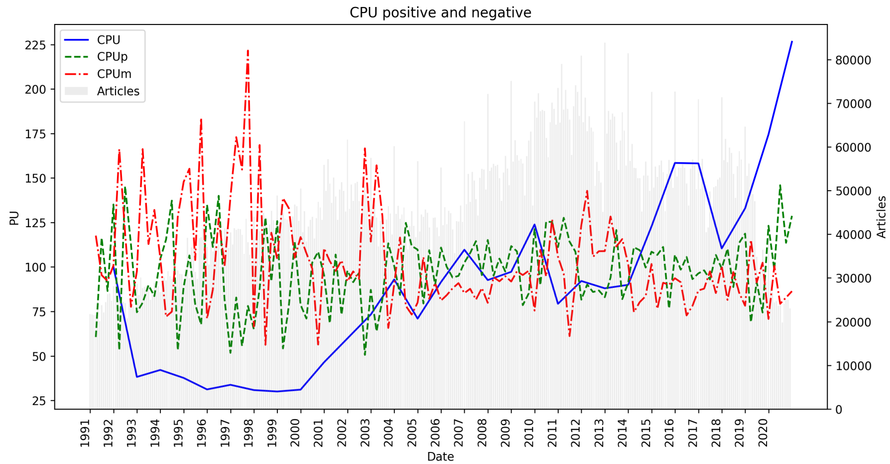

**(a)** Germany




**(b)** France




**(c)** Italy




**(d)** Spain

**Figure 4.2:** Policy uncertainty and policy stance indexes for sample countries

I select as $CPUp$ (CPU plus), articles flagged as true positives during the labeling procedures, and for which any of the two second answers reflect a strengthening of climate policies. Similarly, I flag $CPUm$ (CPU minus), or articles pointing towards a weakening of climate policies or with negative consequences for the environment. Based on this strategy, I derive novel time series, for the four European countries. Rather than dividing the number of articles by the total monthly number of articles in the archive, I divide $CPUp$ and $CPUm$ number of articles by the total number of CPU articles, in order for the series to reflect the relative importance of $CPUp$ or $CPUm$ rather than a general increase of CPU. In Figure 4.2 I plot the time series for $CPUm$ and $CPUp$, in quarterly moving averages, showing significant variation both over time and across countries.

There are several advantages and disadvantages to the use of LLM technologies for the labeling of articles. Compared to keywords, the labeling process is more of a black-box, while the former is fully reproducible. However, the selection of keywords can be subject to biases and discretionary selections. The multilingual capabilities of LLMs, render this approach well-suited for this dataset, which features four non-English languages. In addition, this flexibility extends in time, mitigating the possible recency bias. Journalism and the use of language changed throughout time, and an approach based solely on frequency might be biased towards more recent policy discussions. Finally, the complexity of syntax-aware methods is particularly important in terms of mapping policy stance. While keyword-dictionaries are based on the simple occurrence of words within documents, LLM architectures are syntax-aware, and better able to capture false positives, handling common issues such as negation.[6]

In the Appendix, I provide a sample of articles' titles matched as $CPUp$ and $CPUm$. Interestingly, while belonging to the domain of climate issues, it is clear that other components are still at play, which might be affecting firms' behaviors differently. Many articles are correctly captured under the correct category: President Bush's government agenda for liberalization, opinion pieces on the risks of environmentalism ($CPUm$) or policy announcements about an increase in kerosene tax ($CPUp$). Other news are of local nature (articles about local smog levels or waste management). In this sense, a promising avenue for further research on the measurement and validation of policy uncertainty measures, could be mixing LLM-labeling methods with unsupervised learning to unpack the universe of news into topics, as proposed in the case of EPU by Larsen (2021).

A number of issues remain open with this approach, and will be further discussed in the limitations section. First, while the results seem promising, false positives and noise still affect in the measurement. A more formal validation of the sub-indexes and the labeling performance remains necessary. The performance of LLMs in comparison with human judgment, in social science applications, requires validation, which is not currently implemented for the

[6]For example, the sentence "innovation subsidies will not slow down the climate transition", would be captured by the terms "slow down" in a dictionary-based approach.

sake of this analysis, given the need for human-annotated datasets (Pangakis, Wolken and Fasching 2023; Törnberg 2024). In particular, the prompts I have created to label the news could be further refined by formalizing a validation method for LLM labeling, given a rather cognitively complex task. In future work, I plan on formalizing and validating the prompt form that more accurately can capture the direction of CPU, limiting its subjectivity and increasing accuracy (Juroš et al. 2024).

Nevertheless, this approach seems particularly promising for social sciences, given the large amount of news-data sources now employed in economics, and its potential to develop teacher-student architectures in machine learning applications. In this architectures, the LLM-generated labels are used as training inputs for training smaller text-models (Pangakis and Wolken 2024). I apply this architecture to my dataset and test the reproducibility of artificially generated labels with smaller text models. I discuss its potential merits in the case of CPU indices, and provide benchmark evaluation metrics in the Appendix.

Finally, in order to derive a set of robustness indicators, I isolate events (peaks) in both of the CPU series. I build a peak detection algorithm based on the rolling mean of the monthly time series. This algorithm detects peaks based on the deviation of future data points from a rolling mean of the series. I run the peak detection based a six-months moving average, built for each time series, with a threshold of two standard deviations. Thus, a peak is detected if the new data-points exceed two standard deviations from the rolling mean calculated on the past data points. For the implementation of the algorithm I follow the approached proposed in Brakel (2014), where new peaks detected influence the series. In the Appendix Figure C.1 I show an example of the peak detection process. All indicators are included in the econometric analysis as yearly moving averages.

**Firms and patents dataset**

In line with the literature on the economics of innovation, I make use of patent data to proxy the technological efforts of firms. The use of patent data as a proxy for innovation has a long tradition. Despite the numerous criticisms, patent databases represent a valuable source of information for firms' technological efforts, and have been shown to map effectively the knowledge generated by firms, regions and countries.

I use the OECD's REGPAT database (Maraut et al. 2008) for deriving patent-based indicators. Patents widely vary in quality, and can be filed into different jurisdictions at different patent offices. In addition, firms can file several patents to protect the same invention. In order to avoid double-counting patents for the same technology, and to focus on high-quality patents, I make use of the Triadic Patent Families (TPF) database in REGPAT. Within triadic patent families, an invention is filed under the three major patent offices in the world: the European Patent Office (EPO), the United States Patent and Trademark Office (USPTO) and the Japanese Patent office (JPO). Making use of REGPAT, I construct a dataset focusing on patent families rather than single patents, following the approach in Aghion, Dechezleprêtre et al. (2016). Patenting inventions is a costly process for firms, and more valuable innovations with promising market perspectives are filed in all three offices. Hence, in this work I focus on high-quality inventions.

Patents can be filed under a large number of technological classes, defined under the International Patent Classification (IPC) and under the Cooperative Patent Classification (CPC). I classify technologies under three categories. In line with the eco-innovation literature, I consider green patents technologies that have potential for mitigation or adaptation of climate change. I match patents based on the methodology recently proposed in Favot et al. (2023), building on previous work (Ghisetti and Quatraro 2017). I match codes at different digits based on the OECD's ENVTECH classification (Haščič and Migotto 2015a) and the algorithm proposed by Favot et al. (2023) on both IPC and CPC codes for patent families. I expand this search by manually adding codes at higher level from the Y02/Y04S technological classification. The number of TPFs identified as green technologies represent roughly 9% of total patent families (in line with the results in Favot et al. (2023)). I provide a detailed summary of the codes employed in Table C.5 of the Appendix.[7]

[7]Where a description of sub-codes' purpose is provided, I match codes at a lower depth than the 4-digits macro group indicated, only considering a subset of those technologies. For a more detailed description of the codes employed, please refer to Favot et al. (2023) and Haščič and Migotto (2015a) and the most recent OECD's ENVTECH search strategy.

In turn, I identify "dirty" patents families, matching polluting inventions, linked to the the emission of greenhouse gases. I consider as dirty patent families for the production of fossil-fuel, combustion engines, electricity production from non-renewable sources, in addition to steam and gas technologies. Again, I adapt previous work from Aghion, Dechezleprêtre et al. (2016) and Dechezleprêtre and Sato (2017) for polluting technologies, and expand it with a recent classification of fossil technological codes provided by the International Energy Agency (IEA).[8] Table C.7 provides a full description of the technologies considered as dirty. Finally, I create a sub-category of dirty technologies: grey patents. Grey technologies render combustion processes more efficient, and have potential of reducing GHG emissions, while still being polluting technologies. Once again, I follow previous work (Dechezleprêtre and Sato 2017) and provide a breakdown of grey technological codes in Table C.6.

I aggregate total, green, dirty and grey patent families as counts by applicant. Following Aghion, Dechezleprêtre et al. (2016) I only consider applicants with consecutive observations. Using information on the name and country of applicants in REGPAT's TPF database, I match firms in the ORBIS (Bureau van Dijk) database, exploiting the name-search engine provided by ORBIS, for the applicants having at least one green or dirty patent over the sample period, and with their address in Germany, France, Spain or Italy. Using balance sheet information from ORBIS, I map firms to their main sectors, and collect information on the year of foundation, and the first available balance

[8] I adapted to REGPAT the search strategies for fossil-fuel patents available online from the IEA. For full details, see:https://gitlab.com/iea$_{d}ds_{p}ublic/iea_{p}atstat/$

sheet. I build an unbalanced panel dataset for firms in the four countries. I consider the beginning of the panel the year of foundation where available, and if not the first year of available balance-sheet information. Where the information is not available, I consider as a starting year the year prior to that of the first patent application recorded in REGPAT.[9] I consider as the year of invention of each patent family the earliest filing among the patents belonging to that family.

Table 4.3 describes the number of firms and patents by country, breaking down patent counts for each technological category. In order to control for the path-dependency of the innovation process, again borrowing from Aghion, Dechezleprêtre et al. (2016) I construct several variables for stocks of previous inventions in dirty and clean technologies, and for geographical spillovers of knowledge available to the focal firm, as detailed in the next Section. In addition, to build control variables, I collect data from Eurostat and the OECD to construct sectoral measures of emissions intensity, following Berestycki et al. (2022). I use data on emissions intensities based on environmentally-extended input-output tables. Emissions intensities are defined as GHG emissions embodied in final demand, from Yamano and Guilhoto (2020), normalized by unit of output.[10] Finally, I collect country-level data for the OECD's Environmental Policy Stringency Index (Botta and

[9]Additionally, I correct for discrepancies between firm and patent data. I consider as the starting year the one prior to the first application filing, for firms in which balance sheet information is available, or recorded only after the filing of the first patent.

[10]I consider a number of alternatives for sector emissions, including CO2 intensity per unit of value added, available from the IEA.

Koźluk 2014).

### 4.3.2 Methodology

To investigate the relationship between CPU and the innovation dynamics in firms, I closely follow and adapt the model proposed by Aghion, Dechezleprêtre et al. (2016). I test the hypothesis making use of two symmetric models. In the first model, I regress the count of green patent, by year and firm, against Climate Policy Uncertainty and a set of controls:

$$\begin{aligned} PAT_{i,t} = \exp(\alpha + \beta_2 CPU_{i,t-3} + \beta_3 EPS_{c,t-1} * GHG_{c,s,t-1} \\ + \beta_4 K_{i,t-1}) + \eta_i + \tau_{c,t} + \psi_{s,t} + \epsilon_{i,t} \end{aligned} \tag{4.1}$$

where:

- $K_{i,t}$ is the firm's past patent stock;
- $\tau_{c,t}$ are country by year fixed effects;
- $\psi_{s,t}$ are sector by year fixed effects;
- $\eta_{i,t}$ are firm fixed effects;
- $\epsilon_{i,t}$ is the idiosyncratic error term.

Patent flows are built for three sets of technologies: green, dirty and grey, as count variables by firm and year. I construct the exposure to CPU for the focal firm $i$, similarly to how Aghion, Dechezleprêtre et al. (2016) construct their variable for fuel prices, reflecting the importance of country $c$ for firm $i$

**Table 4.3:** Sample of Firms by country

| **Country** | **Firms** | **Patents** | **Green** | **Dirty** | **Grey** |
|---|---|---|---|---|---|
| Germany | 2780 | 163089 | 13982 | 12707 | 3619 |
| Spain | 192 | 4000 | 327 | 200 | 27 |
| France | 1112 | 63740 | 5142 | 5285 | 527 |
| Italy | 716 | 17506 | 1310 | 1417 | 262 |
| Total | 4800 | 248335 | 20761 | 19609 | 4435 |

**Table 4.4:** Descriptive Statistics

| | Count | Mean | Std | Min | Median | Max |
|---|---|---|---|---|---|---|
| CPU | 101919 | 82.76 | 45.56 | 0.00 | 75.44 | 243.28 |
| CPUp | 101919 | 81.82 | 31.45 | 0.00 | 88.88 | 139.97 |
| CPUm | 101919 | 81.31 | 31.43 | 0.00 | 89.42 | 145.27 |
| Green | 101919 | 0.20 | 1.71 | 0.00 | 0.00 | 114.00 |
| Dirty | 101919 | 0.17 | 2.72 | 0.00 | 0.00 | 363.00 |
| Grey | 101919 | 0.04 | 1.53 | 0.00 | 0.00 | 236.00 |
| Green stock | 101919 | 0.96 | 7.00 | 0.00 | 0.00 | 356.25 |
| Dirty stock | 101919 | 0.87 | 11.74 | 0.00 | 0.00 | 1075.12 |
| SPILLGreen | 101919 | 4420.78 | 3876.12 | 11.46 | 3458.82 | 37151.78 |
| SPILLDirty | 101919 | 3026.18 | 2312.28 | 9.43 | 2620.57 | 17798.43 |
| EPS | 101919 | 2.45 | 1.51 | 0.33 | 2.46 | 5.17 |
| Emit | 94377 | 0.02 | 0.18 | 0.00 | 0.00 | 8.22 |
| ShPatents | 94377 | 0.09 | 0.24 | 0.00 | 0.00 | 1.00 |

in terms of exposure to policy uncertainty. Firms, in fact, are not subject to Climate Policy Uncertainty deriving from only the country in which they are headquartered, but CPU is weighted by the average share of inventors that the firm has in that country. Inventors' shares are built using REGPAT's database, and each CPU measure is in turn constructed as:

$$CPU_{i,t} = \sum_{c \in C} w_{i,c} * CPU_{c,t} \tag{4.2}$$

Where $w_{i,c}$ is a time-invariant, firm-specific weight, where $w_{i,c}$ is the (average) share of inventors of firm $i$ in country $c$, over the period of observation for the firm. Inventors are drawn from the patents' database, and they are assigned to both a country and a firm, based on available information on inventor's location. While inventors could have moved throughout time, I build this indicator as the average share of inventors that each firm has in each country $c$. I construct identical measures for the directional indicators of positive CPU (variable $CPUp$) and negative CPU (variable $CPUm$). In the main specifications testing for the hypothesis, I include both variables for CPU direction as regressors. In equation 4.1, I include country by year fixed effects in order to control for macroeconomic conditions and business cycle dynamics that might be correlated with the dependent variables. While country-level policy stringency in environmental regulation should be captured by the country-year fixed effects, I also include an additional control for $EPS$, interacted with sectoral $GHG$ emissions' intensity, following Berestycki et al. (2022).[11]. Additionally, I include sector by year fixed effects

[11]In robustness checks, I test a different versions of this control, building it similarly to Equation 4.2

in order to capture sectoral trends that might be correlated with patenting patterns. $K_{i,t-1}$ is the total patent stock of firms, controlling for the size of its innovation portfolio.

Another two symmetric equations are estimated: one for the amount of dirty patents, and one for subset of dirty patents considered grey technologies, as detailed in the previous Section. I time-lag uncertainty and control variables to reflect delayed response, as well as to help mitigate contemporaneous feedback effects. Given the slowdown in investments at a one year lag found in previous exercises, I assume that at least three years should be necessary for the effects to translate onto the outcomes of the innovation process, i.e. on patent applications. I run robustness checks for different lag structures in the robustness Section. In addition, I also lag previous patent stocks, and other controls of one year, reflecting again the path-dependency of the innovation process. Patent stocks are constructed following an inventory rule with depreciation rate $r$ of 20%:

$$K_{i,t} = (1 - r)K_{i,t-1} + PAT_{i,t} \tag{4.3}$$

The total patent stock of a firm, however, does not allow me to distinguish between green and dirty technologies already available to the firm, as well as the geographical spillovers that might affect patenting patterns. In order to explicitly account for the path dependency of the innovation process, I break $K_{i,t}$ into different components, accounting for both internal and external spillovers, again following closely the approach proposed in Aghion, Dechezleprêtre et al. (2016). The stock of knowledge of the firm can be expressed in green and dirty components, both internal and external to the firm:

$$K_{i,t} = GreenStock_{i,t} + DirtyStock_{i,t} + SPGreen_{i,t} + SPDirty_{i,t} \quad (4.4)$$

Where:

- $GreenStock_{i,t}$ is the firm's own green patent stock;
- $DirtyStock_{i,t}$ is the firm's own dirty patent stock;
- $SPGreen_{i,t}$ are country-level green spillovers to firm $i$ in period $t$;
- $SPdirty_{i,t}$ are country-level dirty spillovers to firm $i$ in period $t$;

The stocks of green and dirty patents, for firm $i$, are again constructed using the inventory method, and control for the path-dependency in the innovation process: the probability of patenting in a specific technology depends on the past track-record of technologies patented in that domain. In addition, country-level spillovers to firm $i$, control for the external factors that can affect the focal firm's patenting: green (dirty) patenting, can be influenced by the availability of similar technologies outside the firm in that same country. Firms can learn from the available knowledge pool in green and dirty patents, which affects the probability of applying for more patents in the following years. The construction is again symmetrical for dirty and green technologies, and similar to the approach used for that of policy uncertainty. For green technologies, the spillovers available to firm $i$ at time $t$ are:

$$SPGreen_{i,t} = \sum_{c} w_{i,c} * SPGreen_{c,t} \tag{4.5}$$

The spillover pool in country $c$ ($SPGreen_{c,t}$) is defined as the sum of all other firms' patent stocks of green technologies ($KGreen_{j,t}$):

$$SPGreen_{c,t} = \sum_{j \neq i} w_{j,c} KGreen_{j,t} \tag{4.6}$$

As detailed in Aghion, Dechezleprêtre et al. (2016) and discussed in follow-up works (Schickfus 2021), the baseline Directed Technical Change model estimated with two-way fixed effects, would be inconsistent under strict exogeneity, due to serial correlation of the different patent stocks constructed. Thus, borrowing from their approach, I implement the Blundell-Griffith-Van Reenen (BGVR) estimator (Blundell et al. 1999), which relies on the pre-sample mean of the dependent variable in order to proxy for individual fixed effects. This approach is well-suited to patent data, and in empirical setups where data for the dependent variable is available for the pre-sample period.

I run a set of symmetric models for green, dirty and grey patent flows using a Maximum-Likelihood Poisson estimators, accounting for the count data nature of the dependent variables. In addition to the pre-sample mean of the dependent variable, I add controls for firm-level variables. First, I control for the share of patents of the firm in its country-sector (variable $ShPatents$), controlling for potential competition effects. In robustness checks, I also control explicitly for the size of the firms. Because of the limited availability of consistent balance sheet information (employment or total assets) before the 2010 period, I build a time-invariant variable, collecting the last available data point for the firms' assets. I build a categorical variable for the size of the firm (variable $FirmSize$) based on the quartiles of the distribution of assets. All right-hand side variables are log-transformed, including the pre-sample mean. Finally, given some firms have no lagged patent stocks for some periods, I follow Aghion, Dechezleprêtre et al. (2016) and add three dummy variables if the green or dirty (lagged) stocks are zero, or if both are zero. In Table 4.4 of the Appendix, I provide descriptive statistics for all the variables created.

## 4.4 Results

I first test the baseline models, regressing the counts of green, dirty, and grey patents against the aggregate index for CPU. In Table 4.5, I present estimation results, where in all specifications I include the dummies for the absence of green or dirty patents in the past stock of the firms, which are always significant and not reported. In even columns, I add to the baseline specification the control for the share of country-sector patents of the firm. Standard errors are always clustered at the firm level.

**Table 4.5:** Poisson Regression - Baseline estimates for CPU - Green, Dirty and Grey patents.

| | Green | | Dirty | | Grey | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| CPU | 0.0658*** | 0.0660*** | 0.0908** | 0.0870** | 0.1180 | 0.1159 |
| | (0.0138) | (0.0144) | (0.0359) | (0.0348) | (0.0895) | (0.0905) |
| GreenStock | 0.9670*** | 0.9648*** | 0.0878*** | 0.1007*** | 0.0384 | 0.0570 |
| | (0.0174) | (0.0174) | (0.0195) | (0.0202) | (0.0529) | (0.0580) |
| DirtyStock | 0.0398*** | 0.0536*** | 0.9418*** | 0.9437*** | 0.9952*** | 1.004*** |
| | (0.0121) | (0.0117) | (0.0169) | (0.0166) | (0.0593) | (0.0555) |
| SPILLgreen | 0.4909** | 0.5183*** | 0.1076 | 0.1245 | 0.7554 | 0.8411 |
| | (0.1940) | (0.1958) | (0.3593) | (0.3784) | (0.9403) | (1.004) |
| SPILLdirty | -0.3984* | -0.4170** | -0.1201 | -0.1140 | -1.192 | -1.243 |
| | (0.2051) | (0.2055) | (0.3660) | (0.3842) | (0.9333) | (1.001) |
| pre-sample mean | -0.0206*** | -0.0066 | -0.0468*** | -0.0339*** | 0.0228 | 0.0351 |
| | (0.0058) | (0.0061) | (0.0089) | (0.0086) | (0.0456) | (0.0494) |
| Emit | -0.0480 | -0.0466 | -0.0853 | -0.0884 | 0.4875* | 0.5017* |
| | (0.0639) | (0.0635) | (0.0681) | (0.0660) | (0.2599) | (0.2581) |
| EPS*Emit | 0.0110 | -0.0000 | 0.0449 | 0.0309 | 0.2935 | 0.2620 |
| | (0.0479) | (0.0475) | (0.0484) | (0.0474) | (0.2373) | (0.2236) |
| ShPatents | | -0.0514*** | | -0.0640*** | | -0.0926 |
| | | (0.0120) | | (0.0192) | | (0.0742) |
| Observations | 86,562 | 86,562 | 82,082 | 82,082 | 59,452 | 59,452 |
| Pseudo $R^2$ | 0.65247 | 0.65321 | 0.76757 | 0.76821 | 0.81496 | 0.81546 |
| RMSE | 0.81964 | 0.82541 | 0.86330 | 0.87008 | 0.50891 | 0.51270 |
| Country*Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Industry*Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. *CPU* captures firm-level exposure to climate policy uncertainty. *GreenStock* and *DirtyStock* are, respectively, the depreciated stocks of own-firm past green or dirty Triadic Patent Families. *SPILLGreen* and *SPILLDirty* are firm-level geographical spillovers to the focal firm in green and dirty technologies. *Emit* are country-sector-year emissions, and *EPS* is the index of Environmental Policy Stringency. *ShPatents* is the share of firm patents within its country-sector-year. All models contain dummy variables in case the past stock of green and (or) dirty patents is 0, which are always significant and not reported. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Models are estimated by Maximum-likelihood Poisson regressions. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the firm level.

I find a positive relationship between the aggregate index of CPU for both green and dirty technologies, suggesting that both climate-related technologies are sensitive to increases in aggregate uncertainty. The coefficients for green and dirty own stocks of past patents are positive and significant, as expected. Own patent stocks coefficients have greater sizes depending on the technology observed: green past stocks have a higher magnitude for green technologies than for dirty and grey technologies, and in the latter case are also insignificant. On the contrary, own stocks of dirty innovations have a higher magnitude for dirty and grey technologies. The positive sign for own past stocks of opposite nature (dirty stocks for green, and green stocks for dirty) have a smaller coefficient but are still positive, indicating possible between-technology spillovers within firms. External spillovers $SPILLGreen$ are also positively correlated with green patent flows, while $SILLdirty$ are have a negative correlation, whereas they are insignificant for dirty technologies.

As mentioned, the aggregate index does not allow us to distinguish between the different directions of uncertainty-related indexes. Therefore, in Table 4.6, I test the two complementary hypotheses developed in Section 4.2, and include as independent variables of interest the indexes for $CPUp$ and $CPUm$, respectively reflecting CPU capturing the positive or negative direction of uncertainty in terms of environmental regulation.

In line with the expectations in hypothesis 1, column (2) shows a positive

**Table 4.6:** Poisson Regression - Baseline estimates for CPU - Positive and negative policy stance.

| | Green | | Dirty | | Grey | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| CPUp | 0.6122** | 0.6394** | -0.2737 | -0.2948 | -1.244** | -1.335** |
| | (0.2510) | (0.2523) | (0.3014) | (0.3120) | (0.6017) | (0.6204) |
| CPUm | -0.5453** | -0.5718** | 0.3654 | 0.3836 | 1.354** | 1.446** |
| | (0.2518) | (0.2532) | (0.2959) | (0.3062) | (0.5983) | (0.6123) |
| GreenStock | 0.9671*** | 0.9649*** | 0.0877*** | 0.1007*** | 0.0378 | 0.0566 |
| | (0.0174) | (0.0174) | (0.0195) | (0.0202) | (0.0529) | (0.0580) |
| DirtyStock | 0.0397*** | 0.0536*** | 0.9418*** | 0.9438*** | 0.9957*** | 1.005*** |
| | (0.0121) | (0.0117) | (0.0168) | (0.0166) | (0.0593) | (0.0555) |
| SPILLgreen | 0.5073*** | 0.5379*** | 0.1062 | 0.1245 | 0.7106 | 0.8024 |
| | (0.1944) | (0.1961) | (0.3598) | (0.3797) | (0.9538) | (1.019) |
| SPILLdirty | -0.4175** | -0.4390** | -0.1173 | -0.1117 | -1.141 | -1.195 |
| | (0.2049) | (0.2052) | (0.3663) | (0.3852) | (0.9472) | (1.019) |
| pre-sample mean | -0.0207*** | -0.0066 | -0.0468*** | -0.0338*** | 0.0229 | 0.0355 |
| | (0.0058) | (0.0061) | (0.0089) | (0.0086) | (0.0456) | (0.0495) |
| Emit | -0.0477 | -0.0464 | -0.0838 | -0.0868 | 0.4939* | 0.5099** |
| | (0.0640) | (0.0636) | (0.0680) | (0.0659) | (0.2598) | (0.2579) |
| EPS*Emit | 0.0097 | -0.0014 | 0.0458 | 0.0319 | 0.3022 | 0.2714 |
| | (0.0481) | (0.0477) | (0.0485) | (0.0475) | (0.2387) | (0.2245) |
| ShPatents | | -0.0516*** | | -0.0643*** | | -0.0942 |
| | | (0.0120) | | (0.0192) | | (0.0743) |
| Observations | 86,562 | 86,562 | 82,082 | 82,082 | 59,452 | 59,452 |
| Pseudo $R^2$ | 0.65254 | 0.65328 | 0.76759 | 0.76824 | 0.81505 | 0.81557 |
| RMSE | 0.81957 | 0.82535 | 0.86289 | 0.86967 | 0.50835 | 0.51212 |
| Country*Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Industry*Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. *CPUp* captures firm-level exposure to climate policy uncertainty with an environmentally-positive policy stance. *CPUm* instead captures environmentally-negative policy stance. *GreenStock* and *DirtyStock* are, respectively, the depreciated stocks of own-firm past green or dirty Triadic Patent Families. *SPILLGreen* and *SPILLDirty* are firm-level geographical spillovers to the focal firm in green and dirty technologies. *Emit* are country-sector-year emissions, and *EPS* is the index of Environmental Policy Stringency. *ShPatents* is the share of firm patents within its country-sector-year. All models contain dummy variables in case the past stock of green and (or) dirty patents is 0, which are always significant and not reported. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Models are estimated by Maximum-likelihood Poisson regressions. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the firm level.

relationship between $CPUp$ and green patenting, while turning negative for $CPUm$. Interestingly, while the directions of the signs are in line with the hypothesis, the coefficients for the total of dirty technologies are insignificant. However, in columns (5)-(6), I present the same results only considering the subset of dirty patents comprehending grey technologies, which are in this case significant. In line with hypothesis 1b, coefficients suggest that uncertainty due to potential setbacks in climate policy-making, is positively related with more grey patenting, while the opposite happens in the case of green technologies. These results seem to confirm both hypotheses. Unpacking aggregate Climate Policy Uncertainty reveals a symmetric relationship with green and polluting inventions, suggesting that, depending on its stance, CPU is an important factor for directing technological change.

In order to to confirm these results, I test for different measures of $CPUp$ and $CPUm$. In table 4.7, in odd columns, I include the ratios between the country-level index for positive-leaning uncertainty ($RatioP$) over the general CPU index and its counterpart $RatioM$. In even columns, instead, I calculate the ratio between the number of positive or negative events over the total number of events (variables $PeaksP$ and $PeaksM$) detected with the algorithm described in Section 4.3.1. Again, these results seem to confirm the two hypotheses, suggesting a that there is a significant relationship between the direction of uncertainty in climate policies, and that of the technological efforts undertaken by firms.

**Table 4.7:** Poisson Regression - Baseline estimates for CPU stances - alternative measures.

| | Green | | Dirty | | Grey | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| RatioP | 0.7808** | | -0.4563 | | -1.330* | |
| | (0.3084) | | (0.3831) | | (0.7177) | |
| RatioM | -0.6551** | | 0.5410 | | 1.424** | |
| | (0.3076) | | (0.3689) | | (0.7043) | |
| PeaksP | | 0.1190*** | | -0.0067 | | -0.0312 |
| | | (0.0432) | | (0.0512) | | (0.0688) |
| PeaksM | | -0.0041 | | 0.1085* | | 0.2242** |
| | | (0.0460) | | (0.0597) | | (0.0967) |
| Observations | 72,040 | 72,040 | 67,809 | 67,809 | 49,725 | 49,725 |
| Pseudo $R^2$ | 0.65688 | 0.65693 | 0.76741 | 0.76745 | 0.81231 | 0.81239 |
| RMSE | 0.86039 | 0.86051 | 0.92261 | 0.92246 | 0.52828 | 0.52888 |
| Country*Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Industry*Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Full Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. *RatioP* is the ratio between the total level of CPU and its sub-index of environmentally-positive stance. *RatioM* is the ratio between CPU and environmentally-negative CPU. *PeaksP* and *PeaksM*, respectively, represent the total number of events of positive or negative stance, over the total number of events detected by the peak-detection algorithm. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models contain full controls and dummies for null past stocks of green and (or) dirty patents. Models are estimated by Maximum-likelihood Poisson regressions. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the firm level.

### 4.4.1 Heterogeneity and robustness

In Table 4.8 I divide the sample into different historical periods, running separate estimations focusing on the historical evolution of this relationship. Interestingly, while for the period 1995-2005 only green technologies seem sensitive to the directions of policy uncertainty, the coefficients for dirty technologies are again relevant and much higher for the 2010-2020 period, with notable differences between grey and dirty technologies. Different historical phases, seem to suggest a high degree of heterogeneity, across time, both for acceleration and deceleration of policies, and for the development of climate-relevant technologies. This heterogeneity could be driven by acceleration and deceleration of specific policies. In the last decade, in light of an increased implementation of climate policies, green and dirty might be perceived as diametric alternatives, and policy-direction has probably been more credible in terms of their support of one of the two.

Additionally, innovations evolved over time, and the technological linkages between alternative technologies might be changing over time. Grey technologies (on average) appeared more sensitive to $CPUp$ and $CPUm$ then general dirty ones, while the significance seem to be driven by dirty ones in the most recent period. While speculative, these results seem promising in analyzing the dynamic effects of CPU, depending on technological maturity, and in terms of the degree of substitutability and complementarity of technologies with opposite environmental effects.

In the Appendix, I report several robustness checks. In Table C.1, I run the same baseline estimates for $CPUp$ and $CPUm$, by weighting regressions by the average stock of patents of the firm. The coefficients for $CPUm$ and dirty patents are of higher magnitude, and more significant if compared with Table 4.6, suggesting that $CPUm$, pointing a negative direction, might stimulate overall polluting patenting. Weighting also confirms, at the 10% significance level the correlation between $CPUp$ and green patents. Grey patents remain highly sensitive to both directions of CPU, with coefficients of larger magnitudes.

In Table C.3 I add two controls. First, I include a control for the measure of Economic Policy Uncertainty (Baker et al. 2016), calculated analogously to that of CPU. In addition, I also include the categorical variable for quartiles of firm size. The direction of the effects is consistent with previous results, with $CPUp$ and $CPUm$ having opposite effects on green and dirty technologies, and again the latter are driven by grey patents. In Table C.4 I build a control for Environmental Policy Stringency similar to that I build for my measure of CPU, again substantially confirming the relevance of the CPU sub-indexes for DTC.

In Table C.2 I run a leave-one out exercise, excluding one country at the time from the sample. Interestingly, it seems that Germany is driving the significance in results, as it is possible to see in the last three columns. While the numerosity of firms left in the sample could be an issue, this result is very suggestive on the underlying geographical heterogeneity of CPU. A promising direction in policy uncertainty research, is the investigation of cross-country linkages and spillovers of policy uncertainty (Balli et al. 2017; Abakah et al. 2021). One possible reason for this result could be the relative weight that Germany has in both European climate policy-making, and as a powerhouse for the production of green technologies. Furthermore, the integration of value chains across countries (and between technologies) might also be a factor at play, and the role of technological linkages between products and industries should be further explored.

**Table 4.8:** Poisson Regression. CPU stances: Historical analysis.

| | 1995-2005 | | | 2000-2010 | | | 2005-2015 | | | 2010-2020 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Green | Dirty | Grey | Green | Dirty | Grey | Green | Dirty | Grey | Green | Dirty | Grey |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| CPUp | 0.9749** | -0.4172 | -0.8527 | 0.1774 | -2.897* | -2.660** | 0.2081 | 0.3255 | -2.280** | 2.155*** | -3.308*** | -1.459 |
| | (0.3897) | (0.5638) | (1.130) | (0.6943) | (1.527) | (1.298) | (0.4515) | (0.6777) | (1.048) | (0.8050) | (0.8258) | (1.339) |
| CPUm | -0.8883** | 0.6099 | 0.9368 | -0.1295 | 3.004* | 2.754** | -0.1710 | -0.2453 | 2.339** | -2.121*** | 3.285*** | 1.485 |
| | (0.3847) | (0.5533) | (1.112) | (0.6925) | (1.537) | (1.349) | (0.4587) | (0.6781) | (1.087) | (0.8076) | (0.8196) | (1.336) |
| Num. Firms | 3952 | 3952 | 3952 | 4011 | 4011 | 4011 | 3908 | 3908 | 3908 | 3678 | 3678 | 3678 |
| Observations | 23,107 | 22,800 | 18,207 | 24,069 | 23,370 | 18,016 | 23,594 | 22,150 | 16,300 | 21,560 | 18,728 | 11,971 |
| Pseudo $R^2$ | 0.60821 | 0.80119 | 0.86894 | 0.67785 | 0.76176 | 0.79337 | 0.68525 | 0.75026 | 0.71657 | 0.64898 | 0.68547 | 0.62463 |
| RMSE | 0.72646 | 1.2568 | 0.67817 | 1.0306 | 0.86549 | 0.39836 | 1.0658 | 0.70783 | 0.37594 | 0.71811 | 0.53636 | 0.28074 |
| Country*Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Industry*Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Full Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. *CPUp* captures firm-level exposure to climate policy uncertainty with an environmentally-positive policy stance. *CPUm* instead captures environmentally-negative policy stance. All models contain full controls and dummy variables in case the past stock of green and (or) dirty patents is 0, which are always significant and not reported. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Models are estimated by Maximum-likelihood Poisson regressions. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the firm level.

More puzzling, instead, are the dynamics on the timing of these associations. Figure C.2 in the Appendix plots the coefficients and standard errors of $CPUp$ and $CPUm$ tested at different time lags. In the case of green patents, the only significant lag is at 3 years, and the effects disappears in the longer term for all dependent variables. The significance across technologies at t-3, could be confirming the idea that the innovation output takes time to react to a rise in uncertainty. Interestingly, however, short-term correlations with grey inventions seem to be in the same direction as that of green technologies. This refers again to the nature of different technologies as substitutes or complements, revealing a potential dynamic complementarity between grey and green patents. Firms might be adopting a strategic behavior in reducing emissions of their products in the short term, while switching to alternative green technologies in the longer term.

However, further research is needed in this direction to account for the high-degree of volatility in patenting activity. Grey patents only represent a small fraction of total patenting activity, and a macro-level analysis modeling more precisely sectoral dynamics could help clarify this evidence. Moreover, a source of noise could also be the use of the earliest filing applications for patents. Filing years of technologies are an approximation of the timing of innovation activities, but are also the byproduct of legal proceedings, and crucially depend on invention quality. In this sense, keeping in mind that these correlations regard high-end innovation, for which arguably the cost in R&D is higher, looking at the whole spectrum of patent quality could reveal a different pattern of strategic behavior relevant for directed technical change. Evidence on uncertainty and the qualitative features of innovation could render important insights (Bhattacharya et al. 2017). In this framework, the interplay between the business-cycle features of uncertainty and heterogeneity in innovation development could shed further light on the forces directing clean and dirty technological change (Manso et al. 2023).

### 4.4.2 Limitations and further research

A number of other limitations apply to this study. First, while being a promising research avenue, linking innovation activities proxied by patents and uncertainty measures, suffers from a discrepancy in frequencies of the data employed. Patent filing dates are relevant at the yearly level, but, as shown in Section 4.3.1, much variation in CPU indexes is lost by aggregating at the yearly level. This loss of information about higher-frequency dynamics limits the understanding of short and long-term behaviors of firms. Other studies employing quarterly measures for investments, relying on firms' reporting, exploit this variation, losing on the technological heterogeneity by considering aggregate investments. Future research could exploit higher-frequency measures, for example in survey data, to bring further evidence on the time-dynamics of green vs brown technologies. As mentioned, these dynamics might also be explored in the light of the linkages between low and high-carbon inventions. As these technologies present spillovers and path-dependencies, there might be supply-chain related factors at play, which policymakers should be considering.

The geographical coverage of this analysis is mainly driven by the availability of adequate sources of news data. The external validity of this study, therefore, warrants a more thorough analysis, especially considering the peculiar evolution of news markets in each specific country.

Furthermore, while the fixed effects strategy employed in this paper should capture relevant confounders, more econometric work is necessary in order to asses the causality of the relationships uncovered. First, the use of pre-sample means and the BGVR estimator could not be fully capturing firm fixed effects. Similar approaches, both in terms of control function estimations and structural modeling could be suitable to address these issues. Second, the nature of firms is increasingly global. Better data is necessary in order to disentangle their geographical presence, both for emissions (which can have strongly local components), and for the geography of their intellectual property protection. Additionally, a number of political economy concerns could be biasing these results, specifically in terms of reverse causality. Bigger firms and more dominant technological actors, arguably have a much higher potential for influencing government action via lobbying efforts, or by setting an anti-environmental agenda in the media. While these concerns are mitigated by the fact that climate policies are often discussed in an inter-governmental setup, and that the consensus for action is confirmed by international treaties, lobbying could cause increases in CPU, or in one of its sub-components.

Nevertheless, this exercise uncovers another potential avenue for future research. The lack of sectoral variability in CPU indexes (and most of the available EPU ones) could be an important gap to fill. Sub-indexes derived for $CPUp$ and $CPUm$, and are blind to sectors and technologies. Arguably, both climate policies and CPU are indeed sector-specific, if not technology-specific, thinking for example to innovation subsidies. Recent papers (Juhász et al. 2022; Evenett et al. 2024) applied text-as-data techniques to the categorization of policy texts, quantifying (green) industrial policy efforts. In this sense, empirical applications based on rich textual data, as in this work, could give nuance to uncertainty measures, adding sectoral or technological dimensions. In the same vein of (Gugler et al. 2024), heterogeneity in CPU linked to specific policy instruments (subsidies, carbon taxes, etc.) could be explored. A promising approach for future applications is the mix of policy-stance with content analysis, breaking down policy-uncertainty into sub-components.

## 4.5 Conclusions

In this chapter, I investigate the role of Climate Policy Uncertainty for directing technical change. I build a novel dataset by scraping newspaper archives for four European countries: Germany, France, Italy and Spain. I apply text-as-data techniques to derive sub-measures for policy stance underlying CPU, showing a high-degree of variation between positive-leaning and negative-leaning articles. I bring forward additional empirical evidence by constructing a panel of European firms, and testing relevance of CPU sub-indexes in directing environmentally-sensitive technologies. I employ a model of directed technical change and study patenting activity of firms in both low-carbon and polluting technologies.

I find that CPU pointing towards stronger climate policy implementation is positively associated to the development of green technologies, and instead negatively to polluting ones. On the contrary, the measure of CPU implying setbacks weakening climate policy, shows symmetric results, by favoring dirty innovation and discouraging low carbon inventions. These results suggest that CPU might be affecting firms' expectations about the future value of environmentally-sensitive inventions and the direction of their R&D efforts. I propose a novel approach to identify CPU articles from newspaper data, showing that it could be flexibly applied to different data sources in multilingual contexts, through the use of supervised deep learning architectures.

In line with the extant literature, I find that not only realized climate policy but also uncertainty about the probability of future policies affects firm behaviors (Basaglia et al. 2021; Khalil and Strobel 2023). These findings suggest a complex and dynamic response, in terms of firms behavior, to differently-leaning uncertainty, and show the relevance of CPU directionality in the belief revision of firms. I add to the previous literature by showing that the direction of this probability has opposite effects on green and dirty innovations. Directing the economy away from a carbon intensive equilibrium to a cleaner growth path is a priority for policymakers, and governments have been experimenting with mission-oriented policy agendas for sustainability (Mazzucato 2018), and new forms of green industrial policy (Rodrik 2014). Consensus, clarity and communication surrounding climate and green industrial policies is deemed even more relevant in light of the significant effects on patenting patterns. Legislators can accelerate divestment from fossil technologies and foster green growth by committing to a decisive climate policy agenda. Governments can provide clear signals to the market in support of low-carbon growth, altering firms' expectations and directing innovative efforts towards a cleaner growth path. While preliminary, these results suggest that potential cost of climate policy could be lowered by ensuring coherent market signals to firms. On top of policies, government certainty and commitment to a strong climate agenda could spur virtuous circles, steering economic growth towards a low-carbon future.

# Chapter 5

# Conclusions

This dissertation contributes, throughout its three core chapters, to our understanding of the spatial and environmental dynamics of technological change. The first two essays investigate innovation from a fine-grained geographical perspective. The first part investigates the determinants of recombinant novelty at the level of NUTS3 European provinces between 2003 and 2017, making use of patent data and exploring the role of both inward and outward FDIs as potential vectors of external technological components. The findings reveal that inward FDIs are positively associated with local recombinant novelty, facilitating regional access to global knowledge pools, and stimulating knowledge transfers. In contrast, outward FDIs generally exhibit a negative association with novel recombination in patents, potentially leading to a "hollowing-out" effect where local innovation capabilities are reduced. The study also highlights the importance of knowledge proximity, showing that the less related the incoming knowledge is to the local base, the greater the potential for novel recombination. The findings indicate that FDIs play a nuanced role not only depending on their direction. The geographical and functional heterogeneity of these associations is also evident, with knowledge-intensive FDIs in Resarch and Development activities (R&D) showing a more positive impacts compared to non-R&D investments. In particular, these functional characteristics appear relevant in terms of the potential negative effects of OFDIs on local innovation. In addition, this chapter uncovers a large geographical heterogeneity. The study explores both the geography of the origins and destinations of FDIs, and employ spatial econometric models to uncover spatial spillovers. These are found to be generally negative both in terms of outward and inward FDIs, uncovering further the spatial dynamics of external connectivity and regional recombinant novelty.

The second essay investigates the impact of green Foreign Direct Investments (FDIs) and local skill composition on regional green technological diversification in the United States, across 287 cities (Metropolitan Statistical Areas, MSAs) from 2003 to 2018, adopting the framework of evolutionary economic geography. It finds that MSAs with higher levels of green FDIs and abstract skill intensity are more likely to diversify into green technological domains. The study highlights that both external knowledge from FDIs and local capabilities, especially abstract skills, play crucial roles in enabling green diversification. Additionally, local skills moderate the effects of green FDIs through compensation and reinforcing mechanisms. Abstract skills, supporting exploratory capabilities, compensate for weaker external knowledge flows and contribute positively to the process of branching towards green domains. In contrast, routine skills, important to the application of knowledge, enhance the impact of green FDIs by facilitating the transformation of external knowledge flows in green innovation. The results support the hypothesis that global knowledge flows, combined with local skills, are essential drivers of the green transition.

The third essay of this thesis explores the effects of Climate Policy Uncertainty (CPU) on directed technical change (DTC) across European firms from 1990 to 2020. Using novel CPU indices derived from newspaper articles and text-as-data techniques, the study examines how CPU can be a factor affecting the direction of firms' technological efforts towards (or away from) low-carbon inventions. This research makes use of firm-level data and studies patenting activity of firms both in green and polluting technologies. The findings reveal that the direction implied in policy uncertainty, either by signaling advancements or setbacks in climate policies, is a relevant factor for directing firms' expectations. CPU implying a positive environmental direction is positively associated with green patenting and negatively with polluting patenting. Conversely, negative CPU, indicating a weakening of climate policy, positively associates to the development of polluting technologies, potentially hindering green technological development. This reinforces the idea that firms adjust their innovation efforts based on expected future policy environments, contributing to steer the direction of technical change.

The empirical evidence presented in this dissertation has relevant implications for the current economic policy debate. The first essay suggests that promoting inward FDIs, particularly those that introduce diverse and unrelated knowledge, can enhance regional innovation. Policymakers should strive to prevent potential negative spillovers reducing regional innovation potential, particularly in terms of outward FDIs. Generally, the results from this chapter calls for the attention of policymakers to nuanced considerations of the spatial aspects of connectivity policy, in a highly contextualized territorial dimension. The second essay shows that attracting green FDIs, while enhancing the local workforce's abstract skills can accelerate green technological diversification. Policymakers, in this sense, should focus on creating environments that attract green investments and encourage exploration and innovation, with a particular attention to skill development in the workforce. This approach will enable regions to access and utilize global knowledge flows, advancing green innovation. Finally, the empirical evidence in the third essay suggests that reducing climate policy uncertainty and providing clearer, long-term signals can accelerate the transition to low-carbon technologies. Policymakers must ensure consistent and strong climate policy commitments, and be aware that negative uncertainty in the climate policy-making process can also delay the transition, steering innovation away from green technologies.

# Chapter 2 References


Acs, Zoltan J et al. (2002). "Patents and innovation counts as measures of regional production of new knowledge". In: *Research policy* 31.7, pp. 1069–1085.

Aghion, Philippe, Ufuk Akcigit et al. (2019). "Innovation and top income inequality". In: *The Review of Economic Studies* 86.1, pp. 1–45.

Aitken, Brian J and Ann E Harrison (1999). "Do domestic firms benefit from direct foreign investment? Evidence from Venezuela". In: *American economic review* 89.3, pp. 605–618.

Albulescu, Claudiu Tiberiu and Daniel Goyeau (2019). "The interaction between trade and FDI: The CEECs experience". In: *International Economics and Economic Policy* 16, pp. 489–509.

Amendolagine, Vito, Lucia Piscitello et al. (2022). "The impact of OFDI in global cities on innovation by Indian multinationals". In: *Applied Economics* 54.12, pp. 1352–1365.

Amendolagine, Vito, Andrea F Presbitero et al. (2019). "Local sourcing in developing countries: The role of foreign direct investments and global value chains". In: *World Development* 113, pp. 73–88.

Antonelli, Cristiano (2007). "The system dynamics of collective knowledge: From gradualism and saltationism to punctuated change". In: *Journal of Economic Behavior & Organization* 62.2, pp. 215–236.

— (2014). *The economics of innovation, new technologies and structural change*. Routledge.

Antonietti, Roberto, Raffaello Bronzini et al. (Mar. 2015). "Inward Greenfield FDI and Innovation". In: *Economia e Politica Industriale* 42.1, pp. 93–116. ISSN: 0391-2078, 1972-4977. DOI: 10.1007/s40812-014-0007-9.

Antonietti, Roberto and Chiara Franco (Mar. 2021). "From FDI to Economic Complexity: A Panel Granger Causality Analysis". In: *Structural Change and Economic Dynamics* 56, pp. 225–239. ISSN: 0954349X. DOI: 10.1016/j.strueco.2020.11.001.

Arts, Sam, Jianan Hou et al. (Mar. 2021). "Natural Language Processing to Identify the Creation and Impact of New Technologies in Patent Text: Code, Data, and New Measures". In: *Research Policy* 50.2, p. 104144. ISSN: 00487333. DOI: 10.1016/j.respol.2020.104144.

Arts, Sam and Reinhilde Veugelers (Dec. 2015). "Technology Familiarity, Recombinant Novelty, and Breakthrough Invention". In: *Industrial and Corporate Change* 24.6, pp. 1215–1246. ISSN: 0960-6491, 1464-3650. DOI: 10.1093/icc/dtu029.

Ascani, Andrea, Pierre-Alexandre Balland et al. (June 2020a). "Heterogeneous Foreign Direct Investment and Local Innovation in Italian Provinces". In: *Structural Change and Economic Dynamics* 53, pp. 388–401. ISSN: 0954-349X. DOI: 10.1016/j.strueco.2019.06.004.

Ascani, Andrea, Luca Bettarelli et al. (Oct. 2020). "Global Networks, Local Specialisation and Regional Patterns of Innovation". In: *Research Policy* 49.8, p. 104031. ISSN: 00487333. DOI: 10.1016/j.respol.2020.104031.

Ascani, Andrea and Luisa Gagliardi (2015). "Inward FDI and local innovative performance. An empirical investigation on Italian provinces". In: *Review of Regional Research* 35, pp. 29–47.

— (2020). "Asymmetric spillover effects from MNE investment". In: *Journal of World Business* 55.6, p. 101146.

Bannò, Mariasole et al. (2014). "The Impact of Public Support on SMEs' Outward FDI: Evidence from I taly". In: *Journal of Small Business Management* 52.1, pp. 22–38.

Bathelt, Harald and Maximilian Buchholz (2019). "Outward foreign direct investments as a catalyst of urban-regional income development? Evidence from the United States". In: *Economic Geography* 95.5, pp. 442–466.

Bathelt, Harald, Maximilian Buchholz and John A Cantwell (2023). "OFDI activity and urban-regional development cycles: a co-evolutionary perspective". In: *Competitiveness Review: An International Business Journal* 33.3, pp. 512–533.

Bathelt, Harald, Anders Malmberg et al. (2004a). "Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation". In: *Progress in human geography* 28.1, pp. 31–56.

Bello, Michela et al. (2023). *The role of inward innovative foreign direct investments in the acquisition of foreign knowledge: the case of EU green and digital inventions*. Publications Office of the European Union.

Berkes, Enrico and Ruben Gaetani (2021). "The geography of unconventional innovation". In: *The Economic Journal* 131.636, pp. 1466–1514.

Bertrand, Olivier and Laurence Capron (2015). "Productivity enhancement at home via cross-border acquisitions: The roles of learning and contemporaneous domestic investments". In: *Strategic Management Journal* 36.5, pp. 640–658.

Boschma, Ron (2017a). "Relatedness as driver of regional diversification: A research agenda". In: *Regional Studies* 51.3, pp. 351–364.

— (2022a). "Global value chains from an evolutionary economic geography perspective: a research agenda". In: *Area Development and Policy* 7.2, pp. 123–146.

Boschma, Ron, Gaston Heimeriks et al. (2014). "Scientific knowledge dynamics and relatedness in biotech cities". In: *Research policy* 43.1, pp. 107–114.

Boschma, Ron, Víctor Martín et al. (2017). "Neighbour regions as the source of new industries". In: *Papers in Regional Science* 96.2, pp. 227–246.

Buchholz, Maximilian et al. (2020). "Income divergence and global connectivity of US urban regions". In: *Journal of International Business Policy* 3, pp. 229–248.

Burger, Martijn J. and Evert J. Meijers (Mar. 2016). "Agglomerations and the Rise of Urban Network Externalities: Agglomerations and the Rise of Urban Network Externalities”. In: *Papers in Regional Science* 95.1, pp. 5–15. ISSN: 10568190. DOI: 10.1111/pirs.12223.

Cantwell, John and Simona Iammarino (2005). *Multinational corporations and European regional systems of innovation.* Taylor & Francis.

Carnabuci, Gianluca and Elisa Operti (2013a). “Where do firms’ recombinant capabilities come from? Intraorganizational networks, knowledge, and firms’ ability to innovate through technological recombination”. In: *Strategic management journal* 34.13, pp. 1591–1613.

Castaldi, Carolina, Koen Frenken et al. (May 2015). “Related Variety, Unrelated Variety and Technological Breakthroughs: An Analysis of US State-Level Patenting”. In: *Regional Studies* 49.5, pp. 767–781. ISSN: 0034-3404, 1360-0591. DOI: 10.1080/00343404.2014.940305.

Castaldi, Carolina and Bart Los (2017). “Geographical patterns in US inventive activity 1977–1998: The “regional inversion” was underestimated”. In: *Research Policy* 46.7, pp. 1187–1197.

Castellani, Davide, Davide Castellani et al. (2006). *Multinational firms, innovation and productivity.* Edward Elgar Publishing.

Castellani, Davide, Maria Luisa Mancusi et al. (2015). *Exploring the links between offshoring and innovation.*

Castellani, Davide, Giovanni Marin et al. (Jan. 2022a). “Greenfield Foreign Direct Investments and Regional Environmental Technologies”. In: *Research Policy* 51.1, p. 104405. ISSN: 00487333. DOI: 10.1016/j.respol.2021.104405.

Castellani, Davide, Valentina Meliciani et al. (2016). “The determinants of inward foreign direct investment in business services across European regions”. In: *Regional Studies* 50.4, pp. 671–691.

Castellani, Davide and Fabio Pieri (2016). “Outward investments and productivity: Evidence from European regions”. In: *Regional Studies* 50.12, pp. 1945–1964.

Castellani, Davide and Antonello Zanfei (2007). “Internationalisation, innovation and productivity: how do firms differ in Italy?” In: *World Economy* 30.1, pp. 156–176.

Cohen, Wesley M and Daniel A Levinthal (1990). "Absorptive capacity: A new perspective on learning and innovation". In: *Administrative science quarterly*, pp. 128–152.

Cozza, Claudio et al. (2015). "The impact of outward FDI on the performance of Chinese firms". In: *China Economic Review* 36, pp. 42–57.

Crescenzi, Riccardo, Marco Di Cataldo et al. (2021). "FDI inflows in Europe: Does investment promotion work?" In: *Journal of International Economics* 132, p. 103497.

Crescenzi, Riccardo, Arnaud Dyèvre et al. (2022). "Innovation catalysts: how multinationals reshape the global geography of innovation". In: *Economic Geography* 98.3, pp. 199–227.

Crescenzi, Riccardo, Luisa Gagliardi et al. (2015). "Foreign multinationals and domestic innovation: Intra-industry effects and firm heterogeneity". In: *Research Policy* 44.3, pp. 596–609.

Crescenzi, Riccardo and Simona Iammarino (2018). "Global investments and regional development trajectories: the missing links". In: *Transitions in regional economic development*. Routledge, pp. 171–203.

Crinò, Rosario (2009). "Offshoring, multinationals and labour market: a review of the empirical literature". In: *Journal of Economic Surveys* 23.2, pp. 197–249.

D'agostino, Lorena M (2015). "The neglected effects of R&D captive offshoring in emerging countries on the creation of knowledge at home". In: *Economia e Politica Industriale* 42, pp. 61–91.

Damijan, Jože et al. (2017). "Not every kind of outward FDI increases parent firm performance: the case of new EU member states". In: *Emerging Markets Finance and Trade* 53.1, pp. 74–97.

Damioli, Giacomo and Giovanni Marin (2024). "The effects of foreign entry on local innovation by entry mode". In: *Research Policy* 53.3, p. 104957.

De Noni, Ivan and Fiorenza Belussi (2021). "Breakthrough invention performance of multispecialized clustered regions in Europe". In: *Economic Geography* 97.2, pp. 164–186.

Elekes, Zoltán et al. (2019). "Foreign-owned firms as agents of structural change in regions". In: *Regional Studies* 53.11, pp. 1603–1613.

Ernst, Dieter and Linsu Kim (2002). "Global production networks, knowledge diffusion, and local capability formation". In: *Research policy* 31.8-9, pp. 1417–1429.

Fleming, Lee (Jan. 2001). "Recombinant Uncertainty in Technological Search". In: *Management Science* 47.1, pp. 117–132. ISSN: 0025-1909, 1526-5501. DOI: 10.1287/mnsc.47.1.117.10671.

Fosfuri, Andrea and Massimo Motta (1999). "Multinationals without advantages". In: *Scandinavian Journal of Economics* 101.4, pp. 617–630.

Frenken, Koen et al. (July 2007). "Related Variety, Unrelated Variety and Regional Economic Growth". In: *Regional Studies* 41.5, pp. 685–697. ISSN: 0034-3404. DOI: 10.1080/00343400601120296.

Fu, Xiaolan et al. (2018). "Unpacking the relationship between outward direct investment and innovation performance: evidence from Chinese firms". In: *World Development* 102, pp. 111–123.

Gagliardi, Luisa et al. (2021). "Exposure to OFDI and regional labour markets: evidence for routine and non-routine jobs in Great Britain". In: *Journal of Economic Geography* 21.5, pp. 783–806.

Garcia, Renato, Veneziano Araujo, Suelene Mascarini, Emerson Gomes Santos et al. (2023). "How Industrial Diversity Shapes the Effects of Foreign Direct Investment Spillovers on Regional Innovation". In: *International Regional Science Review* 46.1, pp. 98–122.

Garcia, Renato, Veneziano Araujo, Suelene Mascarini, Emerson Gomes Santos et al. (2024). "Unveiling the role of different types of local absorptive capacity: The effects of inward FDI spillovers on regional innovation". In: *Papers in Regional Science* 103.5, p. 100044.

García, Francisco et al. (Feb. 2013). "Does Inward Foreign Direct Investment Improve the Innovative Performance of Local Firms?" In: *Research Policy* 42.1, pp. 231–244. ISSN: 00487333. DOI: 10.1016/j.respol.2012.06.005.

Giorgi, Julien et al. (2024). *Inter-regional highly skilled worker mobility and technological novelty*. Insee, Institut national de la statistique et des études économiques.

Helpman, Elhanan et al. (2004). "Export versus FDI with heterogeneous firms". In: *American economic review* 94.1, pp. 300–316.

Huang, Youxing and Yan Zhang (2020). "The innovation spillovers from outward and inward foreign direct investment: A firm-level spatial analysis". In: *Spatial Economic Analysis* 15.1, pp. 43–59.

Iammarino, Simona (2018). "FDI and regional development policy". In: *Journal of International Business Policy* 1, pp. 157–183.

Iammarino, Simona and Philip McCann (2013). *Multinationals and economic geography: Location, technology and innovation*. Edward Elgar Publishing.

Iammarino, Simona, Philip McCann and Raquel Ortega-Argilés (2018). "International business, cities and competitiveness: recent trends and future challenges". In: *Competitiveness Review: An International Business Journal* 28.3, pp. 236–251.

Javorcik, Beata Smarzynska (2004). "Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages". In: *American economic review* 94.3, pp. 605–627.

Jin, Byungchae et al. (July 2019). "Inward Foreign Direct Investment and Local Firm Innovation: The Moderating Role of Technological Capabilities". In: *Journal of International Business Studies* 50.5, pp. 847–855. ISSN: 0047-2506, 1478-6990. DOI: 10.1057/s41267-018-0177-1.

Li, Xing et al. (2021). "Research on the impact of OFDI on the home country's global value chain upgrading". In: *International Review of Financial Analysis* 77, p. 101862.

Lybbert, Travis J. and Nikolas J. Zolas (Apr. 2014b). "Getting Patents and Economic Data to Speak to Each Other: An 'Algorithmic Links with Probabilities' Approach for Joint Analyses of Patenting and Economic Activity". In: *Research Policy* 43.3, pp. 530–542. ISSN: 0048-7333. DOI: 10.1016/j.respol.2013.09.001.

Mewes, Lars (2019). "Scaling of atypical knowledge combinations in American metropolitan areas from 1836 to 2010". In: *Economic geography* 95.4, pp. 341–361.

Miguelez, Ernest and Andrea Morrison (2023). "Migrant inventors as agents of technological change". In: *The Journal of technology transfer* 48.2, pp. 669–692.

Montresor, Sandro, Gianluca Orsatti et al. (2023). "Technological novelty and key enabling technologies: evidence from European regions". In: *Economics of Innovation and New Technology* 32.6, pp. 851–872.

Neffke, Frank et al. (Jan. 2018a). "Agents of Structural Change: The Role of Firms and Entrepreneurs in Regional Diversification". In: *Economic Geography* 94.1, pp. 23–48. ISSN: 0013-0095, 1944-8287. DOI: 10.1080/00130095.2017.1391691.

Nelson, Richard R and Sidney G Winter (1982). "An evolutionary theory of economic change". In: *Harvard Business School Press, Cambridge*.

Orsatti, Gianluca, Francesco Quatraro and Michele Pezzoni (2020b). "The antecedents of green technologies: The role of team-level recombinant capabilities". In: *Research Policy* 49.3, p. 103919.

Orsatti, Gianluca, Francesco Quatraro and Alessandra Scandura (2023). "Green technological diversification and regional recombinant capabilities: the role of technological novelty and academic inventors". In: *Regional Studies*, pp. 1–15.

— (2024). "Green technological diversification and regional recombinant capabilities: the role of technological novelty and academic inventors". In: *Regional Studies* 58.1, pp. 120–134.

Petit, Maria-Luisa and Francesca Sanna-Randaccio (2000). "Endogenous R&D and foreign direct investment in international oligopolies". In: *International Journal of Industrial Organization* 18.2, pp. 339–367.

Plunket, Anne and Felipe Starosta de Waldemar (2023). "Regional recombinant novelty, related and unrelated technologies: a patent-level approach". In: *Regional Studies* 57.7, pp. 1267–1288.

Quatraro, Francesco (2012). *The economics of structural change in knowledge*. Routledge.

Reddy, Ketan et al. (2022). "Outward foreign direct investment and domestic innovation efforts: Evidence from India". In: *Journal of Economics and Business* 122, p. 106084.

Rojec, Matija and Mark Knell (2018). “Why is there a lack of evidence on knowledge spillovers from foreign direct investment?” In: *Journal of Economic Surveys* 32.3, pp. 579–612.

Storper, Michael and Anthony J Venables (2004). “Buzz: face-to-face contact and the urban economy”. In: *Journal of economic geography* 4.4, pp. 351–370.

Strumsky, Deborah and José Lobo (2015). “Identifying the sources of technological novelty in the process of invention”. In: *Research Policy* 44.8, pp. 1445–1461.

Sturgeon, Timothy J. (2008). “Mapping Integrative Trade: Conceptualising and Measuring Global Value Chains”. In: *International Journal of Technological Learning, Innovation and Development* 1.3, p. 237. ISSN: 1753-1942, 1753-1950. DOI: 10.1504/IJTLID.2008.019973.

Valacchi, Giulia et al. (2021). “Impact of outward foreign direct investments on patent activity of Greenfield multinationals”. In: *Technological Forecasting and Social Change* 173, p. 121168.

Verhoeven, Dennis et al. (2016). “Measuring Technological Novelty with Patent-Based Indicators”. In: *Research Policy* 45.3, pp. 707–723.

Weitzman, Martin L. (May 1998). “Recombinant Growth*”. In: *The Quarterly Journal of Economics* 113.2, pp. 331–360. ISSN: 0033-5533. DOI: 10.1162/003355398555595.

Yang, Ruilin and Harald Bathelt (2022). “China’s outward investment activity: Ambiguous findings in the literature and empirical trends in greenfield investments”. In: *Growth and Change* 53.1, pp. 313–341.

Yeung, Henry Wai-chung (2021). “Regional worlds: from related variety in regional diversification to strategic coupling in global production networks”. In: *Regional Studies* 55.6, pp. 989–1010.

Zhang, Yongmin et al. (2024). “Innovation and OFDI along the Belt and Road”. In: *Technological Forecasting and Social Change* 200, p. 123136.

Zhou, Yu et al. (2019). “Green spillovers of outward foreign direct investment on home countries: Evidence from China’s province-level data”. In: *Journal of cleaner production* 215, pp. 829–844.

# Chapter 3 References


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

Amendolagine, Vito, Ulrich Elmer Hansen et al. (2023). "Do green foreign direct investments increase the innovative capability of MNE subsidiaries?" In: *World Development* 170, p. 106342.

Amendolagine, Vito, Rasmus Lema et al. (2021). "Green foreign direct investments and the deepening of capabilities for sustainable innovation in multinationals: Insights from renewable energy". In: *Journal of Cleaner Production* 310, p. 127381.

Antonelli, Cristiano (2006). "Diffusion as a process of creative adoption". In: *The Journal of Technology Transfer* 31, pp. 211–226.

Ascani, Andrea, Pierre-Alexandre Balland et al. (2020b). "Heterogeneous foreign direct investment and local innovation in Italian Provinces". In: *Structural Change and Economic Dynamics* 53, pp. 388–401.

Ascani, Andrea, Riccardo Crescenzi et al. (2016). "Economic institutions and the location strategies of European multinationals in their geographic neighborhood". In: *Economic Geography* 92.4, pp. 401–429.

Autor, David H. and David Dorn (2013). "The growth of low-skill service jobs and the polarization of the US Labor Market". In: *American Economic Review* 103.5, pp. 1553–1597. ISSN: 00028282. DOI: 10.1257/aer.103.5.1553. arXiv: arXiv:1011.1669v3.

Autor, David H., Frank Levy et al. (2003). "The skill content of recent technological change: An empirical exploration". In: *Quarterly Journal of Economics* 118.4, pp. 1279–

1333. ISSN: 00335533. DOI: 10.1162/003355303322552801. arXiv: arXiv:1406.6943v1.

Barbieri, Nicolò, Claudia Ghisetti et al. (2016). “A Survey of the Literature on Environmental Innovation Based on Main Path Analysis”. In: *Journal of Economic Surveys* 30.3, pp. 596–623.

Barbieri, Nicolò, Alberto Marzucchi et al. (2020). “Knowledge sources and impacts on subsequent inventions: Do green technologies differ from non-green ones?” In: *Research Policy* 49.2, p. 103901.

Bathelt, Harald and Pengfei Li (2020). “Processes of building cross-border knowledge pipelines”. In: *Research Policy* 49.3, p. 103928.

Bathelt, Harald, Anders Malmberg et al. (2004b). “Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation”. In: *Progress in Human Geography* 28.1, pp. 31–56.

Boschma, Ron (2017b). “Relatedness as driver of regional diversification: A research agenda”. In: *Regional Studies* 51.3, pp. 351–364.

— (2022b). “Global value chains from an evolutionary economic geography perspective: a research agenda”. In: *Area Development and Policy* 7.2, pp. 123–146.

Boschma, Ron, Lars Coenen et al. (2018). “Towards a theory of regional diversification: Combining insights from evolutionary economic geography and transition studies”. In: *Transitions in Regional Economic Development*. Routledge, pp. 55–81.

Carnabuci, Gianluca and Elisa Operti (2013b). “Where do firms’ recombinant capabilities come from? Intraorganizational networks, knowledge, and firms’ ability to innovate through technological recombination”. In: *Strategic Management Journal* 34 (13), pp. 1591–1613. ISSN: 01432095. DOI: 10.1002/smj.2084.

Castellani, Davide, Giovanni Marin et al. (2022b). “Greenfield foreign direct investments and regional environmental technologies”. In: *Research Policy* 51 (1). ISSN: 00487333. DOI: 10.1016/j.respol.2021.104405.

Cicerone, Gloria et al. (2023). “Regional artificial intelligence and the geography of environmental technologies: does local AI knowledge help regional green-tech specialization?” In: *Regional Studies* 57.2, pp. 330–343.

Cohen, Wesley M and Daniel A Levinthal (1989). “Innovation and learning: the two faces of R&D”. In: *The Economic Journal* 99.397, pp. 569–596.

Consoli, Davide, Fabrizio Fusillo et al. (2021). “Skill endowment, routinisation and digital technologies: evidence from U.S. Metropolitan Areas”. In: *Industry and Innovation* 28 (8). ISSN: 14698390. DOI: 10.1080/13662716.2021.1904842.

Consoli, Davide, Giovanni Marin et al. (2016). “Do green jobs differ from non-green jobs in terms of skills and human capital?” In: *Research Policy* 45.5, pp. 1046–1060.

Consoli, Davide and Francesco Rentocchini (2015). “A taxonomy of multi-industry labour force skills”. In: *Research Policy* 44.5, pp. 1116–1132.

Corrocher, Nicoletta et al. (2024). “Green technological diversification: The role of international linkages in leaders, followers and catching-up countries”. In: *Research Policy* 53.4, p. 104972.

Crescenzi, Riccardo, Carlo Pietrobelli et al. (2014). “Innovation drivers, value chains and the geography of multinational corporations in Europe”. In: *Journal of Economic Geography* 14.6, pp. 1053–1086.

De Marchi, Valentina et al. (2022). “Multinational subsidiaries and green innovation”. In: *International Business Review* 31.6, p. 102027.

Eck, Nees Jan van and Ludo Waltman (2009). “How to normalize cooccurrence data? An analysis of some well-known similarity measures”. In: *Journal of the American Society for Information Science and Technology* 60.8, pp. 1635–1651.

Fusillo, Fabrizio (2023). “Green Technologies and diversity in the knowledge search and output phases: Evidence from European Patents”. In: *Research Policy* 52.4, p. 104727.

Fusillo, Fabrizio, Davide Consoli et al. (2022). “Resilience, Skill Endowment, and Diversity: Evidence from US Metropolitan Areas”. In: *Economic Geography* 98.2, pp. 170–196. ISSN: 19448287. DOI: 10.1080/00130095.2021.2008797.

Fusillo, Fabrizio, Francesco Quatraro et al. (2022). "Going green: the dynamics of green technological alliances". In: *Economics of Innovation and New Technology* 31.5, pp. 362–386.

Gambardella, Alfonso (1992). "Competitive advantages from in-house scientific research: The US pharmaceutical industry in the 1980s". In: *Research policy* 21.5, pp. 391–407.

Ghisetti, Claudia, Alberto Marzucchi et al. (2015). "The open eco-innovation mode. An empirical investigation of eleven European countries". In: *Research policy* 44.5, pp. 1080–1093.

Girma, Sourafel (2005). "Absorptive capacity and productivity spillovers from FDI: a threshold regression analysis". In: *Oxford Bulletin of Economics and Statistics* 67.3, pp. 281–306.

Haltiwanger, John et al. (2013). "Who creates jobs? Small versus large versus young". In: *Review of Economics and Statistics* 95.2, pp. 347–361.

Haščič, Ivan and Mauro Migotto (2015a). "Measuring environmental innovation using patent data". In.

— (2015b). "Measuring environmental innovation using patent data". In: *OECD Environment Working Papers* No. 89 (89), p. 59. ISSN: 1997-0900. DOI: 10.1787/5js009kf48xw-en. URL: http://dx.doi.org/10.1787/5js009kf48xw-en.

Hidalgo, César A (2021). "Economic complexity theory and applications". In: *Nature Reviews Physics* 3.2, pp. 92–113.

Kemp, René (2010). "Eco-innovation: definition, measurement and open research issues". In: *Economia Politica* 27.3, pp. 397–420.

Losacker, Sebastian, Hendrik Hansmeier et al. (2023). "The geography of environmental innovation: a critical review and agenda for future research". In: *Review of Regional Research* 43.2, pp. 291–316.

Lund Vinding, Anker (2006). "Absorptive capacity and innovative performance: A human capital approach". In: *Economics of Innovation and New Technology* 15.4-5, pp. 507–517.

Lybbert, Travis J. and Nikolas J. Zolas (2014a). "Getting patents and economic data to speak to each other: An 'Algorithmic Links with Probabilities' approach for joint analyses of patenting and economic activity". In: *Research Policy* (43), pp. 530–542. ISSN: 00487333. DOI: 10.1016/j.respol.2013.09.001.

Malecki, Edward J (2002). "Creating and sustaining competitiveness: local knowledge and economic geography". In: *Knowledge, Space, Economy*. Routledge, pp. 103–119.

Marchi, Valentina De (2012). "Environmental innovation and R&D cooperation: Empirical evidence from Spanish manufacturing firms". In: *Research Policy* 41 (3), pp. 614–623. ISSN: 00487333. DOI: 10.1016/j.respol.2011.10.002.

Marino, Alba and Francesco Quatraro (2022). "Leveraging global recombinant capabilities for green technologies: The role of ethnic diversity in MNEs' dynamics". In: *The Journal of Technology Transfer*, pp. 1–33.

Mason, Geoff et al. (2020). "Which skills contribute most to absorptive capacity, innovation and productivity performance? Evidence from the US and Western Europe". In: *Economics of Innovation and New Technology* 29.3, pp. 223–241.

Mazzucato, Mariana and Carlota Perez (2015). "Innovation as growth policy". In: *The Triple Challenge for Europe*, pp. 229–64.

Montresor, Sandro and Francesco Quatraro (2020). "Green technologies and Smart Specialisation Strategies: a European patent-based analysis of the intertwining of technological relatedness and key enabling technologies". In: *Regional Studies* 54 (10). ISSN: 13600591. DOI: 10.1080/00343404.2019.1648784.

Neffke, Frank et al. (2018b). "Agents of structural change: The role of firms and entrepreneurs in regional diversification". In: *Economic Geography* 94.1, pp. 23–48.

Orsatti, Gianluca, François Perruchas et al. (2020). "Public Procurement, Local Labor Markets and Green Technological Change. Evidence from US Commuting Zones". In: *Environmental and Resource Economics*. ISSN: 15731502. DOI: 10.1007/s10640-020-00405-4.

Orsatti, Gianluca, Francesco Quatraro and Michele Pezzoni (2020a). "The antecedents of green technologies: The role of team-level recombinant capabilities". In: *Research Policy*. ISSN: 00487333. DOI: 10.1016/j.respol.2019.103919.

Parjanen, Satu et al. (2011). "Distances, knowledge brokerage and absorptive capacity in enhancing regional innovativeness: A qualitative case study of Lahti region, Finland". In: *European Planning Studies* 19.6, pp. 921–948.

Quatraro, Francesco and Alessandra Scandura (2019). "Academic inventors and the antecedents of green technologies. A regional analysis of Italian patent data". In: *Ecological Economics* 156, pp. 247–263.

Romijn, Henny and Manuel Albaladejo (2002). "Determinants of innovation capability in small electronics and software firms in southeast England". In: *Research Policy* 31.7, pp. 1053–1067.

Santoalha, Artur et al. (2021). "Digital skills, relatedness and green diversification: A study of European regions". In: *Research Policy* 50.9, p. 104340.

Tanner, Anne Nygaard (2014). "Regional branching reconsidered: Emergence of the fuel cell industry in European regions". In: *Economic Geography* 90.4, pp. 403–427.

— (2016). "The emergence of new technology-based industries: The case of fuel cells and its technological relatedness to regional knowledge bases". In: *Journal of Economic Geography* (16), pp. 661–635. ISSN: 14682710. DOI: 10.1093/jeg/lbv011.

UNEP (2023). *Adaptation Gap Report 2023: Underfinanced. Underprepared. Inadequate investment and planning on climate adaptation leaves world exposed*. Tech. rep. https://www.unep.org gap-report-2023: United Nations Environment Programme.

Vona, Francesco and Davide Consoli (2015). "Innovation and skill dynamics: A life-cycle approach". In: *Industrial and Corporate Change* 24.6, pp. 1393–1415. ISSN: 14643650. DOI: 10.1093/icc/dtu028.

Vona, Francesco, Giovanni Marin and Davide Consoli (2019). "Measures, drivers and effects of green employment: evidence from US local labor markets, 2006–2014". In: *Journal of Economic Geography* 19.5, pp. 1021–1048.

Vona, Francesco, Giovanni Marin, Davide Consoli and David Popp (2018). "Environmental regulation and green skills: an empirical exploration". In: *Journal of the Association of Environmental and Resource Economists* 5.4, pp. 713–753.

Zahra, Shaker A and Gerard George (2002). "Absorptive capacity: A review, reconceptualization, and extension". In: *Academy of Management Review* 27.2, pp. 185–203.

Zeppini, Paolo (2015). "A discrete choice model of transitions to sustainable technologies". In: *Journal of Economic Behavior & Organization* 112, pp. 187–203.

Zeppini, Paolo and Jeroen CJM van den Bergh (2011). "Competing recombinant technologies for environmental innovation: extending Arthur's model of lock-in". In: *Industry and Innovation* 18.03, pp. 317–334.

# Chapter 4 References


Abakah, Emmanuel Joel Aikins et al. (2021). “Economic policy uncertainty: Persistence and cross-country linkages”. In: *Research in International Business and Finance* 58, p. 101442.

Acemoglu, Daron, Philippe Aghion et al. (2012). “The environment and directed technical change”. In: *American economic review* 102.1, pp. 131–166.

Aghion, Philippe, Ketan Ahuja et al. (2023). *Sparking Europe’s new industrial revolution-A policy for net zero, growth and resilience*. Tech. rep. Bruegel.

Aghion, Philippe, Antoine Dechezleprêtre et al. (2016). “Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry”. In: *Journal of Political Economy* 124.1, pp. 1–51.

Altenburg, Tilman and Dani Rodrik (2017). “Green industrial policy: Accelerating structural change towards wealthy green economies”. In: *Green Industrial Policy*.

Bai, Dongbei et al. (2023). “Climate policy uncertainty and corporate green innovation: Evidence from Chinese A-share listed industrial corporations”. In: *Energy Economics* 127, p. 107020.

Baker, Scott R et al. (2016). “Measuring economic policy uncertainty”. In: *The quarterly journal of economics* 131.4, pp. 1593–1636.

Balli, Faruk et al. (2017). “Cross-country determinants of economic policy uncertainty spillovers”. In: *Economics Letters* 156, pp. 179–183.

Basaglia, Pier et al. (2021). “Climate policy uncertainty and firms’ and investors’ behavior”. In: *Unpublished manuscript*.

Berestycki, Clara et al. (2022). “Measuring and assessing the effects of climate policy uncertainty”. In.

Bernanke, Ben S (1983). “Irreversibility, uncertainty, and cyclical investment”. In: *The quarterly journal of economics* 98.1, pp. 85–106.

Bettarelli, Luca et al. (2023). *Environmental policies and innovation in renewable energy*. International Monetary Fund.

Blundell, Richard et al. (1999). “Market share, market value and innovation in a panel of British manufacturing firms”. In: *The review of economic studies* 66.3, pp. 529–554.

Botta, Enrico and Tomasz Koźluk (2014). “Measuring environmental policy stringency in OECD countries: A composite index approach”. In.

Bouri, Elie et al. (2022). “Climate policy uncertainty and the price dynamics of green and brown energy stocks”. In: *Finance Research Letters* 47, p. 102740.

Cascaldi-Garcia, Danilo et al. (2023). “What is certain about uncertainty?” In: *Journal of Economic Literature* 61.2, pp. 624–654.

Cherif, Reda and Fuad Hasanov (2019). *The return of the policy that shall not be named: Principles of industrial policy*. International Monetary Fund.

Cui, Xin et al. (2023). “Economic policy uncertainty and green innovation: Evidence from China”. In: *Economic Modelling* 118, p. 106104.

Dechezleprêtre, Antoine and David Hémous (2022). *Directed technical change and environmental economics*. Tech. rep. Working Paper.

Dechezleprêtre, Antoine, Ralf Martin et al. (2019). “Climate change policy, innovation and growth”. In: *Handbook on green growth*. Edward Elgar Publishing, pp. 217–239.

Dechezleprêtre, Antoine and Misato Sato (2017). “The impacts of environmental regulations on competitiveness”. In: *Review of environmental economics and policy*.

Dell, Melissa et al. (2024). “American stories: A large-scale structured text dataset of historical us newspapers”. In: *Advances in Neural Information Processing Systems* 36.

Devlin, Jacob (2018). “Bert: Pre-training of deep bidirectional transformers for language understanding”. In: *arXiv preprint arXiv:1810.04805*.

Dixit, Avinash K and Robert S Pindyck (1994). *Investment under uncertainty*. Princeton university press.

Dorsey, Jackson (2019). "Waiting for the courts: Effects of policy uncertainty on pollution and investment". In: *Environmental and Resource Economics* 74.4, pp. 1453–1496.

Dugoua, Eugenie et al. (2022). "Text as data in environmental economics and policy". In: *Review of Environmental Economics and Policy* 16.2, pp. 346–356.

Evenett, Simon et al. (2024). "The return of industrial policy in data". In: *The World Economy*.

Fankhauser, Sam et al. (2013). "Who will win the green race? In search of environmental competitiveness and innovation". In: *Global environmental change* 23.5, pp. 902–913.

Favot, Marinella et al. (2023). "Green patents and green codes: How different methodologies lead to different results". In: *Resources, Conservation & Recycling Advances* 18, p. 200132.

Feng, Gen-Fu and Mingbo Zheng (2022). "Economic policy uncertainty and renewable energy innovation: International evidence". In: *Innovation and Green Development* 1.2, p. 100010.

Feng, Yuanshun and Xiaonan Ma (2024). "The impact of perceived environmental uncertainty on green innovation: evidence from text recognition of Chinese listed companies' annual reports". In: *Environmental Science and Pollution Research* 31.5, pp. 7146–7166.

Gavriilidis, Konstantinos (2021). "Measuring climate policy uncertainty". In: *Available at SSRN 3847388*.

Ghisetti, Claudia and Francesco Quatraro (2017). "Green technologies and environmental productivity: A cross-sectoral analysis of direct and indirect effects in Italian regions". In: *Ecological Economics* 132, pp. 1–13.

Gilardi, Fabrizio et al. (2023). "ChatGPT outperforms crowd workers for text-annotation tasks". In: *Proceedings of the National Academy of Sciences* 120.30, e2305016120.

Gugler, Klaus et al. (2024). "Environmental Policies and directed technological change". In: *Journal of Environmental Economics and Management* 124, p. 102916.

Hémous, David and Morten Olsen (2021). “Directed technical change in labor and environmental economics”. In: *Annual Review of Economics* 13.1, pp. 571–597.

Hoang, Khanh (2022). “How does corporate R&D investment respond to climate policy uncertainty? Evidence from heavy emitter firms in the United States”. In: *Corporate Social Responsibility and Environmental Management* 29.4, pp. 936–949.

Hu, Yingde et al. (2023). “How does environmental policy uncertainty influence corporate green investments?” In: *Technological Forecasting and Social Change* 189, p. 122330.

Huang, Wendi (2023). “Climate policy uncertainty and green innovation”. In: *Economics Letters* 233, p. 111423.

IEA (2020). “IEA Flagship Report - Clean energy innovation. https://www.iea.org/reports/clean-energy-innovation.” In.

Iscen, Ahmet et al. (2019). “Label propagation for deep semi-supervised learning”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5070–5079.

Juhász, Réka et al. (2022). “The who, what, when, and how of industrial policy: A text-based approach”. In: *What, When, and How of Industrial Policy: A Text-Based Approach (August 15, 2022).*

Juroš, Jana et al. (2024). “LLMs for Targeted Sentiment in News Headlines: Exploring the Descriptive-Prescriptive Dilemma”. In: *Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis*, pp. 329–343.

Khalil, Makram and Felix Strobel (2023). “Capital reallocation under climate policy uncertainty”. In.

Kyaw, Khine (2022). “Effect of policy uncertainty on environmental innovation”. In: *Journal of Cleaner Production* 363, p. 132645.

Larsen, Vegard Høghaug (2021). “Components of uncertainty”. In: *International Economic Review* 62.2, pp. 769–788.

Li, Xiaoqing et al. (2021). "Environmental regulation, economic policy uncertainty, and green technology innovation". In: *Clean Technologies and Environmental Policy* 23, pp. 2975–2988.

Losacker, Sebastian, Jens Horbach et al. (2023). "Geography and the speed of green technology diffusion". In: *Industry and innovation* 30.5, pp. 531–555.

Lütkenhorst, Wilfried et al. (2014). "Green industrial policy: Managing transformation under uncertainty". In: *Deutsches Institut für Entwicklungspolitik Discussion Paper* 28.

Ma, Yan-Ran et al. (2023). "A news-based climate policy uncertainty index for China". In: *Scientific Data* 10.1, p. 881.

Manso, Gustavo et al. (2023). "Heterogeneous innovation over the business cycle". In: *Review of Economics and Statistics* 105.5, pp. 1224–1236.

Maraut, Stéphane et al. (2008). "The OECD REGPAT database: a presentation". In.

Mazzucato, Mariana (2011). "The entrepreneurial state". In: *Soundings* 49.49, pp. 131–142.

— (2018). "Mission-oriented innovation policies: challenges and opportunities". In: *Industrial and corporate change* 27.5, pp. 803–815.

McDonald, Robert and Daniel Siegel (1986). "The value of waiting to invest". In: *The quarterly journal of economics* 101.4, pp. 707–727.

Niu, Shuhai et al. (2023). "How does climate policy uncertainty affect green technology innovation at the corporate level? New evidence from China". In: *Environmental Research* 237, p. 117003.

Noailly, Joelle et al. (2022). *Does Environmental Policy Uncertainty Hinder Investments Towards a Low-Carbon Economy?* Tech. rep. National Bureau of Economic Research.

Pangakis, Nicholas and Samuel Wolken (2024). "Knowledge Distillation in Automated Annotation: Supervised Text Classification with LLM-Generated Training Labels". In: *arXiv preprint arXiv:2406.17633*.

Pangakis, Nicholas, Samuel Wolken and Neil Fasching (2023). "Automated annotation with generative ai requires validation". In: *arXiv preprint arXiv:2306.00176*.

Pegels, Anna (2014). *Green industrial policy in emerging countries*. Routledge London.

Peng, Xin-Yu et al. (2023). "How does economic policy uncertainty affect green innovation?" In: *Technological and Economic Development of Economy* 29.1, pp. 114–140.

Pindyck, Robert S (2021). "What we know and don't know about climate change, and implications for policy". In: *Environmental and Energy Policy and the Economy* 2.1, pp. 4–43.

Ren, Xiaohang, Yukun Shi et al. (2022). "Climate policy uncertainty and corporate investment: evidence from the Chinese energy industry". In: *Carbon Neutrality* 1.1, p. 14.

Ren, Xiaohang, Xiao Zhang et al. (2022). "Climate policy uncertainty and firm-level total factor productivity: Evidence from China". In: *Energy Economics* 113, p. 106209.

Rodríguez-Pose, Andrés and Federico Bartalucci (2023). "The green transition and its potential territorial discontents". In: *Cambridge Journal of Regions, Economy and Society*.

Rodrik, Dani (2014). "Green industrial policy". In: *Oxford review of economic policy* 30.3, pp. 469–491.

Schickfus, Marie-Theres von (2021). *Institutional investors, climate policy risk, and directed innovation*. Tech. rep. ifo Working Paper.

Schmitz, Hubert et al. (2013). "Rent management–the heart of green industrial policy". In: *IDS Working Papers* 2013.418, pp. 1–26.

Tobback, Ellen et al. (2018). "Belgian economic policy uncertainty index: Improvement through text mining". In: *International journal of forecasting* 34.2, pp. 355–365.

Törnberg, Petter (2024). "Best Practices for Text Annotation with Large Language Models". In: *arXiv preprint arXiv:2402.05129*.

Wade, Robert H (2014). "Return of industrial policy?" In: *Economic Policies of the New Thinking in Economics*. Routledge, pp. 76–92.

Wang, Jiawei Brooke (2022). "Waiting or Acting: The Effects of Climate Policy Uncertainty". In: *Available at SSRN 4028885*.

Wang, Mengyu et al. (2023). "Policy Uncertainty Reduces Green Investment". In: *Available at SSRN 4474726*.

Wang, Xinru et al. (2024). "Human-LLM collaborative annotation through effective verification of LLM labels". In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 1–21.

Xu, Yi and Zhenhuan Yang (2023). "Economic policy uncertainty and green innovation based on the viewpoint of resource endowment". In: *Technology Analysis & Strategic Management* 35.7, pp. 785–798.

Yamano, Norihiko and Joaquim Guilhoto (2020). "CO2 emissions embodied in international trade and domestic final demand: Methodology and results using the OECD Inter-Country Input-Output Database". In.

Yu, Haiyang and Haisong Chen (2023). "Economic policy uncertainty and green innovation: Does corporate social responsibility matter?" In: *Finance Research Letters* 58, p. 104649.

# Appendix A

# Appendix for Chapter 2

**Table A.1:** Descriptive Statistics of Variables

| **Variable** | **Count** | **Mean** | **Std** | **Min** | **Median** | **Max** |
|---|---|---|---|---|---|---|
| *Total projects* | | | | | | |
| EU FDI in | 17040 | 12.6 | 47.3 | 0.0 | 1.0 | 1,264.0 |
| EU FDI out | 17040 | 11.6 | 63.2 | 0.0 | 1.0 | 2,481.0 |
| ROW FDI in | 17040 | 14.9 | 67.1 | 0.0 | 0.0 | 2,280.0 |
| ROW FDI out | 17040 | 0.3 | 2.1 | 0.0 | 0.0 | 77.0 |
| EU15 FDI in | 17040 | 10.7 | 41.0 | 0.0 | 1.0 | 1,128.0 |
| EU15 FDI out | 17040 | 6.6 | 41.6 | 0.0 | 0.0 | 1,843.0 |
| nonEU15 FDI in | 17040 | 1.8 | 8.2 | 0.0 | 0.0 | 364.0 |
| nonEU15 FDI out | 17040 | 4.7 | 23.9 | 0.0 | 0.0 | 638.0 |
| *Only R&D projects* | | | | | | |
| IFDI | 17040 | 1.1 | 4.8 | 0.0 | 0.0 | 136.0 |
| OFDI | 17040 | 1.3 | 10.5 | 0.0 | 0.0 | 401.0 |
| EU IFDI | 17040 | 0.5 | 2.4 | 0.0 | 0.0 | 83.0 |
| EU OFDI | 17040 | 0.5 | 3.5 | 0.0 | 0.0 | 130.0 |
| ROW IFDI | 17040 | 0.6 | 2.6 | 0.0 | 0.0 | 78.0 |
| ROW OFDI | 17040 | 0.0 | 0.1 | 0.0 | 0.0 | 6.0 |
| EU15 IFDI | 17040 | 0.4 | 2.1 | 0.0 | 0.0 | 75.0 |
| EU15 OFDI | 17040 | 0.3 | 2.5 | 0.0 | 0.0 | 98.0 |
| nonEU15 IFDI | 17040 | 0.1 | 0.4 | 0.0 | 0.0 | 10.0 |
| nonEU15 OFDI | 17040 | 0.1 | 1.1 | 0.0 | 0.0 | 32.0 |
| *Only non-R&D projects* | | | | | | |
| IFDI | 17040 | 19.5 | 76.8 | 0.0 | 2.0 | 2,359.0 |
| OFDI | 17040 | 25.1 | 146.1 | 0.0 | 1.0 | 6,172.0 |
| EU IFDI | 17040 | 12.0 | 45.3 | 0.0 | 1.0 | 1,219.0 |
| EU OFDI | 17040 | 11.1 | 60.1 | 0.0 | 1.0 | 2,351.0 |
| ROW IFDI | 17040 | 7.4 | 33.6 | 0.0 | 0.0 | 1,140.0 |
| ROW OFDI | 17040 | 0.3 | 2.0 | 0.0 | 0.0 | 75.0 |
| EU15 IFDI | 17040 | 10.2 | 39.2 | 0.0 | 1.0 | 1,093.0 |
| EU15 OFDI | 17040 | 6.3 | 39.3 | 0.0 | 0.0 | 1,745.0 |
| nonEU15 IFDI | 17040 | 1.7 | 7.8 | 0.0 | 0.0 | 354.0 |
| nonEU15 OFDI | 17040 | 4.5 | 23.1 | 0.0 | 0.0 | 606.0 |

**Table A.2:** OLS two-way FE - Robustness checks

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| netIFDI | 0.0318*** | | | | | | | | | |
| | (0.0087) | | | | | | | | | |
| netCapex in | | 0.0082*** | | | | | | | | |
| | | (0.0025) | | | | | | | | |
| Capex in | | | 0.0138** | | | | | | | |
| | | | (0.0066) | | | | | | | |
| Capex in PIM5 | | | | 0.0124* | | | | | | |
| | | | | (0.0066) | | | | | | |
| Capex in PIM10 | | | | | 0.0108* | | | | | |
| | | | | | (0.0065) | | | | | |
| Capex per GDP in | | | | | | 0.0243*** | | | | |
| | | | | | | (0.0089) | | | | |
| Capex per GDP in PIM5 | | | | | | | 0.0224** | | | |
| | | | | | | | (0.0089) | | | |
| Capex per GDP in PIM10 | | | | | | | | 0.0203** | | |
| | | | | | | | | (0.0088) | | |
| IFDI noMA | | | | | | | | | 0.0435** | |
| | | | | | | | | | (0.0219) | |
| IFDI | | | | | | | | | | 0.0219* |
| | | | | | | | | | | (0.0117) |
| Full controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 |
| $R^2$ | 0.98289 | 0.98282 | 0.98280 | 0.98279 | 0.98279 | 0.98283 | 0.98282 | 0.98281 | 0.97359 | 0.98606 |
| Within $R^2$ | 0.04939 | 0.04568 | 0.04451 | 0.04421 | 0.04394 | 0.04621 | 0.04570 | 0.04521 | 0.00599 | 0.05105 |
| region FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

emphClustered (region) standard-errors in parentheses Dep var: Number of novel patents in NUTS3 regions. $IFDI$ is the cumulative count of inward FDI projects in the region. $netIFDI$ is the net stock of FDIs (IFDI - OFDI), $netCapex$ is the same measure for the amount of capital expenditures (CapexIN - CapexOUT). $CapexinPIM5$ is the amount of inward Capital expenditures depreciated at the 5% discount rate. $CapexinPIM10$ is depreciated at the 10% discount rate. Per GDP are similarly constructed, but normalized by regional real GDP at the NUTS3 level (ARDECO database). $IFDInoMA$ is constructed without the 4-year moving average transformation, in this case the dependent variable is the 4-years rolling stock of novel patents. In column 10, the dependent variable is constructed as the number of novel patents per capita. Explanatory variables are lagged by five years. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models are estimated through two-way fixed effects OLS estimators. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the NUTS3 level.

**Table A.3:** OLS two-way FE - Robustness checks

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Capex out | -0.0278*** | | | | | | | |
| | (0.0069) | | | | | | | |
| Capex out PIM5 | | -0.0268*** | | | | | | |
| | | (0.0068) | | | | | | |
| Capex out PIM10 | | | -0.0255*** | | | | | |
| | | | (0.0067) | | | | | |
| Capex per GDP out | | | | -0.0353*** | | | | |
| | | | | (0.0086) | | | | |
| Capex per GDP out PIM5 | | | | | -0.0342*** | | | |
| | | | | | (0.0085) | | | |
| Capex per GDP out PIM10 | | | | | | -0.0326*** | | |
| | | | | | | (0.0084) | | |
| OFDI noMA | | | | | | | -0.0515*** | |
| | | | | | | | (0.0179) | |
| OFDI | | | | | | | | -0.0175 |
| | | | | | | | | (0.0107) |
| Full controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 |
| $R^2$ | 0.98288 | 0.98287 | 0.98286 | 0.98288 | 0.98288 | 0.98287 | 0.97363 | 0.98603 |
| Within $R^2$ | 0.04910 | 0.04867 | 0.04819 | 0.04934 | 0.04888 | 0.04836 | 0.00732 | 0.04896 |
| region FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dep var: Number of novel patents in NUTS3 regions. $OFDI$ is the cumulative count of inward FDI projects in the region. $CapexoutPIM5$ is the amount of outward Capital expenditures depreciated at the 5% discount rate. $CapexoutPIM10$ is depreciated at the 10% discount rate. Per GDP are similarly constructed, but normalized by regional real GDP at the NUTS3 level (ARDECO database). $OFDInoMA$ is constructed without the 4-year moving average transformation, in this case the dependent variable is the 4-years rolling stock of novel patents. In column 8, the dependent variable is constructed as the number of novel patents per capita. Explanatory variables are lagged by five years. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models are estimated through two-way fixed effects OLS estimators. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the NUTS3 level.

**Table A.4:** OLS two-way FE

| | Novelty | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| IFDI | 0.0109 | 0.0130 | 0.0152 | 0.0263* | 0.0434*** | 0.0561*** | 0.0662*** | 0.0766*** |
| | (0.0119) | (0.0132) | (0.0144) | (0.0152) | (0.0159) | (0.0166) | (0.0175) | (0.0174) |
| Observations | 12,496 | 11,360 | 10,224 | 9,088 | 7,952 | 6,816 | 5,680 | 4,544 |
| $R^2$ | 0.97950 | 0.97875 | 0.97902 | 0.98065 | 0.98283 | 0.98531 | 0.98772 | 0.99034 |
| Within $R^2$ | 0.20887 | 0.12522 | 0.06964 | 0.04799 | 0.04643 | 0.03893 | 0.02613 | 0.01863 |
| Full controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Region FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dep var: Number of novel patents in NUTS3 regions. $IFDI$ is the cumulative count of inward FDI projects in the region. Explanatory variables are lagged by five years. In each column, the time lag for IFDI is increased, from $t-1$ (column 1) to $t-8$ (column 8). Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models are estimated through two-way fixed effects OLS estimators. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the NUTS3 level.

**Table A.5:** OLS two-way FE

| | Novelty | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| OFDI | 0.0159 | 0.0057 | -0.0131 | -0.0346** | -0.0441*** | -0.0397*** | -0.0239 | 0.0038 |
| | (0.0118) | (0.0129) | (0.0137) | (0.0141) | (0.0142) | (0.0145) | (0.0151) | (0.0162) |
| Observations | 12,496 | 11,360 | 10,224 | 9,088 | 7,952 | 6,816 | 5,680 | 4,544 |
| $R^2$ | 0.97955 | 0.97879 | 0.97906 | 0.98073 | 0.98284 | 0.98527 | 0.98763 | 0.99024 |
| Within $R^2$ | 0.21106 | 0.12714 | 0.07171 | 0.05229 | 0.04715 | 0.03640 | 0.01927 | 0.00804 |
| region FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dep var: Number of novel patents in NUTS3 regions. $OFDI$ is the cumulative count of outward FDI projects in the region. Explanatory variables are lagged by five years. In each column, the time lag for OFDI is increased, from $t-1$ (column 1) to $t-8$ (column 8). Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models are estimated through two-way fixed effects OLS estimators. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the NUTS3 level.

**Table A.6:** OLS two-way FE - FDI variables

| | All EU | | | | | | EU15 | | | | nonEU15 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| IFDI | 0.0434$^{**}$ | | | | | | | | | | | | | |
| | (0.0188) | | | | | | | | | | | | | |
| OFDI | | -0.0366$^{**}$ | | | | | | | | | | | | |
| | | (0.0151) | | | | | | | | | | | | |
| EU IFDI | | | 0.0583$^{***}$ | | | | | | | | | | | |
| | | | (0.0187) | | | | | | | | | | | |
| EU OFDI | | | | -0.0577$^{***}$ | | | | | | | | | | |
| | | | | (0.0153) | | | | | | | | | | |
| ROW IFDI | | | | | 0.0281 | | | | | | | | | |
| | | | | | (0.0198) | | | | | | | | | |
| ROW OFDI | | | | | | -0.0353$^{*}$ | | | | | | | | |
| | | | | | | (0.0192) | | | | | | | | |
| EU15 IFDI | | | | | | | 0.0365$^{*}$ | | | | 0.0921$^{*}$ | | | |
| | | | | | | | (0.0190) | | | | (0.0534) | | | |
| EU15 OFDI | | | | | | | | -0.0306$^{**}$ | | | | 0.0391 | | |
| | | | | | | | | (0.0136) | | | | (0.0578) | | |
| nonEU15 IFDI | | | | | | | | | 0.0266$^{*}$ | | | | 0.0396 | |
| | | | | | | | | | (0.0157) | | | | (0.0657) | |
| nonEU15 OFDI | | | | | | | | | | -0.0144 | | | | -0.0109 |
| | | | | | | | | | | (0.0149) | | | | (0.0665) |
| Full controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 7,952 | 6,426 | 6,426 | 6,426 | 6,426 | 1,526 | 1,526 | 1,526 | 1,526 |
| $R^2$ | 0.98283 | 0.98282 | 0.98286 | 0.98286 | 0.98277 | 0.98276 | 0.98402 | 0.98401 | 0.98401 | 0.98397 | 0.95094 | 0.95048 | 0.95072 | 0.95042 |
| Within $R^2$ | 0.04643 | 0.04577 | 0.04772 | 0.04779 | 0.04326 | 0.04217 | 0.01800 | 0.01728 | 0.01702 | 0.01507 | 0.08773 | 0.07912 | 0.08360 | 0.07798 |
| region FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Conley standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dep var: Number of novel patents in NUTS3 regions. $IFDI$ is the cumulative count of inward FDI projects in the region. $OFDI$ is the cumulative count of number of outward FDI projects from the region. The geographical breakdown of variables is computed as the flows, for the focal NUTS3 regions, only coming from (or going to) European countries ($EUFDI$), non-European countries ($ROWFDI$) or to respectively EU15 and nonEU15 countries, as detailed in section 2.3. Explanatory variables are lagged by five years. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models are estimated through two-way fixed effects OLS estimators. Conley standard errors are calculated with a radius of 100 kilometers are reported in parenthesis.

**Table A.7:** Spatial Durbin Model: Marginal effects

Panel A:

| | All EU | | | | | |
|---|---|---|---|---|---|---|
| | IFDI | OFDI | EU IFDI | EU OFDI | ROW IFDI | ROW OFDI |
| Direct Effects | 0.0437*** | -0.0056 | 0.0498*** | -0.0182*** | 0.0332*** | -0.0035 |
| Indirect Effects | -0.0070 | -0.1052*** | 0.0237* | -0.1166*** | -0.0534*** | -0.1721*** |
| Total Effects | 0.0366*** | -0.1108*** | 0.0734*** | -0.1347*** | -0.0202*** | -0.1756*** |
| Spatial Coefficient | 0.3101*** | 0.291*** | 0.3091*** | 0.294*** | 0.3021*** | 0.2951*** |
| Region FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 7952 | 7952 | 7952 | 7952 | 7952 | 7952 |
| Adj. R2 | 0.0909 | 0.0972 | 0.0929 | 0.0992 | 0.0899 | 0.0913 |

Panel B:

| | EU15 | | | | nonEU15 | | | |
|---|---|---|---|---|---|---|---|---|
| | EU15 IFDI | EU15 OFDI | nonEU15 IFDI | nonEU15 OFDI | EU15 IFDI | EU15 OFDI | nonEU15 IFDI | nonEU15 OFDI |
| Direct Effects | 0.0274*** | -0.0143*** | 0.0209*** | -0.0004 | 0.1609*** | 0.0656*** | 0.1151*** | 0.0407* |
| Indirect Effects | 0.0014 | -0.0473*** | -0.0702*** | -0.0494*** | -0.0841*** | -0.2019*** | -0.1402*** | -0.1787*** |
| Total Effects | 0.0288*** | -0.0617*** | -0.0493*** | -0.0498*** | 0.0768 | -0.1364** | -0.0252 | -0.1380** |
| Spatial Coefficient | 0.2481*** | 0.2481*** | 0.25*** | 0.2481*** | 0.2687*** | 0.2537*** | 0.2517*** | 0.2507*** |
| Region FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 6426 | 6426 | 6426 | 6426 | 1526 | 1526 | 1526 | 1526 |
| Adj. R2 | 0.0282 | 0.0305 | 0.0288 | 0.0285 | 0.1222 | 0.1193 | 0.12 | 0.115 |

Dep var: Number of novel patents in NUTS3 regions. Full controls included. $IFDI$ is the cumulative count of inward FDI projects in the region. $OFDI$ is the cumulative count of number of outward FDI projects from the region. The geographical breakdown of variables is computed as the flows, for the focal NUTS3 regions, only coming from (or going to) European countries ($EUFDI$), non-European countries ($ROWFDI$) or to respectively EU15 and nonEU15 countries, as detailed in section 2.3. Explanatory variables are lagged by five years. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Marginal effects are reported, the spatial coefficients refer to the first-order spatial autoregressive terms in the Spatial Durbin Model for recombinant novel patents. The spatial weights matrix is built with a 5-nearest-neighbour algorithm applied to NUTS3 European regions. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the NUTS3 level. *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.*

**Table A.8:** Spatial Durbin Model for Research and Development Projects: Marginal effects

Panel A:

| | All EU | | | | | |
|---|---|---|---|---|---|---|
| | IFDI | OFDI | EU IFDI | EU OFDI | ROW IFDI | ROW OFDI |
| Direct Effects | 0.0230*** | 0.0044 | 0.0711*** | -0.0112 | -0.0039 | 0.0559 |
| Indirect Effects | -0.0383** | -0.1310*** | 0.0549** | -0.1156*** | -0.1255*** | 0.0551 |
| Total Effects | -0.0153 | -0.1266*** | 0.1260*** | -0.1268*** | -0.1294*** | 0.1110* |
| Spatial Coefficient | 0.3031*** | 0.304*** | 0.304*** | 0.3091*** | 0.3041*** | 0.3071*** |
| Region FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 7952 | 7952 | 7952 | 7952 | 7952 | 7952 |
| Adj. R2 | 0.0888 | 0.0926 | 0.091 | 0.0905 | 0.0914 | 0.0885 |

Panel B:

| | EU15 | | | | nonEU15 | | | |
|---|---|---|---|---|---|---|---|---|
| | EU15 IFDI | EU15 OFDI | nonEU15 IFDI | nonEU15 OFDI | EU15 IFDI | EU15 OFDI | nonEU15 IFDI | nonEU15 OFDI |
| Direct Effects | 0.0625*** | 0.0261** | 0.0656*** | 0.0320** | 0.0788*** | -0.0308 | -0.0848 | 0.5444* |
| Indirect Effects | 0.1111*** | -0.0418 | -0.1857*** | -0.0147 | -0.1133* | -0.9546** | 0.0360 | 0.1135 |
| Total Effects | 0.1737*** | -0.0157 | -0.1201*** | 0.0172 | -0.0345 | -0.9855* | -0.0488 | 0.6579 |
| Spatial Coefficient | 0.2461*** | 0.2481*** | 0.262*** | 0.2501*** | 0.2597*** | 0.2467*** | 0.2567*** | 0.2577*** |
| Region FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 6426 | 6426 | 6426 | 6426 | 1526 | 1526 | 1526 | 1526 |
| Adj. R2 | 0.0308 | 0.0271 | 0.0272 | 0.0272 | 0.1143 | 0.1148 | 0.1101 | 0.1091 |

Dep var: Number of novel patents in NUTS3 regions. Full controls included. $IFDI$ is the cumulative count of inward FDI projects in R&D in the region. $OFDI$ is the cumulative count of number of outward FDI projects from the region. The geographical breakdown of variables is computed as the flows, for the focal NUTS3 regions, only coming from (or going to) European countries ($EUFDI$), non-European countries ($ROWFDI$) or to respectively EU15 and nonEU15 countries, as detailed in section 2.3. Explanatory variables are lagged by five years. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Marginal effects are reported, the spatial coefficients refer to the first-order spatial autoregressive terms in the Spatial Durbin Model for recombinant novel patents. The spatial weights matrix is built with a 5-nearest-neighbour algorithm applied to NUTS3 European regions. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the NUTS3 level. *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.*

**Table A.9:** Spatial Durbin Model excluding Research and Development Projects: Marginal effects

Panel A:

| | All EU | | | | | |
|---|---|---|---|---|---|---|
| | IFDI | OFDI | EU IFDI | EU OFDI | ROW IFDI | ROW OFDI |
| Direct Effects | | | | | | |
| | 0.0474*** | -0.0073 | 0.0526*** | -0.0189*** | 0.0480*** | -0.0036 |
| Indirect Effects | | | | | | |
| | -0.0054 | -0.1036*** | 0.0247* | -0.1171*** | -0.0608*** | -0.1737*** |
| Total Effects | | | | | | |
| | 0.0420*** | -0.1110*** | 0.0773*** | -0.1360*** | -0.0128 | -0.1773*** |
| Spatial Coefficient | 0.3141*** | 0.2971*** | 0.302*** | 0.2941*** | 0.316*** | 0.2941*** |
| Region FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 7952 | 7952 | 7952 | 7952 | 7952 | 7952 |
| Adj. R2 | 0.0914 | 0.0972 | 0.0934 | 0.0992 | 0.0905 | 0.0914 |

Panel B:

| | EU15 | | | | nonEU15 | | | |
|---|---|---|---|---|---|---|---|---|
| | EU15 IFDI | EU15 OFDI | nonEU15 IFDI | nonEU15 OFDI | EU15 IFDI | EU15 OFDI | nonEU15 IFDI | nonEU15 OFDI |
| Direct Effects | | | | | | | | |
| | 0.0303*** | 0.0033 | 0.0225*** | -0.0007 | 0.1605*** | 0.0708*** | 0.1209*** | 0.0395* |
| Indirect Effects | | | | | | | | |
| | -0.0023 | -0.0207* | -0.0609*** | -0.0477*** | -0.0754** | -0.1613** | -0.1610*** | -0.1687*** |
| Total Effects | | | | | | | | |
| | 0.0280*** | -0.0174** | -0.0384*** | -0.0484*** | 0.0852 | -0.0905 | -0.0401 | -0.1292** |
| Spatial Coefficient | 0.2561*** | 0.2541*** | 0.25*** | 0.247*** | 0.2587*** | 0.2547*** | 0.2577*** | 0.2467*** |
| Region FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 6426 | 6426 | 6426 | 6426 | 1526 | 1526 | 1526 | 1526 |
| Adj. R2 | 0.0285 | 0.0272 | 0.0285 | 0.0285 | 0.1219 | 0.1141 | 0.1217 | 0.1144 |

Dep var: Number of novel patents in NUTS3 regions. Full controls included. $IFDI$ is the cumulative count of inward FDI projects in R&D in the region. $OFDI$ is the cumulative count of number of outward FDI projects from the region. The geographical breakdown of variables is computed as the flows, for the focal NUTS3 regions, only coming from (or going to) European countries ($EUFDI$), non-European countries ($ROWFDI$) or to respectively EU15 and nonEU15 countries, as detailed in section 2.3. Explanatory variables are lagged by five years. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Marginal effects are reported, the spatial coefficients refer to the first-order spatial autoregressive terms in the Spatial Durbin Model for recombinant novel patents. The spatial weights matrix is built with a 5-nearest-neighbour algorithm applied to NUTS3 European regions. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the NUTS3 level. *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.*

# Appendix B

# Appendix for Chapter 3

**Table B.1:** Green Diversification, skill composition, and share of Green FDI Capex

| | entryGT | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| shGreenCapex | 0.4767*** | 0.3180*** | 0.2971*** | 0.1571** | 0.3528*** | 0.0192 |
| | (0.0847) | (0.0757) | (0.0745) | (0.0726) | (0.1125) | (0.0878) |
| (d) ASH | 0.3365*** | 0.2111*** | 0.2019*** | 0.1596*** | 0.2131*** | 0.1547*** |
| | (0.0460) | (0.0410) | (0.0403) | (0.0353) | (0.0405) | (0.0346) |
| (d) RSH | 0.0330 | 0.0410 | 0.0006 | -0.0237 | -0.0269 | -0.1006** |
| | (0.0446) | (0.0396) | (0.0370) | (0.0345) | (0.0343) | (0.0426) |
| TechRel | | 0.0148*** | 0.0154*** | 0.0128*** | 0.0128*** | 0.0128*** |
| | | (0.0009) | (0.0009) | (0.0008) | (0.0008) | (0.0008) |
| GDPpc | | | 0.3502*** | 0.1832* | 0.1831* | 0.1904* |
| | | | (0.1008) | (0.0978) | (0.0973) | (0.0982) |
| EmpGrowth | | | 0.1175 | -0.0952 | -0.1314 | -0.0969 |
| | | | (0.4040) | (0.3902) | (0.3873) | (0.3905) |
| ShPatents | | | -0.1446*** | -0.1474*** | -0.1465*** | -0.1467*** |
| | | | (0.0269) | (0.0273) | (0.0268) | (0.0271) |
| (d) shGreenEst | | | | 0.3171*** | 0.3186*** | 0.3162*** |
| | | | | (0.0826) | (0.0813) | (0.0818) |
| GreenPrevRTA | | | | 0.7824*** | 0.7981*** | 0.7752*** |
| | | | | (0.1860) | (0.1841) | (0.1832) |
| shGreenCapex × (d) ASH | | | | | -0.3030** | |
| | | | | | (0.1225) | |
| shGreenCapex × (d) RSH | | | | | | 0.3546*** |
| | | | | | | (0.1167) |
| Observations | 2,102,562 | 2,102,562 | 2,102,562 | 2,102,562 | 2,102,562 | 2,102,562 |
| Log-Likelihood | -108,388.2 | -107,996.9 | -107,883.2 | -107,801.5 | -107,793.9 | -107,789.9 |
| Adjusted Pseudo $R^2$ | 0.06639 | 0.06975 | 0.07070 | 0.07139 | 0.07145 | 0.07148 |
| Year*Tech fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| State fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Clustered (MSA) standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dep var: MSA entry in a green technological specialization. *shGreenCapex* is the share of green capital expenditures in inward FDIs over the total capital expenditures in inward FDIs. *ASH* and *RSH* are dichotomous variables equal to one if the share of, respectively, abstract and routine skills is above the national median. Explanatory variables are lagged by one year. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models are estimated through fixed effects logit estimators and include an indicator variable for the absence of inward FDIs in the MSA, whose estimated coefficient is not reported. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the MSA level. *p<0.1; **p<0.05; ***p<0.01

**Table B.2:** Green Diversification, skill composition, and shares of Green FDI (counts)

| | entryGT | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| shGreenFDI | 0.5139*** | 0.3483*** | 0.3274*** | 0.1693** | 0.3664*** | 0.0063 |
| | (0.0952) | (0.0852) | (0.0829) | (0.0814) | (0.1241) | (0.0992) |
| (d) ASH | 0.3413*** | 0.2136*** | 0.2032*** | 0.1601*** | 0.2156*** | 0.1559*** |
| | (0.0463) | (0.0410) | (0.0405) | (0.0353) | (0.0410) | (0.0346) |
| (d) RSH | 0.0361 | 0.0428 | 0.0017 | -0.0225 | -0.0251 | -0.1084** |
| | (0.0444) | (0.0395) | (0.0367) | (0.0343) | (0.0341) | (0.0433) |
| TechRel | | 0.0148*** | 0.0154*** | 0.0128*** | 0.0128*** | 0.0128*** |
| | | (0.0009) | (0.0009) | (0.0008) | (0.0008) | (0.0008) |
| GDPpc | | | 0.3559*** | 0.1871* | 0.1896* | 0.1966** |
| | | | (0.1004) | (0.0977) | (0.0969) | (0.0978) |
| EmpGrowth | | | 0.1024 | -0.1040 | -0.1441 | -0.1015 |
| | | | (0.4044) | (0.3901) | (0.3865) | (0.3903) |
| ShPatents | | | -0.1448*** | -0.1474*** | -0.1464*** | -0.1462*** |
| | | | (0.0270) | (0.0273) | (0.0267) | (0.0269) |
| (d) shGreenEst | | | | 0.3132*** | 0.3130*** | 0.3126*** |
| | | | | (0.0829) | (0.0818) | (0.0821) |
| GreenPrevRTA | | | | 0.7913*** | 0.8046*** | 0.7789*** |
| | | | | (0.1860) | (0.1833) | (0.1818) |
| shGreenFDI × (d) ASH | | | | | -0.3240** | |
| | | | | | (0.1354) | |
| shGreenFDI × (d) RSH | | | | | | 0.4218*** |
| | | | | | | (0.1301) |
| Observations | 2,102,562 | 2,102,562 | 2,102,562 | 2,102,562 | 2,102,562 | 2,102,562 |
| Log-Likelihood | -108,393.0 | -107,998.3 | -107,883.9 | -107,802.0 | -107,794.9 | -107,789.0 |
| Adjusted Pseudo $R^2$ | 0.06635 | 0.06974 | 0.07070 | 0.07139 | 0.07144 | 0.07149 |
| Year*Tech fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| State fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Clustered (MSA) standard-errors in parentheses*

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dep var: MSA entry in a green technological specialization. *shGreenFDI* is the share of green FDIs, as a count of inward FDI flows, over the total count of inward FDI flows. *ASH* and *RSH* are dichotomous variables equal to one if the share of, respectively, abstract and routine skills is above the national median. Explanatory variables are lagged by one year. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models are estimated through fixed effects logit estimators and include an indicator variable for the absence of inward FDIs in the MSA, whose estimated coefficient is not reported. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the MSA level. *p<0.1; **p<0.05; ***p<0.01

**Table B.3:** Green Diversification, skill composition, and Green Capex (4-digits tech FE)

| | entryGT | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| GreenCapex | 0.0476*** | 0.0311*** | 0.0421*** | 0.0220** | 0.0401*** | 0.0058 |
| | (0.0107) | (0.0101) | (0.0096) | (0.0091) | (0.0119) | (0.0107) |
| BrownCapex | 0.0388*** | 0.0204* | 0.0420*** | 0.0254*** | 0.0279*** | 0.0273*** |
| | (0.0114) | (0.0107) | (0.0102) | (0.0095) | (0.0096) | (0.0096) |
| (d) ASH | 0.3145*** | 0.2455*** | 0.2235*** | 0.1818*** | 0.2512*** | 0.1774*** |
| | (0.0531) | (0.0481) | (0.0454) | (0.0396) | (0.0463) | (0.0392) |
| (d) RSH | 0.0387 | 0.0487 | -0.0183 | -0.0325 | -0.0385 | -0.1124** |
| | (0.0514) | (0.0477) | (0.0434) | (0.0400) | (0.0393) | (0.0491) |
| TechRel | | 0.0132*** | 0.0123*** | 0.0058*** | 0.0060*** | 0.0060*** |
| | | (0.0016) | (0.0014) | (0.0012) | (0.0012) | (0.0012) |
| GDPpc | | | 0.3193*** | 0.1678 | 0.1766* | 0.1800* |
| | | | (0.1157) | (0.1073) | (0.1059) | (0.1071) |
| EmpGrowth | | | 0.4733 | 0.2568 | 0.2591 | 0.2668 |
| | | | (0.5191) | (0.4968) | (0.4919) | (0.4938) |
| ShPatents | | | -0.2007*** | -0.1718*** | -0.1606*** | -0.1629*** |
| | | | (0.0314) | (0.0318) | (0.0311) | (0.0322) |
| (d) shGreenEst | | | | 0.3102*** | 0.3005*** | 0.3020*** |
| | | | | (0.0973) | (0.0956) | (0.0959) |
| GreenPrevRTA | | | | 1.358*** | 1.389*** | 1.359*** |
| | | | | (0.2167) | (0.2146) | (0.2147) |
| GreenCapex × (d) ASH | | | | | -0.0362*** | |
| | | | | | (0.0132) | |
| GreenCapex × (d) RSH | | | | | | 0.0333** |
| | | | | | | (0.0138) |
| Observations | 233,470 | 233,470 | 233,470 | 233,470 | 233,470 | 233,470 |
| Log-Likelihood | -66,041.3 | -65,888.6 | -65,715.7 | -65,609.8 | -65,598.5 | -65,599.4 |
| Adjusted Pseudo $R^2$ | 0.03603 | 0.03820 | 0.04063 | 0.04211 | 0.04226 | 0.04225 |
| Tech*Year fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| State*Year fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Clustered (MSA) standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dep var: MSA entry in a green technological specialization. $GreenCapex$ and $BrownCapex$ are expressed as continuous variables. $(d)ASH$ and $(d)RSH$ are dichotomous variables equal to one if the share of, respectively, abstract and routine skills is above the national median. Explanatory variables are lagged by one year. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models are estimated through fixed effects logit estimators. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the MSA level. Technology-by-year fixed effects are built with 4-digit CPC codes, and State-by-year fixed-effects are included.

# Appendix C

# Appendix for Chapter 4

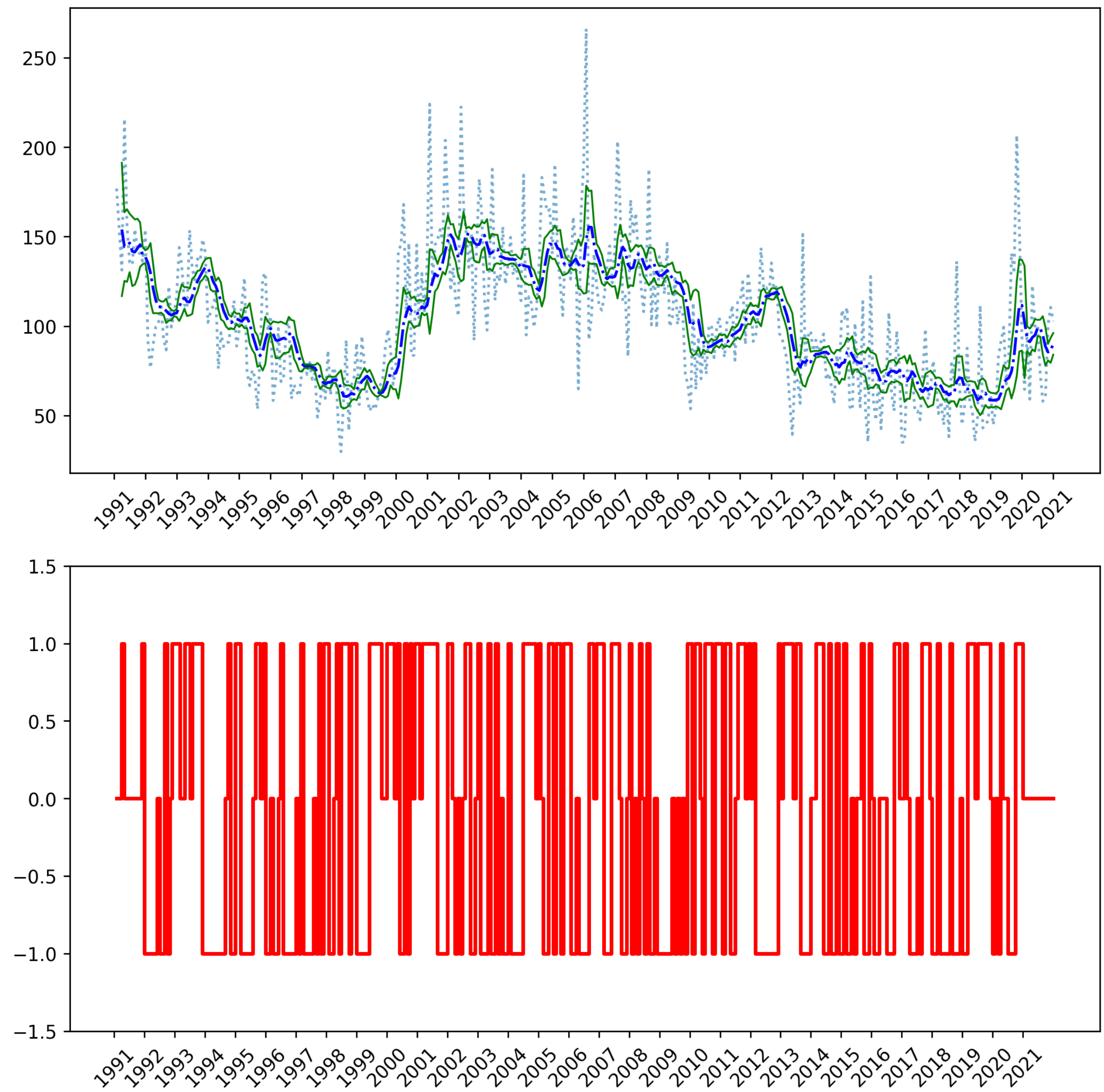



**Figure C.1:** Peak detection algorithm.
The dotted line plots the monthly series of CPU for Italy. In blue, I represent the moving average, and in green the threshold standard deviations for detecting peaks. In the bottom panel, events are flagged as 1 (peaks) or -1 (troughs). I only consider positive deviations (peaks) in the count of events for each time series derived.

**Table C.1:** Weighted Poisson Regression - Baseline estimates for CPU - Positive and negative policy stance.

| | Green | | Dirty | | Grey | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| CPUp | 0.6190* | 0.6487* | -0.5998 | -0.6014 | -1.773** | -1.793** |
| | (0.3597) | (0.3615) | (0.3840) | (0.4053) | (0.8453) | (0.8381) |
| CPUm | -0.5311 | -0.5518 | 1.078** | 1.074** | 2.574*** | 2.603*** |
| | (0.3538) | (0.3545) | (0.4188) | (0.4322) | (0.8620) | (0.8372) |
| GreenStock | 1.034*** | 1.017*** | 0.1459*** | 0.1573*** | -0.0278 | -0.0256 |
| | (0.0477) | (0.0427) | (0.0322) | (0.0331) | (0.1203) | (0.1203) |
| DirtyStock | 0.0171 | 0.0338* | 0.9404*** | 0.9401*** | 1.282*** | 1.282*** |
| | (0.0225) | (0.0192) | (0.0289) | (0.0260) | (0.1046) | (0.1041) |
| SPILLgreen | 1.331 | 1.428 | 1.358** | 1.429** | 3.299*** | 3.346*** |
| | (0.9559) | (0.9558) | (0.6264) | (0.6669) | (1.157) | (1.159) |
| SPILLdirty | -1.389 | -1.430 | -1.494** | -1.478** | -3.421*** | -3.437*** |
| | (1.047) | (1.046) | (0.6465) | (0.6800) | (1.126) | (1.133) |
| pre-sample mean | -0.0380*** | -0.0211 | -0.0771*** | -0.0647*** | -0.1033** | -0.1008* |
| | (0.0143) | (0.0130) | (0.0163) | (0.0156) | (0.0525) | (0.0558) |
| Emit | 0.0096 | 0.0045 | 0.0036 | -0.0006 | 0.8018* | 0.7961* |
| | (0.1049) | (0.1058) | (0.1008) | (0.0984) | (0.4110) | (0.4084) |
| EPS*Emit | 0.0705* | 0.0614 | 0.0819 | 0.0844 | -0.3819* | -0.3807* |
| | (0.0419) | (0.0412) | (0.0620) | (0.0554) | (0.2197) | (0.2163) |
| ShPatents | | -0.0438* | | -0.0582** | | -0.0159 |
| | | (0.0256) | | (0.0254) | | (0.0499) |
| Observations | 86,562 | 86,562 | 82,082 | 82,082 | 59,452 | 59,452 |
| RMSE | 0.77593 | 0.77660 | 0.71955 | 0.71647 | 0.37788 | 0.37771 |
| Country*Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Industry*Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Full Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. *CPUp* captures firm-level exposure to climate policy uncertainty with an environmentally-positive policy stance. *CPUm* instead captures environmentally-negative policy stance. *GreenStock* and *DirtyStock* are, respectively, the depreciated stocks of own-firm past green or dirty Triadic Patent Families. *SPILLGreen* and *SPILLDirty* are firm-level geographical spillovers to the focal firm in green and dirty technologies. *Emit* are country-sector-year emissions, and *EPS* is the index of Environmental Policy Stringency. *ShPatents* is the share of firm patents within its country-sector-year. All models contain dummy variables in case the past stock of green and (or) dirty patents is 0, which are always significant and not reported. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Models are estimated by Maximum-likelihood Poisson regressions, weighted by the average number of patents for each firm, over the period of observation. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the firm level.

**Table C.2:** Poisson Regression. CPU stances: Split sample analysis.

| | Excluding France | | | Excluding Italy | | | Excluding Spain | | | Excluding Germany | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Green | Dirty | Grey | Green | Dirty | Grey | Green | Dirty | Grey | Green | Dirty | Grey |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| CPUp | 0.6848** | -0.4028 | -1.460 | 0.4973 | -0.8001* | -2.020*** | 0.7551** | -0.3604 | -1.876*** | 0.5442 | -0.3826 | -0.3225 |
| | (0.3392) | (0.4092) | (0.9879) | (0.3690) | (0.4502) | (0.7339) | (0.3359) | (0.4096) | (0.6122) | (0.3909) | (0.6181) | (0.7653) |
| CPUm | -0.6244* | 0.4503 | 1.485 | -0.4372 | 0.8890** | 2.140*** | -0.6899** | 0.4775 | 1.975*** | -0.5086 | 0.4382 | 0.3334 |
| | (0.3380) | (0.3960) | (0.9903) | (0.3704) | (0.4458) | (0.7351) | (0.3382) | (0.4119) | (0.6107) | (0.3904) | (0.6020) | (0.7618) |
| Firms | 3229 | 3229 | 3229 | 3572 | 3572 | 3572 | 4007 | 4007 | 4007 | 1705 | 1705 | 1705 |
| Observations | 55,106 | 52,403 | 37,760 | 61,602 | 57,937 | 42,262 | 69,026 | 65,217 | 48,660 | 26,563 | 21,544 | 9,870 |
| Pseudo $R^2$ | 0.66901 | 0.77451 | 0.83826 | 0.67090 | 0.78022 | 0.82482 | 0.66006 | 0.76938 | 0.81445 | 0.59669 | 0.72441 | 0.62612 |
| RMSE | 0.86068 | 0.84612 | 0.54918 | 0.91030 | 0.95265 | 0.54967 | 0.87477 | 0.93573 | 0.53189 | 0.58320 | 0.80251 | 0.41412 |
| Country*Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Industry*Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Full Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. $CPUp$ captures firm-level exposure to climate policy uncertainty with an environmentally-positive policy stance. $CPUm$ instead captures environmentally-negative policy stance. All models contain full controls and dummy variables in case the past stock of green and (or) dirty patents is 0, which are always significant and not reported. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Models are estimated by Maximum-likelihood Poisson regressions. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the firm level.

**Table C.3:** Poisson Regression - Baseline estimates for CPU stances - Additional controls.

| | Green | | Dirty | | Grey | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| CPUp | 0.6803** | | -0.4285 | | -1.645** | |
| | (0.3099) | | (0.3878) | | (0.6528) | |
| CPUm | -0.6011* | | 0.5408 | | 1.924*** | |
| | (0.3151) | | (0.3913) | | (0.6860) | |
| PeaksP | | 0.1208*** | | -0.0182 | | -0.0523 |
| | | (0.0452) | | (0.0563) | | (0.0786) |
| PeaksM | | -0.0144 | | 0.0870 | | 0.2436*** |
| | | (0.0493) | | (0.0616) | | (0.0914) |
| EPU | -0.0211 | 0.0073 | -0.0374 | 0.0408 | -0.2377*** | -0.0474 |
| | (0.0384) | (0.0191) | (0.0569) | (0.0382) | (0.0794) | (0.0585) |
| Observations | 72,040 | 72,040 | 67,809 | 67,809 | 49,725 | 49,725 |
| Pseudo $R^2$ | 0.65696 | 0.65713 | 0.76813 | 0.76811 | 0.81637 | 0.81626 |
| RMSE | 0.86229 | 0.86229 | 0.92212 | 0.92288 | 0.50474 | 0.50629 |
| Country*Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Industry*Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Full Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Firm Size Dummy | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. $CPUp$ captures firm-level exposure to climate policy uncertainty with an environmentally-positive policy stance. $CPUm$ instead captures environmentally-negative policy stance. $PeaksP$ and $PeaksM$, respectively, represent the total number of events of positive or negative stance, over the total number of events detected by the peak-detection algorithm. $EPU$ is the firm-level exposure to Economic Policy Uncertainty. All models contain full controls and dummy variables in case the past stock of green and (or) dirty patents is 0, which are always significant and not reported. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Models are estimated by Maximum-likelihood Poisson regressions. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the firm level.

**Table C.4:** Poisson Regression - Baseline estimates for CPU stances - alternative construction for EPS.

| | Green | | Dirty | | Grey | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| CPUp | 0.6972** | | -0.4363 | | -1.518** | |
| | (0.3096) | | (0.3911) | | (0.7262) | |
| CPUm | -0.6125** | | 0.5286 | | 1.612** | |
| | (0.3099) | | (0.3848) | | (0.7278) | |
| PeaksP | | 0.1200*** | | -0.0091 | | -0.0483 |
| | | (0.0432) | | (0.0522) | | (0.0766) |
| PeaksM | | -0.0005 | | 0.1113* | | 0.2342** |
| | | (0.0461) | | (0.0593) | | (0.1058) |
| EPS | -0.1253** | -0.0665 | -0.0895 | -0.0267 | -0.2624* | -0.2159* |
| | (0.0515) | (0.0456) | (0.0668) | (0.0657) | (0.1371) | (0.1145) |
| Observations | 72,058 | 72,058 | 67,826 | 67,826 | 49,725 | 49,725 |
| Pseudo $R^2$ | 0.65697 | 0.65704 | 0.76804 | 0.76793 | 0.81585 | 0.81583 |
| RMSE | 0.86151 | 0.86161 | 0.91886 | 0.92111 | 0.50531 | 0.50689 |
| Country*Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Industry*Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Full Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Firm Size Dummy | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Cluster S.E. | Firm | Firm | Firm | Firm | Firm | Firm |

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. $CPUp$ captures firm-level exposure to climate policy uncertainty with an environmentally-positive policy stance. $CPUm$ instead captures environmentally-negative policy stance. $PeaksP$ and $PeaksM$, respectively, represent the total number of events of positive or negative stance, over the total number of events detected by the peak-detection algorithm. $EPS$ is the Environmental Policy Stringency index, constructed analogously to CPU. All models contain full controls and dummy variables in case the past stock of green and (or) dirty patents is 0, which are always significant and not reported. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Models are estimated by Maximum-likelihood Poisson regressions. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the firm level.

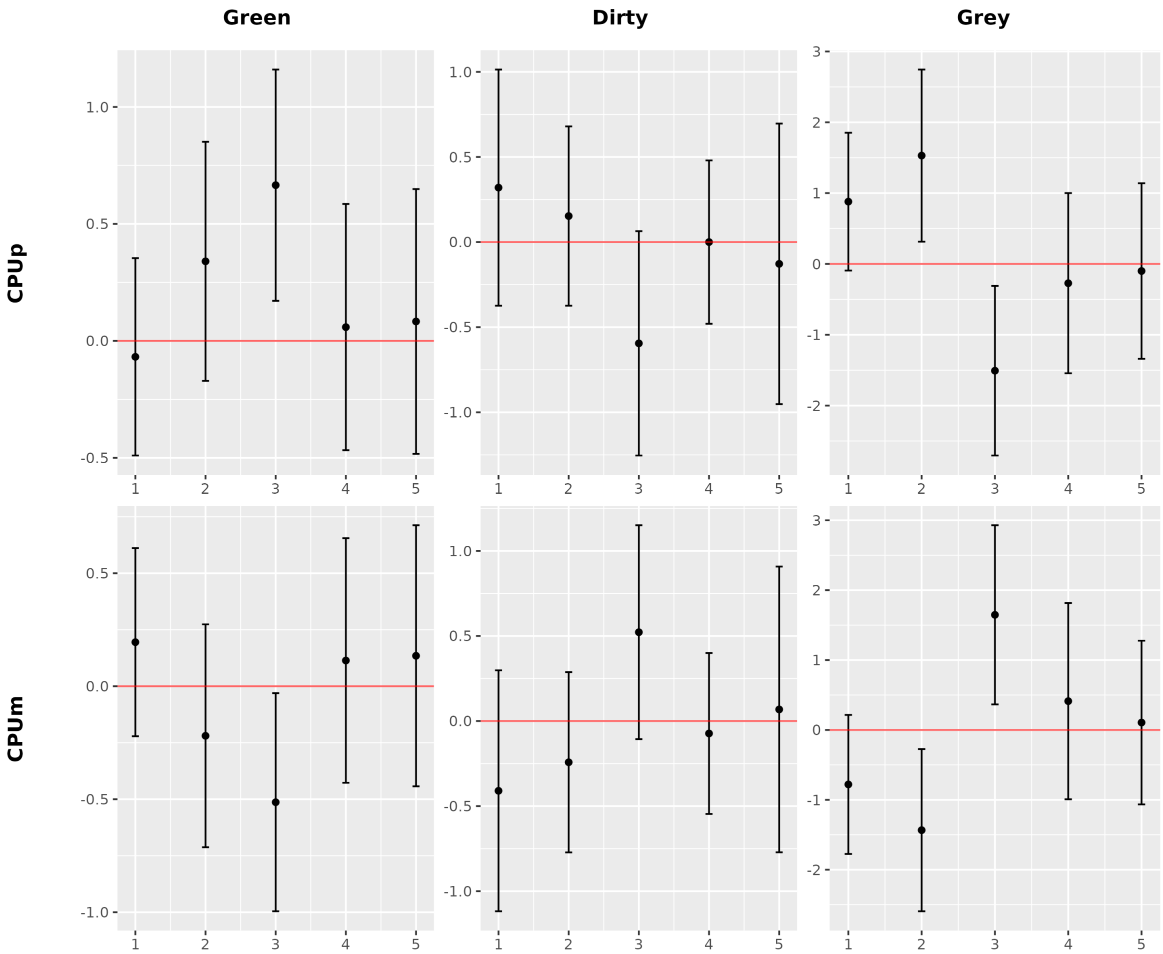



**Figure C.2:** Timing of different lags.
Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. *CPUp* captures firm-level exposure to climate policy uncertainty with an environmentally-positive policy stance. *CPUm* instead captures environmentally-negative policy stance. All models contain full controls and dummy variables in case the past stock of green and (or) dirty patents is 0, which are always significant and not reported. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Models are estimated by Maximum-likelihood Poisson regressions. Standard errors at a 5% level are plotted in the Figure.

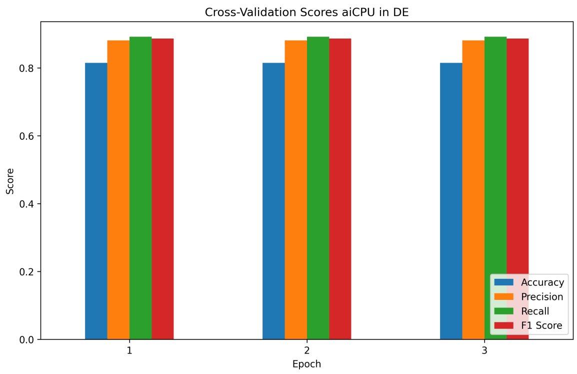
### An application of semi-supervised learning in the case of Climate Policy Uncertainty

I provide here an example of architecture for out-of-sample labelling of news data, testing the flexibility of LLM-based methods for labelling of policy uncertainty in newspapers' articles. The objective of this exercise is similar to those of semi-supervised label propagation algorithm (Iscen et al. 2019). In these setups, semi-supervised deep learning exploits a small number of human-curated labelled data, in cases where larger unlabelled data of similar nature are available. The case for this application builds on recent literature using artificial LLM labels as ground truth, and train semi-supervised algorithms on unlabelled data. The algorithm, rather simple in its nature, can be flexibly exploited in the case of policy uncertainty exercises, in which larger amounts of data from a diverse set of news archives can be added.

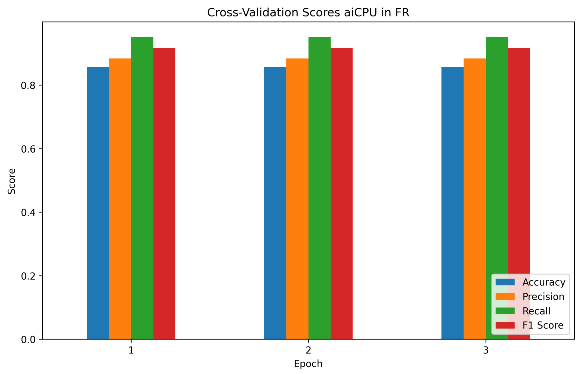
The architecture is based on recent literature employing LLM labelled dataset, in which a "teacher" algorithm is labelling data subsequently used by another "student" model (Pangakis and Wolken 2024). This approach, promising in social sciences, can be leveraged to reduce cost of labelling, and potentially achieving human-level quality (Gilardi et al. 2023). After labelling the sample of articles as explained in Section 4.3.1, I extract a random sample of articles amounting to 50% of the dataset, for each country (and therefore each language). In turn, I train a "student" model for each language. Leveraging the superior syntactic properties of tensor-based architectures (in comparison with word-occurrence models), I train Google's Bidirectional Encoder Representation from Transformers (BERT) on two training exercises. First, I train BERT for the prediction of true positives (intended as the LLM-labelled positives) in the sample of Climate Policy Uncertainty articles. Second, I train BERT in a multi-class labelling exercise, based on the direction of climate policy uncertainty labels. The training is performed on the BERT model pre-trained on a large corpus of multilingual data (Devlin 2018). The relevance of this exercise lies in the possibility to exploit labelled data, costly to obtain, to other news sources within the same language, over a number of different directions. In terms of parameters, BERT is a much smaller model, compared with the latest ChatGPT-4o. Scores for evaluation metrics, across three epochs, are reported in Figure C.3, for general CPU, and in Figure C.4 for $CPUp$ and $CPUm$. I report scores for accuracy, precision, recall and F1 scores.

The metrics suggest that the prediction of the binary outcome has a quite high predictive power, with F1 scores well above 0.8 in all languages. This result shows that teacher-student architectures are relevant for recognizing news articles about CPU and able to filter out effectively false negatives, also in a multi-lingual context. Trained models, much smaller and cheaper than LLM-based labelling could be applied to new data sources.

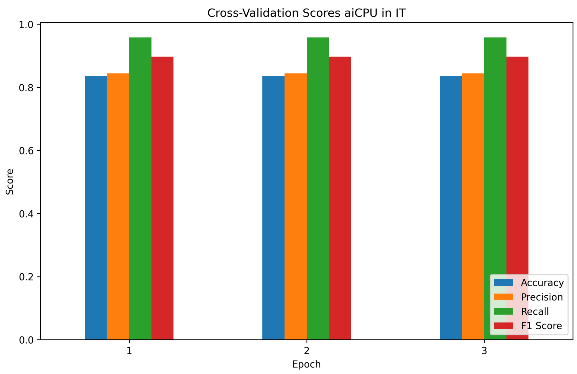
For what concerns the directions of effects, instead, the performance of multi-label is lower. While this attempt could be perfected, human intervention in prompt-tuning, and in the generation of artificial labels might be necessary for achieving a higher out-of-sample performance. However, precision and recall are around 0.6 for all languages, and the series derived from predicted labels (following the methodology explained in Section 4.3.1, correlate at the country level at 0.63 for $CPUp$ and 0.78 for $CPUm$. Overall, these results are very promising for the use of LLM, deep learning and teacher-student architectures in applications related to policy uncertainty, with the potential of making the labelling process cheaper, open source, and flexible and adaptable to novel data sources.



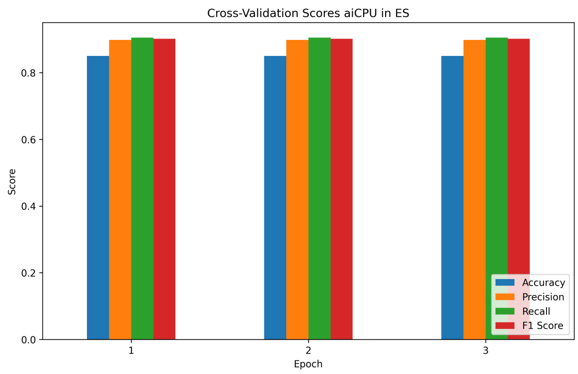

**(a)** Germany




**(b)** France

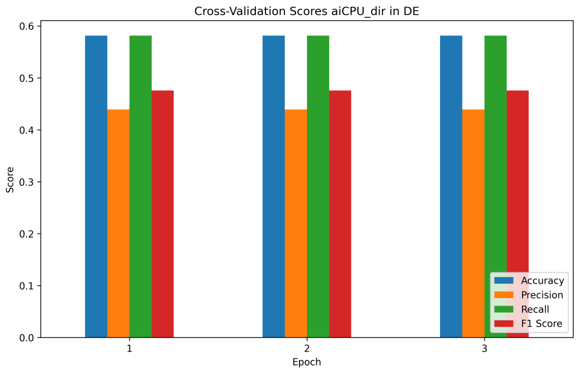



**(c)** Italy

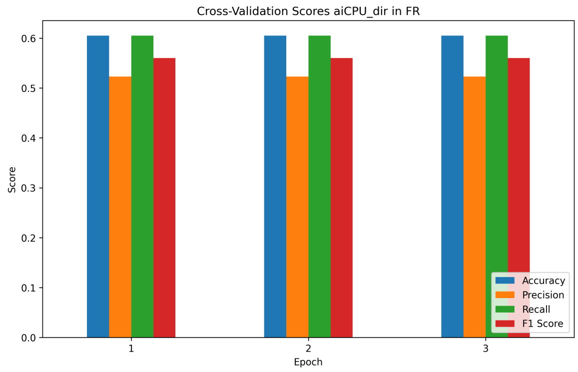



**(d)** Spain

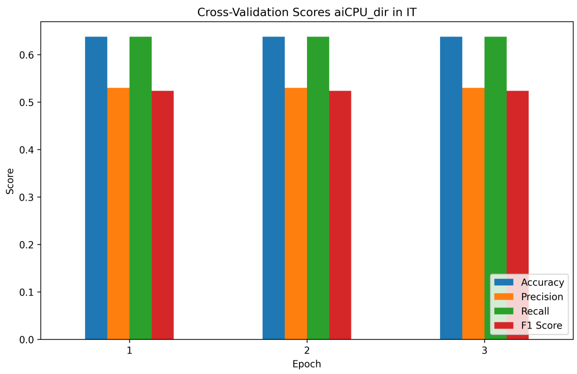
**Figure C.3:** Evaluation scores for three epochs in training - CPU labelling



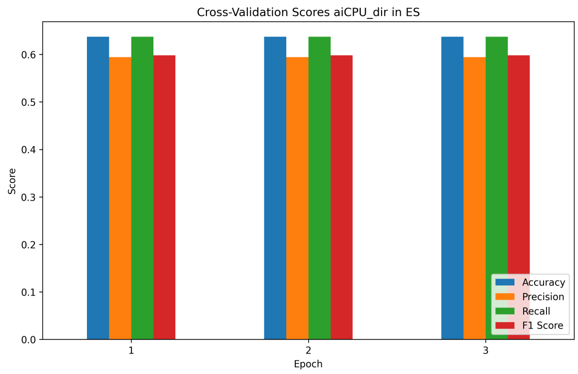

**(a)** Germany




**(b)** France




**(c)** Italy




**(d)** Spain

**Figure C.4:** Evaluation scores for three epochs in training - CPU directions

**Table C.5:** Technological categories for green patents

| Description | Codes | Lower digit purpose |
|---|---|---|
| Separation; Purification of air, liquids, or gases | B01D | Used in environmental control technologies such as filtration and water treatment. |
| Manufacture of iron or steel | C21B | Includes processes that reduce environmental impact in steel manufacturing. |
| Processing of pig-iron or steel | C21C | Focuses on methods to improve energy efficiency and reduce emissions. |
| Methods or apparatus for combustion using solid fuel | F23B | Related to cleaner combustion technologies for reduced pollution. |
| Combustion apparatus using fluid or pulverized fuel | F23C | Involves systems designed for efficient combustion with minimal environmental impact. |
| Cremation; Incineration of waste | F23G | Waste treatment technologies that minimize emissions. |
| Removal or treatment of combustion products | F23J | Techniques for controlling and reducing air pollution. |
| Furnaces; Kilns; Ovens | F27B | Energy-efficient designs and emissions reduction in industrial heating processes. |
| Chemical or physical processes, catalysts | B01J | Used in environmental applications like pollution control and energy-efficient processes. |
| Lubricating of internal combustion engines | F01M | Involves technologies to reduce environmental impact from lubricants. |
| Testing static or dynamic structures, mechanical structures | G01M | Environmental monitoring and testing technologies. |
| Magnetic or electrostatic separation of solid materials | B03C | Used in recycling and waste processing. |
| Fuels not derived from petroleum, including biofuels | C10L | Development of cleaner, renewable energy sources. |
| Exhaust apparatus for combustion engines | F01N | Technologies for reducing vehicular emissions. |
| Auxiliary equipment for ships, including pollution control devices | B63J | Marine pollution control. |
| Treatment of water, waste water, or sewage | C02F | Essential for water purification and environmental protection. |
| Materials for specific applications, including environmental uses | C09K | Involves chemicals for environmental protection like soil conditioners. |
| Water supply installations | E03C | Technologies improving water efficiency and management. |
| Sewage disposal | E03F | Waste management and pollution control. |
| Fertilizers, including those derived from waste | C05F | Involves recycling waste into environmentally friendly fertilizers. |
| Ships or other waterborne vessels; Equipment for ships | B63B | Marine environmental protection technologies. |
| Hydraulic engineering; Dams; Harbors | E02B | Involves managing water resources and protecting the environment. |

| | | |
|---|---|---|
| Cleaning streets; Removing snow, ice, or sand | E01H | Environmental management in urban areas. |
| Collecting or transporting refuse; Containers for refuse | B65F | Waste management technologies. |
| Animal feeding-stuffs; Non-medical feed additives | A23K | Related to sustainable agriculture and environmental protection. |
| Footwear | A43B | Involves materials and processes that reduce environmental impact in manufacturing. |
| Cleaning beaches or sea floor; Other cleaning operations | B03B | Environmental cleanup technologies. |
| Working of metal powder; Manufacture of articles from metal powder | B22F | Involves sustainable materials and processes in manufacturing. |
| Preparation or pre-treatment of plastics or other compositions | B29B | Environmental impact reduction in plastic processing. |
| Presses in general; Pressing | B30B | Includes environmentally friendly pressing methods in manufacturing. |
| Motor vehicles; Trailers | B62D | Technologies for reducing the environmental impact of vehicles. |
| Containers for storage or transport of articles or materials | B65D | Involves packaging technologies that reduce environmental waste. |
| Handling thin or filamentary material | B65H | Includes materials handling in an environmentally friendly way. |
| Manufacture of glass; Glassware | C03B | Technologies to reduce the environmental impact of glass manufacturing. |
| Cements; Concrete; Artificial stone | C04B | Focuses on sustainable construction materials. |
| Working-up of macromolecular substances | C08J | Recycling and environmental impact reduction in polymer processing. |
| Lubricating compositions | C10M | Development of environmentally friendly lubricants. |
| Production and refining of metals | C22B | Includes environmental technologies in metallurgy. |
| Preparation of fibers for spinning; Machines for cotton processing | D01G | Involves technologies for reducing environmental impact in textile manufacturing. |
| Fibrous raw materials for paper-making | D21B | Environmental impact reduction in the paper industry. |
| Production of cellulose by removing non-cellulose constituents | D21C | Cleaner technologies in cellulose production. |
| Pulp compositions; Impregnating materials | D21H | Includes environmentally friendly additives in paper production. |
| Cables; Conductors; Insulators | H01B | Technologies for energy-efficient and environmentally friendly electrical systems. |
| Electric discharge tubes; Gas-filled discharge tubes | H01J | Energy-efficient lighting technologies. |
| Processes or means for direct conversion of chemical energy into electrical energy | H01M | Focus on batteries and fuel cells, including environmentally friendly energy storage. |

| | | |
|---|---|---|
| Sterilization or disinfection techniques; Deodorization | A61L | Environmental impact reduction in medical technology. |
| Crushing, pulverizing, or disintegrating in general | B02C | Used in recycling and waste processing. |
| Disposal of solid waste | B09B | Technologies for efficient and environmentally friendly waste disposal. |
| Cracking hydrocarbon oils; Production of liquid hydrocarbon mixtures | C10G | Cleaner processes in the oil industry. |
| Reclamation of contaminated soil | B09C | Environmental technologies for soil remediation. |
| Signaling or calling systems; Preventing, indicating, or extinguishing fires | G08B | Includes environmental monitoring technologies. |
| Technologies for Adaptation to Climate Change | Y02A | |
| Climate Change Mitigation Technologies related to Buildings | Y02B | |
| Capture, Storage, Sequestration or Disposal of Greenhouse Gases | Y02C | |
| Climate Change Mitigation Technologies in ICT | Y02D | |
| Reduction of GHG Emissions in Energy Generation | Y02E | |
| Climate Change Mitigation Technologies in the Production or Processing of Goods | Y02P | |
| Climate Change Mitigation Technologies related to Transportation | Y02T | |
| Climate Change Mitigation Technologies related to Wastewater Treatment | Y02W | |
| Smart grids | Y04S | |
| Fuel Cells | H01M | |
| Wind motors | F03D | |
| Propulsion Of Electrically-Propelled Vehicles | B60L | |
| Tide or wave power plants | E02B9/08 | |
| Devices for producing mechanical power from geothermal energy | F03G4 | |
| Devices for producing mechanical power from solar energy | F03G6 | |
| Ocean thermal energy conversion | F03G7/05 | |
| Use of solar heat | F24J2 | |
| Production or use of heat, not derived from combustion using geothermal heat | F24J3/08, F26B3/28 | |
| Clean Filters | B01D46, B01D50, B01D35, B01D39, B01D41 | |
| Water Cleaning | E02B15 | |
| Construction waste managment | C04B18 | |

| | |
|---|---|
| Submerged units incorporating electric generators or motors characterized by using wave or tide energy | F03B13/10-26 |

Table C.6: Technological categories for grey patents

| **Description** | **Codes** |
|---|---|
| Idling devices | F02M3/00, F02M3/01, F02M3/02, F02M3/03, F02M3/04, F02M3/05 |
| Injection apparatus in combutions engines | F02M39, F02M41, F02M2041, F02M46, F02M43, F02M45, F02M47, F02M49, F02M51, F02M53, F02M55, F02M57, F02M59, F02M, F02M61, F02M63, F02M65, F02M67, F02M69, F02M71 |
| Adding non-fuel substances to fuel mix | F02M23, F02M25 |
| Electricity control and efficiency | F02D/41, F02B47/06 |
| Combustion technologies with mitigation potential | Y02E20/12, Y02E20/14, Y02E20/16, Y02E20/18, Y02E20/30, Y02E20/32, Y02E20/34 |

Table C.7: Technological categories for dirty patents

| **Description** | **Codes** |
|---|---|
| Internal-combustion piston engines | F02B |
| Controlling combustion engines | F02D |
| Cylinders, pistons, or casings for combustion engines; arragement of sealings in combustion engines | F02F |
| Supplying combustion engines with combustiles mixtures or constituents thereof | F02M |
| Starting of combustion engines | F02N |
| Ignition (other than compression ignition) for internal-combustion engines | F02P |
| Oil extraction and refining | C10G |
| Fuel | C10L1 |
| Separating Plants for Oil-related | B03B9/02, B03D2203/006 |
| Gas-turbine plants | F02C |
| Production of fuel gases by carburetting air or other gases | C10J |
| Hydraulic engineering | E02B |
| Steam engine plants (and similar) | F01K |
| Steam generation | F22 |
| Combustion apparatus or processes | F23 |
| Furnaces | F27 |
| Heat exchange in general | F28 |
| Lightning | F21H |

| | |
|---|---|
| Conventional and unconventional oil and gas exploratation and extraction | E21B B63B35/4413, B63B2035/442, B63B2035/448, B63B75/00, C09K8, C10L5/04, E02B17/0, E02B2017/003, E02B2017/004, E02B2017/005, E02B2017/006,E21B E02B2017/007, E02B2201, |
| Exploration and mining | B03B9/0, B03B1, B03D2203/006, B61D11, E21C |
| Gas conditioning | F25J3/0209, F25J3/0214, F25J3/0615, F25J3/061, |
| Solid Fuel conditioning | C10F, C10L |
| Coal to gas processes | C10B47, C10B49, C10B51, C10B53, C10B55, C10B57, C10J1, C10J3 |
| Hydrogen fuel production | C01B3/22, C01B3/3, C01B3/4 |

## Environmental Keywords

***Italian:***
((riscaldamento AND (globale OR del pianeta)) OR (emissioni AND NOT (obbligazionarie OR del tesoro)) OR energia OR energetic* OR ambiente OR ambiental* OR ecologic* OR climatic* OR carbonio OR gas serra OR effetto serra OR anidride carbonica OR CO2 OR metano OR CH4 OR inquinament* OR inquinante OR (ossid. AND di zolfo) OR SOx OR diossido di zolfo OR biossido di zolfo OR anidride solforosa OR SO2 OR ossido di azoto OR monossido di azoto OR NOx OR diossido di azoto OR biossido di azoto OR NO2 OR (particelle AND (fini OR solide OR piccole)) OR (particolate AND atmosferic* )) OR polveri sottili OR materiale particolato OR PM10 OR PM2.5 OR ozono OR rinnovabil* OR idroelettric* OR idraulic* OR eolic* OR fotovoltaic* OR emissioni OR biomass* OR (auto OR veicol* AND elettric*) OR ((auto OR motore OR alimentazione) AND ibrid*) OR (solar* AND NOT (crema OR eritema OR sistema OR trattamento OR ustione OR anno)))

***French:***
((energi* OR énergétiqu* OR environmenta*OR écologique* OR changement climatique OR réchauffement climatique OR climatiq* OR pollution OR pollutan* OR carbone OR gaz à effet de serre OR dioxyde de carbone OR co2 OR ch4 OR méthane OR oxyde de soufre OR so2 OR dioxyde de soufre OR sox OR oxyde d azote OR dyoxyde d azote OR particule fines OR PM2.5 OR PM10 OR ozone OR éolien* OR solair* AND NOT (crème OR système) OR photovoltaïque* OR hydraulique* OR biomasse OR (énergies renouvelables OR énergie renouvelable) OR (voitures OR voiture AND (électriques OR électrique OR hybride*))))

***Spanish:***
((energ* OR energétic* OR medio ambient* OR ecológic* OR cambio climático OR calentamiento global OR climatic* OR contaminación OR contaminante* OR polución OR carbono OR gases de efecto invernadero OR dióxido de carbono OR CO2 OR metano OR CH4 OR óxido de azufre OR SO2 OR dióxido de azufre OR SOx OR óxido de nitrógeno OR NOx OR dióxido de nitrógeno OR (partículas AND (finas OR en suspensión))

OR PM2.5 OR PM10 OR ozono OR eólic* OR (tecnología* OR panel* OR placa* OR central* AND solar*) OR fotovoltaic* OR (energía AND (hidráulica OR hidroeléctric*)) OR biomasa OR (energías AND (renovables OR verdes OR alternativas OR limpias)) OR (auto* OR coche* AND (eléctrico* OR híbrido*))))

***German:***
((klima* AND NOT (Geschäftsklima OR politisches OR wirtschaftliches OR Wirtschafts OR Regulierung OR regulatorisches OR Rechts OR rechtliches OR gesellschaftliches OR Gesellschafts)) OR Energiewende OR (Erneuerbare AND Energien AND Gesetz) OR EEG.Einspeisevergütung OR EEG.Umlage OR Klimapolitik OR Energiepolitik OR Umweltpolitik OR Lufteinhaltepolitik OR Luftreinhalteplan OR Umwelt OR ökologisch OR klimawandel OR Erderwärmung OR globale Erwärmung OR Umwelt* OR Energie* OR Kohlenstoff* OR Treibhausgas* OR THG* OR Kohlendioxid* OR CO2* OR Methan* OR CH4* OR Schadstoff* OR Umweltverschmutzung* OR Luftverschmutzung* OR verschmutz* OR schwefeloxid* OR SOx OR Schwefeldioxid* OR SO2* OR Stickoxid* OR NOx OR Stickstoffdioxid* OR NO2* OR (Partikel* AND (Fein OR Feinpartikel OR Feinstaub)) OR PM2.5 OR PM10* OR Ozon* OR erneuerbar* OR Hydro* OR Windenergie* OR Windpark* OR Windkraftanlage* OR Photovoltaik* OR PV OR Solar* OR Biomasse* OR (Elektrofahrzeug* OR Elektroauto* OR E-auto*) OR (Hybridfahrzeug* OR Hybridauto*)))

## Policy Keywords

***Italian:***
((politica AND NOT monetaria) OR regolament* OR legislazion* OR legge OR tasse OR canon* OR (standard AND NOT (& OR and OR e OR poors OR poor's)) OR certificat* OR certificazion* OR sussidi OR sussidio OR sovvenzion* OR ETS OR Sistema ES OR feed-in-tariff* OR conto energia OR (scambio AND di quote) OR regime di scambio OR sistema di scambio OR decarbonizzazione OR effetto serra OR cap and trade OR mercato dei diritti di emissione OR (mercato AND (dell OR di AND emission*)) OR (etichett* AND (ambiental* OR ecologic*)) OR eco-etichett* OR eco-label OR normative OR normativa)

***French:***
((politiq AND NOT monétaire) OR réglementation* OR lois OR loi OR redevance* OR tax* OR impôt* OR norme* OR tarification* OR tarif de rachat OR certificat* OR subvention* OR ETS OR (marché AND d emissions) OR droit* à polluer OR système d échange OR SEQE)

***Spanish:***
((política AND NOT monetaria) OR regulación* OR ley OR leyes OR impuesto* OR estándar* OR tarifa de alimentación OR certificado* OR subsidio* OR (mercado AND de emision*) OR derecho* OR contaminar OR sistema de comercio OR ETS)

***German:***
((politik AND NOT geld) OR richtlinie* OR reform* OR regulierung* OR vorschrift* OR gesetz* OR gebühr* OR abgabe* OR maßnahme* OR steuer* OR standard* OR zertifikat* OR subvention* OR preisgestaltung OR emissionshandel OR ETS OR einspeisetarif* OR einspeisevergütung* OR handelssystem* OR cap and trade OR (label OR kennzeichen AND umweltzeichen OR umweltabzeichen) OR umlage)

## Uncertainty Keywords

***Italian:***
(può OR potrebbe OR probabile OR probabilmente OR possibile OR possibilmente OR potenziale OR potenzialmente OR immaginare OR assumere OR assunzione OR credere OR sostenere OR stimare OR ipotesi OR ipotetico OR speculare OR speculazione OR sospettare OR supporre OR aspettarsi OR dubbio OR dubitare OR dubbioso OR incerto OR incertezza OR sconosciuto OR non familiare OR discutibile OR discutibilmente OR forse OR sembrare OR apparentemente OR improbabile OR nessun indizio OR nessuna prova OR nessuna idea)

***French:***
(peut OR pourrait OR probable OR probablement OR possible OR possiblement OR potentiel OR potentiellement OR imaginer OR supposer OR supposition OR croire OR prétendre OR estimer OR hypothèse OR hypothétique OR spéculer OR spéculation OR suspecter OR s'attendre à OR doute OR douter OR douteux OR incertain OR incertitude OR inconnu OR non familier OR discutable OR discutablement OR peut-être OR sembler OR apparemment OR improbable OR aucun indice OR aucune preuve OR aucune idée)

***Spanish:***
(puede OR podría OR probable OR probablemente OR posible OR posiblemente OR potencial OR potencialmente OR imaginar OR asumir OR suposición OR creer OR sostener OR estimar OR hipótesis OR hipotético OR especular OR especulación OR sospechar OR suponer OR esperar OR duda OR dudar OR dudoso OR incierto OR incertidumbre OR desconocido OR no familiar OR discutible OR discutiblemente OR quizás OR parecer OR aparentemente OR improbable OR ningún indicio OR ninguna prueba OR ninguna idea)

***German:***
(kann OR könnte OR wahrscheinlich OR möglich OR möglicherweise OR potenziell OR potentiell OR vorstellen OR annehmen OR Annahme OR glauben OR behaupten OR schätzen OR Hypothese OR hypothetisch OR spekulieren OR Spekulation OR verdächtigen OR vermuten OR erwarten OR Zweifel OR zweifeln OR zweifelhaft OR unsicher OR Unsicherheit OR unbekannt OR nicht vertraut OR fragwürdig OR fraglich OR vielleicht OR scheinen OR anscheinend OR unwahrscheinlich OR kein Anzeichen OR kein Beweis OR keine Ahnung)

## Sample titles of articles flagged as CPU plus

- L'Europa e la sfida dell'energia
- Il futuro dei Verdi
- Europa è ora di riprendere il cammino
- «Sui rigassificatori intervenga il governo» Scaroni: mi preoccupa il prossimo inverno, ma ci stiamo attrezzando perevitare il peggio
- Bersani «Avanti con le liberalizzazioni di tv ed energia I tagli per risanare il bilancio non si spalmano»
- Rivoluzione energetica contro la crisi
- Rifiuti, Amaie Energia accetta la sfida
- Ue, l'Italia spinge sull'Accordo commerciale per i beni ambientali
- "La infraestructura verde es un símbolo para el nuevo modelo de ciudades"
- La energía solar sale a flote
- «En 10 años ya no podremos invertir el calentamiento»
- «Hay que proteger el paisaje y a los paisanos»
- Blair asegura que ignorar el cambio climático tendrá consecuencias desastrosas
- Lo hemos hecho posible. Ahora tú decides
- «Abrir los mercados no es el nirvana»
- Es hora de tomar en serio el cambio climático
- 2013: le retour en force de l'Europe?
- Une étude de lAPPA précise limpact de la pollution sur la mortalité et la morbidité
- Mis en œuvre avec pragmatisme, un "Green Deal" européen a le potentiel de remodeler l'économie du continent
- L'éolien français manque de souffle
- Nucléaire, éolien... Que proposent Emmanuel Macron et Marine Le Pen en matière dénergie?
- Le contre-modèle américain
- Aérien: la Commission européenne planche sur une taxe kérosène

## Sample titles of articles flagged as CPU minus

- Tutti i rischi dell ambientalismo
- «Meno tasse, meno regole, più sicurezza» Colloquio con Bush, oggi l'insediamento alla Casa Bianca
- Enigma Trump: benvenuti nell'era dell'incertezza
- La fiducia nell'Opec dipende da Mosca
- Economia italiana stabile ma pesa l'incertezza politica europea
- Che cosa succede se gli Stati Uniti abbandonano l'accordo sul clima
- Allarme smog nel giorno dell afa
- La OMC y el futuro del medio ambiente

- Bush propone el mayor aumento del gasto militar desde la era de Reagan
- Un choque de titanes del petróleo en el peor momento posible
- EE UU y China suavizan sus controles medioambientales por la crisis del coronavirus
- Riesgos de catástrofe global
- La política energética de López Obrador provoca incertidumbre en el sector de las renovables
- Écologie et amateurisme
- Comment les Verts ont disparu dune campagne pourtant marquée par lécologie
- Lécologie nest pas morte, cest lécologie politique qui nexiste plus
- Le protocole de Kyoto est moribond, achevons-le !
- Les climato-sceptiques à lassaut du Giec
- Il y a un vrai problème autour de la capacité des Etats en développement à réduire la déforestation
- «La compensation carbone ne doit pas servir à se dédouaner»
- «Les gilets jaunes, symptôme dun peuple qui refuse un monde en perpétuelle accélération»
- Taxe carbone: pourquoi il ne faut plus laugmenter, et même la diminuer!,
- «Greta Thunberg, icône dun écologisme naïf»