SCUOLA DI DOTTORATO
UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Dipartimento di / Department of

Economia, Metodi Quantitativi e Strategia D'impresa

Dottorato di Ricerca in / PhD program Economics and Statistics Ciclo / Cycle XXXV

Curriculum in (se presente / if it is) Economics

Essays on Energy Economics and Econometrics

Cognome / Surname Romagnoli Nome / Name Matteo

Matricola / Registration number 854282

Tutore / Tutor: prof. Matteo Manera

Supervisors: prof. Matteo Manera and prof. Anna Cretì

Coordinatore / Coordinator: prof. Matteo Manera

ANNO ACCADEMICO / ACADEMIC YEAR 2022/2023

Introduction

This thesis consists of two independent essays that empirically explore different research questions at the intersection of energy economics and climate change. The energy sector is responsible for about three quarters of human greenhouse gas emissions. [1] It is therefore not surprising that numerous policies have been adopted around the world to promote a safe and equitable transition to a net-zero emissions energy system. [2] The underlying idea that connects the following essays is that applied econometric studies and policy evaluation studies can play a key role in shedding light on the realised effects of these policies and thus provide a valuable tool to guide the decarbonisation of the energy sector.

The first chapter examines the relationship between electricity liberalisation and the direction of technological change in the electricity sector. Technological change has long been identified as a necessary condition for transitioning to a net-zero emissions energy system while maintaining or improving current living standards. In this chapter, I show that electricity liberalisation can play an important role in steering innovation away from "dirty" technologies for electricity generation, transmission and distribution and towards "cleaner" alternatives. Using patent-level quality indicators, I also find that "clean" energy patents developed in liberalised energy markets tend to be more exploratory (i.e. less incremental), while the same effect is not observed for "dirty patents".

In the second chapter, a join work with Anna Cretì and Marzia Sesini, we examine the impact of bioethanol support in France on wheat consumption and wheat prices and assess its implications for food security. A clear understanding of the relationship between conventional biofuels and food security is important to determine the extent to which these fuels can play a role in decarbonising the transport sector. To date, however, there are few empirical studies examining the relationship between biofuel support policies and food security in European countries. On the one hand, we find that biofuel support in France in the early 2000s led to a significant increase in the amount of wheat used outside the food sector, resulting in a significant decrease in the share of wheat used in the food sector. On the other hand, we find no evidence of an increase in wheat prices. If, as economic theory predicts, higher demand for wheat in the fuel sector puts upward pressure on wheat prices, in our analysis this effect was likely overshadowed by the high volatility of wheat prices worldwide due to the global food crisis of 2007/2008.

[1] https://ourworldindata.org/ghg-emissions-by-sector

[2] (IEA (2022), World Energy Outlook 2022, IEA, Paris https://www.iea.org/reports/world-energy-outlook-2022, License: CC BY 4.0 (report); CC BY NC SA 4.0 (Annex A))

Contents

Chapter 1

Clean Sweep: Electricity Liberalisation and the Direction of Technological Change in the Electricity Sector

Abstract

This paper examines the impact of electricity liberalisation on the direction of technological change in the electricity sector. To this end, I use data on patents filed over the period 1990-2017 related to electricity generation, transmission and distribution in combination with a set of patent-level quality indicators and an instrumental variable approach. The results show that electricity liberalisation leads to the development of less dirty patents and more clean energy technologies. Clean energy innovations developed in more competitive electricity markets also tend to be more exploratory in nature, as they rely more on knowledge spillovers from other technological fields. The same is not true for dirty technologies. Of the clean energy technologies considered in the analysis, innovation in energy storage technologies appears to benefit most from increased competition in the electricity sector.

Keywords: Clean-energy Technologies, Electricity Liberalisation, Climate Change, Patent Data. **Jel Codes**: L94, 031, Q42, Q55, Q01

1.1 Introduction

Technological change and innovation are key to achieve a secure and just transition to a net-zero emissions energy system (IEA, 2022). In light of this, economists have long been interested in understanding what factors can reduce path dependence in polluting technologies and steer technological change towards cleaner alternatives (Aghion et al., 2023; Aghion et al., 2016; Acemoglu et al., 2012). In this paper, I investigate whether electricity liberalisation can have this effect for technologies related to electricity generation, transmission and distribution. To this end, I examine the impact of increased competition in the electricity sector on the number of clean and dirty patents developed in the industry and on the exploratory or exploitative nature of these patents.

The question of whether and how electricity liberalisation affects the direction of technological change in the electricity sector is particularly important given the lively debate on the regulation and design of the electricity market in recent years. In the European Union, this debate was recently reignited by the Russian invasion of Ukraine and the subsequent electricity price crisis (ACER, 2022). In the United States, the Texas blackout in 2021 triggered a similar discussion, focusing in particular on the role of electricity liberalisation in the failure of the power system (e.g., Mays et al., 2022).

The main data sources for the analysis are the European Patent Office's Patent Statistical database (PATSTAT) and the OECD Patent Quality Indicators Database. The PATSTAT database contains a wealth of information on worldwide patents, including but not limited to the country of origin of the inventor(s), the year in which the patent was first filed and the technological fields to which the patent is assigned. This information is used to create a country-level panel dataset for clean and dirty innovations related to electricity generation, transmission and distribution filed between 1990 and 2017. Clean energy innovations are identified using the Cooperative Patent Classification code Y02E ("Reduction of Greenhouse Gas Emissions, Related to Energy Generation, Transmission or Distribution"). Dirty innovations are identified as those that belong to class H02 (Generation; Conversion or Distribution of Electric Power) but are not assigned to class Y02E (Popp et al., 2022).

The list of all technologies included in technology class Y02E can be found in Appendix A.1. Part of the contribution of this paper is that it focuses on a broader range of clean energy technologies than is typically the case in the literature on liberalisation and green innovation in the electricity sector. This literature focuses mainly on renewable energy innovations, while the Y02E class also includes other technologies that can contribute to the decarbonisation of the energy sector, such as those related to energy storage or combustion technologies with mitigation potential. However, it is important to note that two of the technological fields categorised under the Y02E code are likely to have a relationship with electricity liberalization that is distinctly different from the others These are nuclear energy (Y02E 30) and hydrogen-related technologies (Y02E 60/30 and Y02E 60/50).

Nuclear energy is likely to suffer from the liberalisation of the electricity market (Markard et al., 2020, Newbary 2010). According to IEA (2019a): *"The biggest barrier to new nuclear construction is mobilising investment. Plans to build new nuclear plants face concerns about competitiveness with other power generation technologies and the very large size of nuclear projects that require billions of dollars in upfront investment. Those doubts are especially strong in countries that have introduced competitive wholesale markets."* [IEA, 2019a, pag 4]. As for hydrogen-related technologies, it seems difficult to imagine that changes in electricity market regulation during the period of under analysis (1990-2017) could have played a significant role in their development. The hydrogen market is still considered nascent today, and despite growing interest in the use of hydrogen in the power sector, its deployment remains negligible to date (IEA, 2022a, IEA 2019b). Data on hydrogen demand paint a similar picture. Demand for hydrogen continues to be driven by traditional refining and industrial applications. New applications such as transport, high-temperature industrial heating, energy and buildings account for less than 0.1% of global demand, with road transport being the largest contributor to hydrogen demand among these new applications [1]. Due to these considerations, I do not include in the main analysis patents related to nuclear power and hydrogen technologies. Throughout the remainder of this paper, the term "clean energy technologies" refers to all Y02E technologies except those in these technological fields. For the sake of completeness, Appendix A.8 shows and discusses the results when these patent applications are also included in the sample.

The OECD Patent Quality Indicators Database contains a wide range of indicators calculated at the patent level that capture various important characteristics of the underlying technologies (Squicciari et al., 2013). A subset of these indicators is used in the following analysis to proxy the exploratory or exploitative nature of patents related to electricity generation, transmission or distribution. This allows me to investigate whether electricity liberalisation has led to the development of more exploratory innovations; a pattern that would suggest less path dependence in the development of new technologies in the electricity market.

The use of patent-level indicators to investigate these types of questions is now well established in the literature (see e.g. Acemoglu et al., 2022; Popp et al., 2022; De Noni et al.,

[1] Source: https://www.iea.org/energy-system/low-emission-fuels/hydrogen

2021; Barbieri et al., 2020). However, Higham et al. (2021) show that results obtained with different patent indicators can be inconsistent and therefore stress the importance of carefully selecting the best indicators in relation to the research question of the analysis. To measure the exploratory nature of the patents in the sample, I follow the literature and choose two widely used indicators that focus on the breath, complexity and diversity of the search space of these patents: the radicalness index (Shane, 2001) and the originality index (Trajtenberg et al., 1997). These indicators are described in detail in Section 1.3.

The search space of a patent can be defined as the totality of the knowledge inputs on which the invention is based, i.e. the set of sources used to develop the invention. A narrow search space is characterised by knowledge inputs from few technological fields and is generally associated with the development of incremental or exploitative innovations (Squicciarini et al., 2013, Quintana-García et al., 2008; Shane, 2001; March, 1991). A diverse search space, on the other hand, uses knowledge from a variety of fields and is thus evidence that the patent builds on many strands of knowledge. An established body of literature in innovation theory relies on the radicalness index and/or the originality index to measure the breadth and diversity of a patent's search space and to identify radical and exploratory innovations (Acemoglu et al., 2022, De Noni et al., 2021, Barbieri et al., 2020; Squicciarini et al., 2013; Quintana-García et al., 2008; Shane, 2001; Trajtenberg et al., 1997). The underlying assumption is that exploratory innovation requires the combination of diverse ideas. Furthermore, the characteristics measured by the radicalness and originality indices are not only good indicators for capturing the exploratory nature of a patent, but also have relevant policy implications themselves. A diverse search space has been shown to be particularly important for the development of green innovations, as these often require more diverse knowledge inputs (Barbieri et al., 2020; De Marchi, 2012). Moreover, the use of "knowledge accumulated in other technological fields is an often overlooked factor that is of vital importance to technology policy, because the benefits of spillovers can be harnessed at relatively low cost and can avoid or reduce the need for additional R&D" (IEA, 2020, p. 338). For example, knowledge spillovers from the production of silicon for microprocessors played an important role in the development of solar PV panels, and the carbon anode used in Li-ion batteries today was first developed by petrochemical companies (IEA, 2020). [2]

To uncover the causal effect of interest, I rely on an Instrumental Variable (IV) approach. The proposed strategy follows Nicolli and Vona (2019) and uses regulation in telecommunications as an instrument for electricity regulation. The intuition behind this approach is that

[2] Another example cited in the IEA report is the development of the first gas turbine jet engine in 1939, based on government-funded military R&D in the United Kingdom.

the liberalisation of the telecommunications sector preceded that of the electricity market and is generally considered as instrumental in giving momentum to the latter (Nicolli and Vona, 2019; Pollitt 2012; Joskow, 2008). Høj et al. (2006) provide evidence of this kind of spillover effects between product market reforms in different sectors and Nicolli and Vona (2019) argue that widespread liberalisation signals greater commitment to this reform and greater efforts on the part of policymakers and public authorities. Consequently, extensive liberalization across various sectors is expected to correlate with more effective product market reforms. At the same time, the liberalisation of the telecommunications sector can be considered as independent of the direct lobbying power of actors in the energy sector (Nicolli and Vona, 2019) and technological developments in energy supply technologies.

The results of the empirical analysis show that electricity liberalization increases the number of clean energy patents in electricity generation, transmission and distribution, while the number of dirty patents in these fields decreases. This reform also makes clean energy patents more exploratory, as they rely more on the use of knowledge spillovers from other technological fields. The same pattern is not observed for dirty patents. Among the Y02E technological fields included in the analysis, energy storage technologies appear to benefit most from increased competition in the electricity market. These results suggest that electricity liberalisation can play an important role in changing the direction of technological change in the electricity sector, steering it away from dirty technologies and towards more and more exploratory clean energy technologies.

The remainder of the chapter is organised as follows. Section 1.2 reviews the relevant literature and presents the research hypotheses. Section 1.3 describes the data used and Section 1.4 discusses the empirical strategy. Section 1.5 presents the results and Section 1.6 concludes.

1.2 Liberalisation and Innovation in the Electricity Sector

Previous empirical studies on the relationship between electricity liberalisation and innovation have mainly focused on the impact of this reform on the amount of new technologies developed in the market (e.g. the number of patents) or the inputs used in the research process (e.g. investment in R&D). A useful way to categorise this literature is to distinguish between studies that focus on clean energy technologies and those that look at all innovations developed in the electricity sector.

When focusing on all innovations developed in the market, the results are mixed. A first strand of literature suggests that overall R&D spending and patent activity decline after electricity liberalisation, e.g. Sanyal and Gosh (2013), Sterlacchini (2012), Jamasb and Pollitt (2008) and Dooley (1998). A more recent study by Cambini et al. (2016) finds a positive effect of electricity market reform on both R&D spending and patenting. According to their results, these effects are driven by the vertical unbundling of large firms in the electricity sector. Wang and Mogi (2017) find that Japanese electric utilities filed more patents after liberalisation, but in conjunction with a decrease in R&D spending. Finally, Marino et al. (2019) find an inverted-U relationship between electricity liberalisation and the number of patents filed in the electricity sector.

In contrast, the literature focusing specifically on clean-energy innovation consistently shows a positive effect of electricity liberalisation on clean energy innovation (Popp, 2019). Nesta et al. (2014) show that electricity liberalisation has a positive impact on renewable energy innovation, which is driven by the reduction of barriers to entry. Nicolli and Vona (2016) extend this result and show that the effect of electricity liberalisation is heterogeneous across different renewable energy technologies and is particularly significant for solar, wind and, to a lesser extent, waste technologies. Jamasb and Pollitt (2011) find that electricity-related patents in non-nuclear and renewable technologies increased after electricity liberalisation, but also observe a permanent decline in R&D in the post-liberalisation period, which they suggest may have a long-term negative impact on technological progress in the sector. Jacobsson and Bergek (2004) provide anecdotal evidence of the role that new entrants have played in the development of the German wind energy sector. Finally, in reviewing of the literature on the diffusion of renewable energy technologies, Negro et al. (2012) conclude that incumbents' preference for incremental and near-to-market innovation is a significant barrier to innovation in renewable energy technologies and to their diffusion.

This paper contributes to the above literature by explicitly categorising technologies related to electricity generation, transmission or distribution into dirty and clean innovations and comparing the impact of electricity liberalisation on both categories. The analysis is therefore not only related to the literature on electricity liberalisation and innovation, but also to the literature on directed technological change and the environment (Aghion et al., 2023; Aghion et al., 2016; Calel and Dechezleprêtre, 2016; Acemoglu et al., 2012; Popp et al., 2002). This literature has shown how the direction of technological change can change from dirty to clean technologies in response to public policies such as taxes or subsidies (Acemoglu et al., 2012). For example, Aghion et al. (2016) find that in response to higher tax-adjusted fuel prices, clean innovations in the automotive sector increase while dirty innovation decreases. Calel and Dechezleprêtre (2016) find that the EU Emission Trading System has increased low-carbon innovation in regulated firms. Particularly relevant in this context is the work of Aghion et al.

(2023), which finds that firms in the automotive sector innovate more in clean technologies when they are exposed to pro-environmental attitudes and that this effect is particularly strong for higher levels of product market competition.

In studying the effect of liberalisation on the direction of technological change in the electricity sector, it is important to underline that this sector is one of the best examples of a large technical system (Hughes, 1987). Joerges (1998) defines large technical systems as complex systems of physical structures and machineries, integrated over space and time and supporting other technical systems. The strong interactions between the different components of a large technical system, as well as the interactions between the system itself and the other technical systems it supports, mean that innovation in these environments is often path-dependent and characterised by incremental improvements that rely heavily on knowledge already available in the field (Negro et al. 2012; Markard and Truffer 2006; Hughes, 1987).

In such an environment, liberalisation can be expected not only to affect the number of clean and dirty technologies developed in the electricity market, but also to change the nature of these technologies making them less exploitative and more exploratory or, in other words, less incremental and more radical. As discussed in Section 1.1, I follow much of the existing literature on this topic and proxy the exploratory nature of an innovation with the breath and complexity of its knowledge base (De Noni et al., 2021, Barbieri et al., 2020; Squicciarini et al., 2013; Quintana-García et al., 2008; Shane, 2001; Trajtenberg et al., 1997).

Markard and Truffer (2006) present evidence that electricity liberalisation leads to the development of more radical clean-energy innovations by analysing a series of 44 interviews conducted in more than 30 utilities from Germany, the Netherlands and Switzerland. Their study concludes that the liberalisation of the electricity sector may be one of these external drivers which increases the diversity of clean-energy technologies' search space and therefore leads to more exploratory clean-energy innovations. [3]

A first channel through which electricity liberalisation can lead to more exploratory innovation is the lowering of barriers to entry, which allows new entrants in the market. The importance of new entrants in developing radical innovations is widely recognised in the economics literature (Akcigit and Kerr, 2018; Klepper 1996; Winter, 1984). The model proposed by Klepper (1996) predicts that low competition reduces the diversity of product R&D and could therefore lead to a narrower search space in the innovation process. In contrast, the entry of new players is expected to lead to a broader approach to R&D and a wider search space,

[3] It is worth noting that their case study is based on fuel cell innovation strategy, a technology that is excluded from the main sample used in this analysis. See Appendix A.8 for the results using a sample which includes also hydrogen related technologies.

especially in the case of the electricity sector, as these players are not tied to the traditional large-scale plants and technologies of incumbents (Nicolli and Vona, 2016; Nesta et al., 2014).

A second channel through which electricity liberalisation may lead to less incremental innovation in the electricity sector is by triggering a shift in the actors responsible for the innovation activity. Dolphin and Pollitt (2020) use patent data from the UK and show that innovation activity shifted from regulated monopolists to electric equipment manufacturers after the reform of the electricity market. As electric equipment manufacturers are less tied to traditional generation technologies, we can expect them to be less affected by path dependence in dirty technologies. Furthermore, new entrants will change the demand faced by electric equipment manufacturer as they are likely to demand new types of technologies (Sanyal and Gosh, 2013). This will provide a further incentive for electric equipment manufacturers to make their innovations less path-dependent and to widen their search space.

Finally, more competition in the electricity market may also have an impact on the kind of innovations developed by former monopolists. Incumbents' preference for incremental innovation has been identified as a potential barrier to the development and diffusion of exploratory clean energy technologies (Negro et al., 2012). The literature on the relationship between innovation and product cannibalisation is relevant to this point, as radical clean energy technologies are often competence-destroying for incumbents in the electricity market and have the potential to threaten their existing asset base (Nesta et al., 2014). When cannibalisation is an issue, competitive pressure is essential to create incentives for incumbents to develop radical innovations (Conner, 1988; Reinganum, 1983). Therefore, technologies developed by electric utilities can be expected to become less incremental and more exploratory as competition in the electricity market increases. Privatisation may also have a similar effect by changing the strategy of incumbents so that their innovations become less incremental. An example in this direction is the empirical evidence suggesting that investor-owned electric utilities respond more to renewable energy policies than state-owned utilities (Nicolini and Tavoni 2017; Delmas and Montes-Sancho 2011; Carley 2009).

Drawing from the discussion in this section, the research hypothesis of this study is that electricity liberalization reduces path dependence in dirty technologies for electricity generation, transition and distribution and redirects innovation towards cleaner alternatives. More competition in the electricity market is also expected to foster the development of more exploratory energy technologies.

1.3 Data and Descriptive Statistics

Using the PATSTAT database, I collected data on patent applications related to the generation, transmission or distribution of electricity, developed in OECD countries and filed at the European Patent Office (EPO) between 1990 and 2017.[4] [5] I follow an approach commonly used in the literature and identify clean energy technologies using the Cooperative Patent Classification code Y02E.[6] As discussed in detail in Section 1.1, patents on nuclear and hydrogen technologies are not included in the sample. For completeness, results including these technologies are presented in Appendix A.8. Dirty patents are identified as those belonging to class H02 and not also assigned to class Y02E.

The resulting patent-level dataset was merged with the OECD Patent Quality Indicators database (February 2022) (Squicciarini et al., 2013), which contains information on a variety of patent quality indicators calculated at the application level. Section 1.3.1 explains in detail each patent indicator used in the analysis.

To avoid double counting of the same invention, only one application from each patent family was selected.[7]. In other words, the following analysis is based on patent families rather than individual patent applications. A well-known issue with this approach when using patent-level quality indicators is that patent applications belonging to the same family often have different values for the same indicator. To address this problem, I follow Barbieri et al. (2020) and Verhoeven et al. (2016) and select only the application with the highest value of the indicator of interest from each patent family. In Appendix A.4, I test the robustness of the results using the application with the lowest value.

[4]Selecting patents from a single patent office is particularly important because part of the analysis uses patent-level indicators as dependent variables. When working with patent-level quality indicators, considering data from several offices at the same time could lead to biased indicators, which capture differences in office practices and regulations rather than differences in the characteristics of the underline patents (Squicciarini, 2013). On the other hand, this leads to the so called home bias effect. This is a common and well know problem in studies based on patents from a unique patent office, but luckily can be fully dealt with in the estimation process (see for instance Conti et al., 2018). The specification proposed in model 1.3 does this by including country fixed effects and the patent family stock of the country in dirty and clean patents. An alternative strategy would be to use the share of clean patents, as this measure clearly abstract from the total number of patents filed at the EPO by each country in the sample. I do this in Appendix A.10.

[5]Due to the small number of Y02E patent families filed during the study period, Greece, Portugal and Turkey are excluded from the sample.

[6]The list of technologies included in class Y02E can be found in Appendix A.1.

[7]A patent family is defined by the EPO as "a collection of patent applications covering the same or similar technical content", see https://www.epo.org/searching-for-patents/helpful-resources/first-time-here/patent-families.html

Each patent is assigned to a country based on the address of the inventor(s). This strategy is often used in the literature because the use of the country of the inventor(s) links the document to the environment and territory in which it was developed (Baudry and Dumont, 2006). More than 60% of the families in the sample come from Germany and Japan combined, two well-known leaders in the development of clean energy technologies. This high share is also explained by the fact that patents from the United States and South Korea could not be included in the sample due to missing values in the OECD Product Market Regulation Database (Vitale et al., OECD, 2020). This database contains the variable I use as a measure of electricity regulation in different countries, which is discussed in more detail at the end of this section. The countries included in the final sample can be see in Appendix A.2, along with the evolution of electricity regulation in each of these countries.

Figure 1.1 shows the number of patents in dirty and clean energy technologies in the sample from 1990 to 2017. To get a clearer picture of the development of the different technologies under class Y02E, the number of clean energy patents is divided into seven different categories, each representing a different technological field.

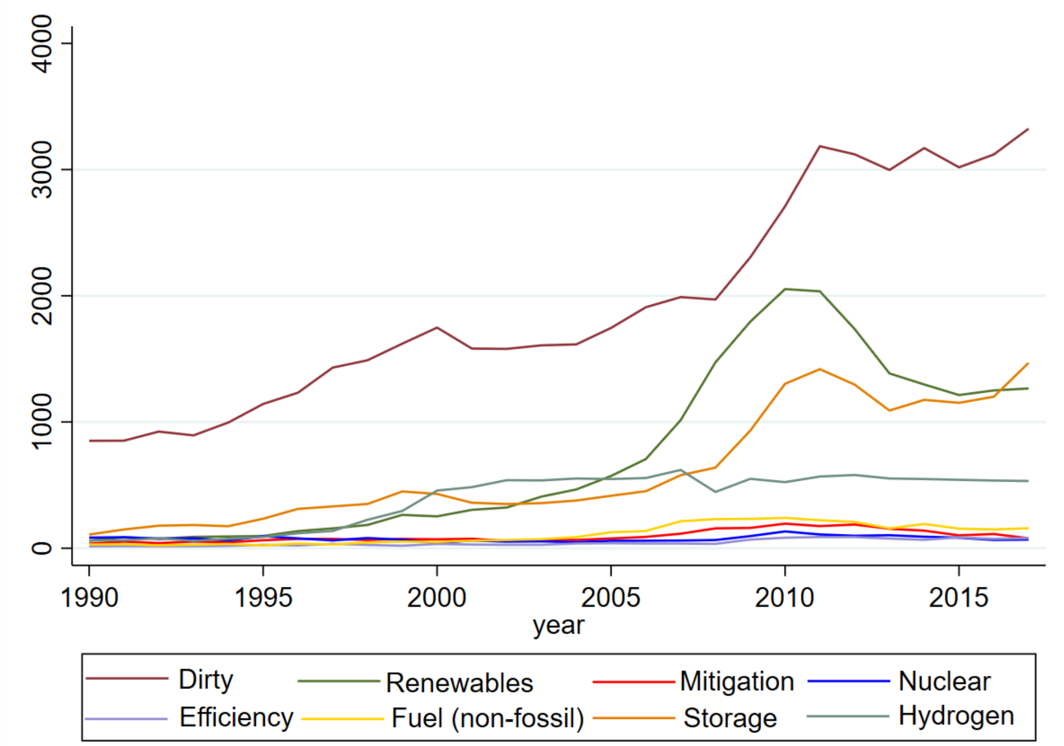
The name of the applicant(s) for each patent application was obtained using the OECD HAN (February 2022).[8] The total number of applicants in the final data set is 16,352. Of these, 1,660 filed at least one patent family in both clean and dirty technologies, 9,900 filed at least one patent in clean energy technologies, and 8,112 filed at least one patent in dirty technologies. The number of patent families filed by these applicants is very heterogeneous. More than 50% of the applicants in the sample have filed only one patent family, while the applicants in the 99th percentile have filed more than 87 patent families and the top 5 applicants in this respect all filed more than 2,000 patent families. Appendix A.6 shows no evidence of an heterogeneous effect of of electricity liberalisation when comparing applicants in the 99th percentile with the others.

Using the information on the applicant, I calculate the cumulative stock of patent families at the applicant level. To do this, I apply the perpetual inventory method with a depreciation rate of 15%, as is common in the literature (Kafouros et al., 2021, Hussinger and Pacher, 2019; Hall, 2005).[9] The same method is also used to calculate the cumulative stock of patent

[8]For applications with more than one applicant, I focus on the applicant who filed the highest number of patent families in the period of interest. Note, however, that more than 90% of the applications in the sample are associated with only one applicant and more than 99% are associated with at most two applicants

[9]According to the perpetual inventory method, the stock of patent families for firm i at time t (K_{it}) can be calculated as $K_{it}P_{it}(1\text{-}\delta)K_{i,t-1}$, where P_{it} is the number of patent families developed by firm i in year t and δ is the depreciation rate. Note that when computing the patent family stock in clean energy technologies I include

Figure 1.1: Number of Clean and Dirty patent families



Notes: Clean and Dirty patent families for countries in the regression sample. Dirty patent families are defined as those assigned to the H02 class but not to class Y02E. The codes used to identify the technological field of clean energy patents are the following: Y02E/10 for Renewable energy technologies (renewables); Y02E/20 for Combustion technologies with mitigation potential (mitigation); Y02E/30 for Energy generation of nuclear origin (nuclear); Y02E/40 for Technologies for an efficient electrical power generation, transmission or distribution (efficiency); Y02E/50 Technologies for the production of fuel of non-fossil origin (fuel, non-fossil); Y02E 60/10 trough Y02E 60/16 for energy storage technologies (storage) and Y02E 60/30 trough Y02E 60/60 for hydrogen related technologies (hydrogen). In case a patent is assigned to more than one technological fields among the ones presented in the graph, the count is fictionalized.

families at the country level.

Finally, the resulting data set is enriched with a variety of country-level control variables. These include the PMR index for the electricity sector (hereafter PMR_{elec}), i.e. the main independent variable of interest in this study. The OECD PMR database (2018) contains a set of time-varying sector-specific indicators calculated to measure the degree of liberalisation in different sectors of the economy (Vitale et al., 2020). The PMR index is a ”de jure” index. This allows for greater comparability among countries and helps the OECD to verify the reliability and accuracy of the index (Vitale et al., 2020). Potential drawbacks of the ”de jure” nature of the index are discussed in Section 1.4.3, along with a discussion of how these

all patents allocated to the Y02E class in order to fully capture the stock of knowledge of the applicant in this field.

potential limitations are addressed in the proposed analysis. The use of the PMR_{elec} index is widespread in the literature on electricity liberalisation, see e.g. Marino et al. (2019), Nicolli and Vona (2019), Nicolli and Vona (2016), Cambini et al. (2016), Nesta et al. (2014). In its latest version, the index covers the period 1975-2018 and ranges from 6 to 0, with higher values signalling a more regulated electricity market. Appendix A.2 shows the evolution of the PMR_{elec} index by country over the period considered in the study.

Descriptive statistics for the variables discussed in this section, as well as for other variables used in the analysis, can be found in Table 1.1.

Table 1.1: Descriptive Statistics

Variable	Variable Description	Observations	Mean	Std. Deviation	Min	Max	Source
Country-level alanlysis							
$\text{Count}_{c,t}^{Clean}$	Count of clean patent families	464	96.74	200.34	0	1407.47	Author's calculations, data from PATSTAT
$\text{Count}_{c,t}^{Dirty}$	Count of dirty patent families	464	114.21	203.27	0	1117.5	Author's calculations, data from PATSTAT
Lag PMR_{elec}	Regulation in the Electricity Sector	464	2.45	1.60	0.14	6.28	OECD PMR database (2018)
Lag EPS_{tech}	Stringency of policies supporting green technologies	464	1.95	1.18	0.5	6	OECD EPS database (2022)
Lag GDP_{pc}	GDP per capita PPP (Thousands of Current $)	464	37.19	16.30	11.53	103.55	World Bank
Lag Oil price (imports)	Crude oil import prices (100$/barrel)	464	0.48	0.34	0.12	1.18	OECD Data
Lag Knowledge Stock (country), Y02E	Cumulative discounted nbr of country's Y02E families, thousands of patents	464	0.39	0.86	0	5.32	Author's calculations, data from PATSTAT
Lag Knowledge Stock (country), Dirty	Cumulative discounted nbr of country's Dirty families, thousands of patents	464	0.36	0.67	0	3.61	Author's calculations, data from PATSTAT
Patent-level Analysis, Clean patents							
Radicalness	Radicalness Index (Squicciarini 2013, Shane 2001)	42,834	0.33	0.26	0	1	OECD Patent Quality database (Feb. 2022)
Originality	Originality Index (Squicciarini 2013, Trajtenberg et al.,1997)	42,797	0.71	0.19	0	0.98	OECD Patent Quality database (Feb. 2022)
Many applicants	Equal to one if more than one applicant, zero otherwise	42,834	0.07	0.25	0	1	OECD Han database (Feb. 2022)
Lag of Knowledge Stock_{clean}	Cumulative discounted nbr of applicant's Y02E families, thousands of patents	42,834	0.07	0.13	0	0.78	Author's calculations, data from PATSTAT
Lag PMR_{elec}	Regulation in the Electricity Sector	42,834	1.37	0.99	0.14	6.28	OECD PMR database (2018)
Lag EPS_{tech}	Stringency of policies supporting green technologies	42,834	3.07	1.26	0.5	6	OECD EPS database (2022)
Lag GDP_{pc}	GDP per capita PPP (Thousands of Current $)	42,834	41.39	10.16	11.53	103.55	World Bank
Lag Oil price (imports)	Crude oil import prices (100$/barrel)	42,834	0.68	0.33	0.12	1.18	OECD Data
Patent-level Analysis, Dirty patents							
Radicalness	Radicalness Index (Squicciarini 2013, Shane 2001)	51,143	0.32	0.26	0	1	OECD Patent Quality database (Feb. 2022)
Originality	Originality Index (Squicciarini 2013, Trajtenberg et al.,1997)	50,265	0.70	0.21	0	0.98	OECD Patent Quality database (Feb. 2022)
Many applicants	Equal to one if more than one applicant, zero otherwise	51,143	0.04	0.20	0	1	OECD Han database (Feb. 2022)
Lag of Knowledge Stock_{dirty}	Cumulative discounted nbr of applicant's dirty families, thousands of patents	51,143	0.09	0.18	0	1.06	Author's calculations, data from PATSTAT
Lag PMR_{elec}	Regulation in the Electricity Sector	51,143	1.73	1.38	Min	0.14	6.28 PMR database (2018)
Lag EPS_{tech}	Stringency of policies supporting green technologies	51,143	2.81	1.32	0.5	6	OECD EPS database (2022)
Lag GDP_{pc}	GDP per capita PPP (Thousands of Current $)	51,143	38.50	11.03	11.53	103.55	World Bank
Lag Oil price (imports)	Crude oil import prices (100$/barrel)	51,143	0.58	0.35	0.12	1.18	OECD Data

Notes: Descriptive statistics for the regression datasets used in Section 1.5 (Tables 1.2, 1.4 and 1.5).

1.3.1 Measuring the exploratory nature of energy patents

To test the hypothesis that electricity liberalisation leads to more exploratory energy innovations, I rely on two well-established patent quality indicators from the OECD Patent Quality Indicators database (February 2022) (Squicciarini et al., 2013). These are the radicalness index (a là Shane 2001) and the originality index (a là Trajtenberg et al., 1997). The decision to focus on these two indicators as proxies for the exploratory nature of a patent follows a broad literature that argues that exploratory innovations are often based on a diverse search space and rely on the recombination of different ideas (De Noni et al., 2021; Barbieri et al., 2020; Squicciarini et al., 2013; Quintana-García et al., 2008; Shane, 2001; Trajtenberg et al., 1997). The two proposed indicators both assess the breath and diversity of a patent search space, but in different ways. The radicalness index measures the extent to which a patent draws on knowledge from outside technological fields, i.e. how much a patent relies on knowledge spillovers from other technologies. The originality index can be interpreted as a measure of the interdisciplinarity of a patent, with interdisciplinary patents being defined as those that cite many technological fields and also belong to many technological fields (Higham et al., 2021). If electricity liberalisation reduces path dependence in the electricity sector, innovation is expected to become less linked with the traditional technologies used by incumbents. New technologies and approaches are likely to emerge in more competitive markets, and the literature suggests that these will be developed, at least to some extent, thanks to the recombination of more diverse knowledge (Squicciarini et al., 2013; Quintana-García et al., 2008; Shane, 2001; Trajtenberg et al., 1997). The remainder of this section describes the radicalness and originality indices in detail.

Radicalness Index

The radicalness index measures how different a patent is from the patents it cites (Shane, 2001). The intuition behind this indicator is that ”when a patent cites previous patents in classes other than the ones it is in, that pattern suggests that the invention builds upon different technical paradigms from the one in which it is applied” (Shane, 2001, p. 210. See also Barbieri et al., 2020; Verhoeven et al. 2016; Squicciarini et al. 2013; Rosenkopf and Nerkar 2001). This indicator is therefore linked to the notion of knowledge spillovers, as it measures how much knowledge the patent takes from outside technological fields. Following Shane (2001), the index for a focal patent, *p*, is defined by Squicciarini et al. (2013) as follows:

$$Radicalness_p = \Sigma_j^{n_p} CT_j / n_p; \quad IPC_{pj} \neq IPC_p \tag{1.1}$$

Where CT_j is the number of 4-digit IPC codes (IPC$_{pj}$) of patent *j* cited by patent *p* that are not assigned to the focal patent *p*. The denominator, n_p, is the number of total IPC classes in the backward citations of patents cited by patent *p*, counted at the most disaggregated level available.[10] The indicator is therefore normalised and its value ranges between zero and one. High values of the radicalness index signify that the patent takes knowledge from outside

[10]Backward citations are defined as the citations that the focal patent makes to older patents to disclose the prior knowledge on which it is based (Squicciarini et al., 2013)

technological fields and applies it to its own field(s). Thus, the higher the index, the more the patent relies on technological spillovers from outside technological fields.

If electricity liberalisation reduces path dependency in the electricity sector, we could expect energy technologies to explore more outside technological fields and borrow more knowledge from them. This kind of behaviour would be reflected in an increase in the radicalness index of these patents.

Originality Index

The originality index measures how much the backward citations of a patent are distributed across different technological fields (Trajtenberg et al., 1997). The intuition behind this indicator is that knowledge recombination processes that rely on a diversified set of knowledge sources are expected to lead to more original results (Barbieri et al., 2020, Dechezleprêtre et al., 2017; Verhoeven et al,. 2016, Squicciarini et al. 2013, Trajtenberg et al.,1997). More generally, the indicator is closely linked to the notion of interdisciplinarity. (Higham et al., 2021). Higham et al. (2021) define interdisciplinary patents as those that cite and are simultaneously associated with many technological fields.

Building on Hall at al. (2001), Squicciarini et al. (2013) calculate the originality index as follows:

$$Originality_p = 1 - \Sigma_j^{n_p} s_{pj}^2 \tag{1.2}$$

Where s_{pj} is the percentage of citations of patent *p* to the 4-digit IPC patent class *j* with respect to the total number of 4-digit IPC patent classes cited by patent p (n$_p$). Note that the indicator is based on a Hirschman-Herfindahl index, which measures the extent to which the backward citations of the focal patent are concentrated in different technological fields (i.e. $\Sigma_j^{n_p} s_{pj}^2$). The originality index ranges from zero to one, and higher values of the index signal patents with backward citations distributed across many different fields.

If electricity liberalisation reduces path dependence in the electricity sector, we could expect to see energy technologies that become more interdisciplinary. The originality index allows us to measure whether this is the case by assessing whether energy patents cite more technological fields while also increasing the number of technology classes to which they belong.

1.4 Methodology and Identification Strategy

The analysis is divided into two parts. The first part, discussed in Section 1.4.1, examines the impact of electricity liberalisation on the direction of technological change in the electricity sector. For this purpose, I use a country-level panel data set and focus on the number of clean and dirty patent families filed in country *c* and year *t*. The second part is discussed in Section 1.4.2 and uses the radicalness and originality indices as dependent variables to test whether electricity liberalisation leads to more exploratory innovations. In this case, the unit of observation is the patent family. In both cases, the analysis covers the period between 1990 and 2017.

1.4.1 The direction of technological change in the electricity sector

Equation (3) is used to estimate the effect of electricity liberalisation on the direction of technological change in the electricity sector.

$$\begin{aligned} E[Count^k_{ct}|\mathbf{X}_{c,t}] = exp(&\beta_1 PMR_{elec,c,t-1} + \beta_2 EPS_{tech,c,t-1} + \beta_3 GDP_{pc,c,t-1} + \\ &\beta_4 OilPriceImp_{c,t-1} + \beta_5 KStock^{clean}_{c,t-1} + \beta_6 KStock^{dirty}_{c,t-1} + \tau_t) \end{aligned} \tag{1.3}$$

Where $Count^k_{ct}$ is number of patents related to electricity generation, transmission or distribution in country *c* and year *t* for technology type *k*. Given the nature of the dependent variable, equation (3) is estimated using Poisson regression. Following the standard approach in similar empirical setting, the count of patents is fictionalized in the case of inventors residing in different countries. The main coefficient of interest is β_1, which quantifies the effect of a change in the degree of electricity liberalisation on the dependent variables of interest. The variables on the right-hand side of the equation are lagged by one year to account for the lag in the effect of policy variables.

As control variables, I include the sub-index of the OECD Environmental Policy Stringency Database (2022) which measures the use of policies to support clean-energy innovation (Kruse et al., 2022), GDP per capita[11], the price of crude oil imports [12] (Calel and Dechezleprêtre, 2016) and the country's knowledge stocks in clean and dirty energy technologies,

[11] Source: World Bank ttps://data.worldbank.org/indicator/NY.GDP.PCAP.CD. Retrived 27 Sept 2023

[12] Source: OECD (2022), Crude oil import prices (indicator). doi: 10.1787/9ee0e3ab-en (Retrieved 22 June 2022)

computed using the perpetual inventory method as described in Section 1.3 (Aghion et al 2016; Nesta et al., 2014; Nicolli and Vona, 2016; Popp 2002).

Standard errors are clustered at the country level (Abadie et al., 2023).

1.4.2 Exploratory Innovation in the Electricity Sector

I now turn to the question of whether electricity liberalisation makes innovation in the electricity market more exploratory. To estimate the relationship between electricity liberalisation and exploratory innovation, I rely on model (4), which I estimate separately for clean and dirty patents. As already discussed, I use the radicalness and originality indices from the OECD Patent Quality Indicators Database (February 2022) as indicators for exploratory patents. Since these indicators are computed at the application level, in this case the unit of observation is the patent. Note that in this specification dealing with patent families that have inventors in more than one country is less straightforward with respect to the framework discussed in the previous section. Since joint applications are relatively rare (about 8% of the applications in the sample), the following analysis focuses only on patents that can be unambiguously assign to a single country based on the address of the inventors. In Appendix A.5 I test the robustness of the results to the inclusion of patent families with inventors in different countries by assigning them to the country with the most liberalised electricity market among inventor's countries.

$$PatInd_i = \beta_1 LagPMR_{elec,i} + \beta_2 Lag\mathbf{X}_i + \beta_3 \mathbf{A}_i + App_i + Tech_i + Country_i + Year_i + \varepsilon_i \quad (1.4)$$

$PatInd_i$ represents one of the patent indicators discussed in Section 1.3, calculated for patent *i*. The variable PMR_{elec} and the control variables in the matrix **X** are lagged by one year to account for the lag in the effect of policy variables.

The coefficient β_1 quantifies the effect of a change in the degree of electricity liberalisation on $PatInd_i$.

The matrix **X** includes the sub-index of the OECD Environmental Policy Stringency Database (2022) that measures support for clean-energy innovation (Kruse et al., 2022), GDP per capita[13] and the price of crude oil imports [14].

[13] Source: World Bank ttps://data.worldbank.org/indicator/NY.GDP.PCAP.CD. Retrived 27 Sept 2023

[14] Source: OECD (2022), Crude oil import prices (indicator). doi: 10.1787/9ee0e3ab-en (Retrieved 22 June 2022)

A is a matrix of applicant-level control variables. To proxy the resources and knowledge available at the applicant level, I include in the regression the stock of clean or dirty patent families calculated for each applicant. In particular, I use the stock of clean-energy patent family when considering clean energy patents and the stock of dirty patent families when estimating the model for dirty patents. These are calculated using the perpetual inventory method with a discount rate of 15% as described in Section 1.3. The relationship between $PatInd_i$ and these knowledge stocks is unlikely to be linear. On the one hand, new entrants are expected to develop more exploratory inventions, which would suggest a negative effect of the patent family stock on $PatInd_i$. On the other hand, this negative effect could be weaker (or become positive) for applicants who file many patents, as these applicants have access to more resources and can build on a larger body of knowledge. To capture these dynamics, I also include the squared value of the applicant's patent family stock as control variable. Finally, I control for patents with multiple applicants using a dummy variable that is equal to one for patent families with more than one applicant and zero otherwise. I do this because more than one applicant working on the same invention could lead to more resources for its development. [15]

The specification is then augmented with fixed effects for the applicant of the patent (App_i), its technology class ($Tech_i$)[16], and the country ($Country_i$) and year ($Year_i$) in which it was filed. The inclusion of applicant fixed effects in the model allows me to control for time-invariant heterogeneity across applicants. However, it also prevents patent families of applicants with only one family from being included in the regression (Correia, 2015). For this reason, I present the results both with and without the inclusion of applicant fixed effects in the model.

In the literature, the use of the radicalness and originality indices as dependent variables in patent-level specifications is sometimes accompanied by application-level control variables selected on the basis of how these indices are constructed. In particular, since these indicators use information on the citations made by the patent, previous work controls for the number of backward citations of the patent (see for example Barieri et al., 2020; Hall et al., 2001). In addition, the scope of the patent, i.e. the number of full-digit IPC codes to which the invention is assigned, is often used as control variable for the radicalness index (Barieri et al.,

[15]Note that more than 90% of the applications in our sample have only one applicant, see Table 1.1.

[16]In the OECD Patent Quality Database, the information on the technological fields of the patent is based on the WIPO taxonomy (Schmoch, 2008). For patents assigned to more than one technological field, only the one with the most IPC codes is retained. If a patent has the same number of IPC codes for different technological fields, it is randomly assigned to one of these fields (Schmoch, 2008

2020; Sapsalis et al., 2006). The reason for this is that a higher scope is usually associated with lower values of the radicalness index, as it is more difficult for a broader patent to cite technological fields to which it is not assigned. In the proposed framework, I cannot control for these variables without running into the problem of "bad controls" (Cinelli et al., 2022). Both the number of backward citations and the scope of the patent can be though as alternative outcome variables and/or channels through which the effect we want to investigate could be expected to unfold. This being the case, I do not include these variables in the specification presented in model (4). In Appendix A.7, I conduct a sensitivity analysis and show that the results for the estimation of model (4) are not affected in magnitude or statistical significance by the inclusion of these potentially problematic variables.

Since the "treatment variable" (i.e. PMR_{elec}) is at the country level, I cluster the standard errors at this level (Abadie et al., 2023). This results in clusters that are heterogeneous in size, as there are large differences in the number of patents filed across countries. In similar contexts, wild cluster bootstrapping has been shown to perform better than inference relying on clustered standard errors based on large-sample theory (Roodman et al., 2019; Cameron and Miller, 2015). For this reason, when presenting the results of the estimation of model (4), I report p-values and confidence intervals obtained implementing a wild cluster bootstrap using the STATA command 'boottest' (Roodman et al., 2019).

To estimate model (4), I rely on linear regression analysis using the reghdfe command in STATA (Correia, 2016). Although the radicalness and originality indices are bounded between zero and one, I opt for a linear model instead of a fractional model for several reasons. First, a linear model allows to easily include high-dimensional fixed effects, thus allowing the inclusion of applicant fixed effects in the regression. Second, this allows me to rely on wild bootstrap-based inference, as opposed to the alternative score bootstrap used with the maximum likelihood estimator. This is an advantage as the former has been shown to be more reliable compared to the latter (Roodman et al., 2019). Similar choices are not uncommon in the literature, see for example Porter and Serra (2019). As a robustness check, Appendix A.3 presents the results of a fractional probit model with inference based on score bootstrap and without including applicant fixed effects. The results obtained with this alternative specification are very similar in magnitude and statistical significance to the ones obtained relying on linear regression and wild bootstrap.

1.4.3 Instrumental Variable strategy

The proposed estimation strategy might suffer from endogeneity stemming from different sources.

First, while the PMR_{elec} index is widely used in the literature (e.g. Marino et al., 2019; Nicolli and Vona, 2019; Nicolli and Vona, 2016; Nesta et al., 2014), it is at best an imperfect proxy of incumbents' effective market power. In particular, one might worry that the PMR index, being a "de jure" index, underestimates the ability of incumbents to retain market power after the reform in a highly concentrated industry such as the energy sector (Vitale et al., OECD, 2020, Nicolli and Vona, 2019). To explain why this is the case, let us take a concrete example and look at Japan's electricity reform. The PMR_{elec} index for Japan in 2009 is 0.93. This means that, according to the PMR index, Japan had made significant progress in liberalising the electricity market. In practise, however, the market share of independent power producers in that year was only 2.8% (Jones and Kim 2013). This suggests that, although the regulation of the electricity sector was pro-competition on paper, incumbents were able to retain much of their market power. To make matters worse, this kind of measurement error is likely correlated with the PMR index itself. For while one can imagine incumbents retaining a high degree of market power in a liberalised electricity sector, a scenario in which incumbents' market power decreases without a change in electricity regulation seems unlikely at best.

Second, the development of more scalable technologies for electricity generation in the 1990s was a key factor that enabled the liberalisation of the electricity sector in the first place (e.g. gas-fired power plants and RETs) (Batlle and Ocaña, 2013). Given this, one might worry about a possible reverse causality issue.

Third, countries that develop more radical clean-energy innovations might also be the ones that reform the electricity market first or more. This could be the case, for example, if "green lobbies" have the ability to influence both the development of radical clean-energy technologies and the regulation of the electricity sector (Nicolli and Vona 2019).

The empirical framework described in the previous section already mitigates some of these concerns, e.g. by controlling for a country support to clean-energy technologies and by using applicant and country fixed effects. However, to account for these potential sources of endogeneity more fully, I follow Nicolli and Vona (2019) and use an instrumental variable strategy where regulation in the telecommunication sector is used as instrument for regulation in the electricity sector. Regulation in the telecommunication sector is measured by the OECD PMR index for this particular industry.

The liberalisation of telecommunications took place before the liberalisation of the electricity sector and played a crucial role in getting the latter off the ground (Nicolli and Vona, 2019; Pollitt, 2012; Joskow, 2008). At the same time, this reform can be considered independent from the lobbying power of actors in the electricity sector and from technological developments in energy supply technologies. This helps to address the second and third sources of potential endogeneity described at the beginning of this section. In addition, electricity liberalisation is more likely to reduce the market power of incumbents in countries where public authorities and policy makers invest greater effort in the liberalisation process and are more committed to this reform. A widespread change in the regulation of different sectors can be interpreted as a signal of effort and commitment in this sense and is therefore likely to correlate with the implementation of a successful liberalisation process in the electricity sector (Nicolli and Vona 2019). This being the case, the proposed strategy also mitigates the problem of measurement error discussed above.

The country-level regression for the estimation of model (3) relies on a fixed-effect Poisson model. To account for the possible endogeneity of PMR_{elec} in this non-linear setting, I apply the Control Function (CF) Approach (Wooldridge, 2015). The CF approach relies on the availability of one or more instrumental variables to run a first-stage regression and uses the generalised residuals from this first-stage regression as *control function*. This *control function* is then added to the second-stage regression and this allow us to estimate the effect of the endogenous regressor appropriately (Wooldridge, 2015). A particularly interesting feature of the CF approach is that the estimated coefficient on the *control function* in the second-stage regression is a regression-based Hausman (1978) test that can easily be made robust to heteroskedasticity and cluster correlation (Wooldridge, 2015). A statistically significant coefficient on the *control function* is evidence that the regressor we suspect to be endogenous is indeed endogenous. A non-significant coefficient indicates that this regressor is exogenous and there is no reason to prefer the IV estimates over the simple Poisson estimates (Wooldridge, 2015). Note that if the control function is statistically significant in the second stage, it is important to adjust the standard errors to account for the two-stage estimation procedure (Wooldridge, 2015). Following Wooldridge (2015), I do this using bootstrap.

The patent-level analysis is carried out relying on linear regression analysis, making the IV estimation relatively straightforward. In recent years, the literature on inference for two stage least squares estimates has been very active. Many papers that have recently addressed this issue recommend the use of Anderson-Rubin confidence intervals in single instrument applications (Kean and Neal, 2023; Lee et al., 2022; Lal et al., 2021; Andrews et al., 2019; Anderson and Rubin, 1949). The main advantage of these confidence intervals is that they are efficient in the presence of weak instruments, but because of their desirable properties, their use is now often recommended regardless of the value of the first-stage F-statistic (Keane and Neal, 2022; Andrews et al., 2019). In light of these recent developments, and as recommended by Keane and Neal (2022), the inference in the IV estimates of model (3) is based on the Anderson-Rubin test rather than the t-test. To also address the problem of having clusters that are heterogeneous in size, discussed in Section 1.4.2, I rely on a bootstrap version of the Anderson-Rubin test based on wild bootstrap. This is done in STATA using the same 'boottest' command discussed in the previous section (Roodman et al., 2019). Since the first-stage estimates are also likely to suffer from this problem, I follow Young (2022) and report the p-value of the first-stage regression based on wild bootstrap inference in addition to the first-stage F-statistic, and ensure that this p-value is less than 0.01.

1.5 Results

In this section, I present the results of the empirical analysis. Section 1.5.1 shows the effect of electricity liberalisation on the number of patents filed in dirty and clean technologies. Section 1.5.2 looks at the impact of electricity liberalisation on the exploratory or exploitative nature of innovations in clean and dirty technologies. The significance of the results is discussed in more detail in Section 1.6. The result tables in Section 1.5 focus on the main coefficient of interest, i.e. the estimated coefficient for PMR_{elec}. See Appendix A.9 for the complete result tables.

1.5.1 Electricity liberalisation and the direction of technological change in the electricity sector

In this section I examine the impact of electricity liberalisation on the number of clean and dirty patents filed for technologies related to the generation, transmission and distribution of electricity. Table 1.2 shows that electricity liberalisation affects the direction of technological change in the electricity sector by reducing the number of dirty technologies and increasing the number of clean energy patents. The results presented in Panel B are obtained using the CF approach and confirm what we see in Panel A. The control function in the regression for dirty and clean technologies is significant, which is evidence that $PMR_{elec,t-1}$ is endogenous. This being the case, standard errors in columns (1) and (2) are adjusted using a bootstrap procedure (Wooldridge, 2015). The coefficients estimated using the CF approach are larger compared to what we observe in Panel A. The best explanation for this pattern is endogeneity in the results from Panel A coming from the measurement error problem discussed in Section 1.4.3. If incumbents can retain high market power even after changes in the laws regulating competition in the electricity sector, we would expect the coefficients in Panel A to be downward biased. The proposed instrumental variable strategy seems better suited to capture the actual degree of competition in a country's electricity market, leveraging the fact that widespread liberalisation in various sectors signals greater effort and commitment on the part of policy makers and market authorities (Nicolli and Vona, 2019).

According to the results in Panel B, a one unit decrease in PMR$_{elec,t-1}$ leads on average to about 47 fewer dirty patents and about 33 more clean patents. To put these values in perspective, a one unit change in PMR$_{elec,t-1}$ amounts to 62% of the standard deviation of this variable, while the changes triggered in the number of dirty and clean energy patents amount to 23% and 16% of the standard deviations of these counts, respectively. [17]

Table 1.3 breaks down the analysis for different technological fields belonging to class Y02E. Note that the same patent can be assigned to more than one of the codes in class Y02E. If this happens, I deal with it by using fractional counts, e.g. a patent assigned to two technological fields is attributed to both but counted as half a patent. [18]

According to the results in Panel A of Table 1.3, the effect of electricity liberalisation is positive and significant for 3 out of 5 classes, namely renewable energies, combustion technologies with mitigation potential and storage technologies. In the estimates obtained with the CF approach (Panel B of Table 1.3), the coefficient remains significant only for electricity storage and renewables, albeit only at the 90% level for the latter. However, for both renewables and combustion technologies with mitigation potential, we find that the control function is not significant, with p-values of 0.299 and 0.289, respectively. In these cases, we should reject the null hypothesis of the Hausman test built into the CF regression and conclude that

[17] The incidence ratio associated with $PMR_{elec,t-1}$ for the estimates in column (1) is 1.41. The average number of dirty patents per country in the sample used for the regression is 114.21. Thus a one unit decrease in $PMR_{elec,t-1}$ leads on average to 46.79 fewer dirty patents (1.41*114.21 - 114.21 = 46.79). For the variable $PMR_{elec,t-1}$, a one unit change is equal to 62% of its standard deviation and for the number of dirty patents, a 46.79 change is equal to 23% of the standard deviation of this variable (46.79/203.27). For the estimates in column (2), the incidence ratio associated with $PMR_{elec,t-1}$ is 0.66 and the average number of clean patents in the sample is 96.74. Thus, on average, a one unit decrease in electricity liberalisation leads to almost 33 additional clean patents (0.66*96.74 - 96.74 - 32.89), a change equivalent to 16.4% of the standard deviation of the outcome variable in column (2) (32.89/200.34).

[18] However, these patents represents less than 8% of the patents in the sample and using alternative methods to deal with this issue does not affect the results. For example, the results are robust if each patent is assigned fully to all Y02E technology classes listed in the document or if patents assigned to multiple codes are excluded from the sample

$PMR_{elec,t-1}$ is not actually endogenous and there is no reason to reject the Poisson estimates in Panel A (Wooldridge, 2015). Again, the estimated coefficients using the CF approach are larger than for our base Poisson estimates, suggesting endogeneity stemming from measurement error in the PMR index. This is especially true for storage technologies, which is to be expected as this is the only case where the control function is significant. This pattern suggests that the ability of incumbents to retain market power after electricity liberalisation is particularly detrimental to the development of energy storage technologies. This result can be explained by noticing that energy storage technologies play a key role in the development of a decentralised generation paradigm, while incumbents' skills and expertise are generally tied to large-scale plants in a highly centralised generation paradigm (Defeuilley, 2019; Nesta et al., 2014). Overall, these findings support the idea that electricity liberalisation can reduce the electricity sector's dependence on dirty technologies and steer innovation towards clean energy technologies.

Table 1.2: Poisson fixed effect estimates of Model (3)

	(1)	(2)
Panel A	$Count_{ct}^{Dirty}$	$Count_{ct}^{Clean}$
PMR$_{elec,c,t-1,}$	0.1469***	-0.0999***
	(0.0439)	(0.0336)
Controls	Yes	Yes
Country FE	Yes	Yes
Year FE	Yes	Yes
Observations	464	464
Panel B	$Count_{ct}^{Dirty}$	$Count_{ct}^{Clean}$
PMR$_{elec,c,t-1,}$	0.3469***	-0.4144***
	[0.0872]	*[0.0898]*
Control Function	-0.2435***	0.3830***
	[0.0928]	*[0.1014]*
Controls	Yes	Yes
Country FE	Yes	Yes
Year FE	Yes	Yes
F-stat first stage	14.70	14.70
First Stage Bootstrap p-value	0.0024	0.0024
Observations	464	464

Notes: All specifications also control for the use of policies to promote clean energy innovation, GDP per capita, the price of crude oil imports, and the country's knowledge stocks in dirty and clean energy technologies. In Panel B regulation in the telecommunications sector is used as instrument for regulation in the electricity sector using a CF approach. Standard errors are clustered at the country level and reported in parentheses (18 clusters). In Panel B, if the control function is significant in the first-stage regression, I adjust the standard errors based on 999 bootstrap replications and report them in italics and in square brackets.

Table 1.3: Poisson fixed effect estimates of Model (3) - Heterogeneity by technology class

	(1)	(2)	(3)	(4)	(5)
Panel A	$Count_{ct}^{Renewables}$	$Count_{ct}^{CombustionMitigation}$	$Count_{ct}^{Efficiency}$	$Count_{ct}^{Fuel,nonfossil}$	$Count_{ct}^{Storage}$
PMR$_{elec,c,t-1,}$	-0.1439***	-0.1074**	0.0117	-0.002	-0.1032**
	(0.0415)	(0.0427)	(0.0960)	(0.0622)	(0.0469)
Controls	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	464	464	464	464	464
Panel B	$Count_{ct}^{Renewables}$	$Count_{ct}^{CombustionMitigation}$	$Count_{ct}^{Efficiency}$	$Count_{ct}^{Fuel,nonfossil}$	$Count_{ct}^{Storage}$
PMR$_{elec,c,t-1,}$	-0.3352*	-0.3218	0.1953	0.1144	-0.4217***
	(0.1905)	(0.2133)	(0.2547)	(0.2112)	*[0.1374]*
Control Function	0.2225	0.2632	-0.2266	-0.1356	0.4056**
	(0.2141)	(0.2485)	(0.3043)	(0.2228)	*[0.1574]*
Controls	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
F-stat first stage	14.70	14.70	14.70	14.70	14.70
First Stage Bootstrap p-value	0.0010	0.0010	0.0010	0.0010	0.0057
Observations	464	464	464	464	464

Notes: All specifications also control for the use of policies to promote clean energy innovation, GDP per capita, the price of crude oil imports, and the country's knowledge stocks in dirty and clean energy technologies. In Panel B regulation in the telecommunications sector is used as instrument for regulation in the electricity sector using a CF approach. Standard errors are clustered at the country level and reported in parentheses (18 clusters). In Panel B, if the control function is significant in the first-stage regression, I adjust the standard errors based on 999 bootstrap replications and report them in italics and in square brackets.

1.5.2 Electricity liberalisation and exploratory innovations in the electricity sector

Table 1.4 and 1.5 show the results of OLS estimation of model (4) for clean and dirty energy patents, respectively. The wild cluster bootstrap does not assume normality and therefore does not compute standard errors (Roodman et al., 2019). I therefore follow Porter and Serra (2019) and report the obtained p-values and 95% confidence intervals in the results tables. Appendix A.3 shows the results when using a fractional model and inference based on score bootstrap.

In Table 1.4, when the dependent variable is the radicalness index (columns 1 and 2) the coefficient associated with PMR_{elec} has the expected sign and is always statistically significant, at least at the 95% level. Instead, I find no effect of electricity liberalisation on the originality index. This pattern is confirmed by the various robustness checks presented in Appendix A. Panel B shows the results for the IV strategy discussed in Section 1.4.3. These results confirm the conclusions from Panel A and the magnitude of the estimated effect is similar. The effect of electricity liberalisation on the radicalness index appears not only statistically significant but also economically meaningful. Comparing the coefficients estimated in columns (1) and (2) of Panel B with the standard deviation of the radicalness index in the estimation sample, i.e. 0.25, we see that the estimated effect amounts to 10% and 6% of this value, respectively. The standard deviation of PMR_{elec} in the estimation samples used in Table 1.4 is roughly one (see Table 1.1), which means that the coefficient can be naturally interpreted as the effect of a one standard deviation change in PMR_{elec} on the dependent variable.

In Table 1.5 we see that the effect of electricity liberalisation on the exploratory nature of dirty patents is never significant. Interestingly, the estimated coefficients in columns (1) and (2) of Table 5 are much closer to zero than their counterparts in Table 1.4.

Table 1.6 examines whether the impact of electricity liberalisation on the radicalness index is heterogeneous among the different technological fields were are considering. [19] As can be seen in Figure 1.1, splitting the sample in this specification can be difficult due to the relatively small number of patent families in most Y02E technological fields which leads to a low number of observations. For this reason, I have divide the sample into only three groups: renewable energy technologies, energy storage technologies, and other technologies (which includes combustion technologies with mitigation potential, higher efficiency technologies, and fuels of non-fossil origin). Since one patent family can be allocated to more than one technological fields, here I assign each patent to all the fields listed in the document.

[19] Similar results for the originality index are not presented as they confirm the lack of a robust relationship with electricity liberalisation.

Two patterns emerge clearly from the results in Table 1.6. The radicalness index of renewable energy technologies is not affected by electricity liberalisation, as the coefficient associated with PMR_{elec} is never significant and sometimes even positive. On the other hand, it is again storage technologies that seem to benefit most from electricity liberalisation. The coefficient of PMR_{elec} is always of the expected sign and significant in 3 of the 4 specifications presented. The only exception is column (1) in Panel A, where I present the estimates of the OLS regression and with applicant fixed effect. Even in this case, however, the magnitude of the coefficient is relatively high, and the lack of significance can be explained by the combination of lower precision of the estimates due to fewer observations and the downward bias of the coefficient in the specifications that do not account for endogeneity. Indeed, despite the very different empirical framework. we once again find that in the 2SLS estimates the estimated effect of electricity liberalisation on energy storage patents increases in magnitude. This is consistent with the results presented in Table 1.3. Finally, the coefficient for other technologies is always negative, but significant only in Panel A. However, the very small number of observations for this group makes it difficult to draw any conclusions.

In summary, the results of this section suggests that clean energy technologies developed in more competitive electricity markets use more knowledge spillovers from outside technological fields. Following the literature, I interpret this as a sign of exploratory behaviour. This pattern is particularly strong for energy storage technologies, which is consistent with the results presented in Section 1.5.1. At the same time, citations of clean energy patents remain concentrated in the same number of technological fields, as shown by the non-significant effect on the originality index. This suggests that while electricity liberalisation increases the use of knowledge spillovers in the development of clean energy technologies, it does not make these technologies more interdisciplinary. The search space of dirty patents is neither positively nor negatively affected by electricity liberalisation.

Table 1.4: OLS and 2SLS estimates of Model (4) - Clean patents

	(1)	(2)	(3)	(4)
	Clean	Clean	Clean	Clean
Panel A: OLS	Rad	Rad	Ori	Ori
Lag PMR$_{elec}$	-0.0171	-0.0181	-0.0021	-0.0018
	(0.0167)	(0.0005)	(0.2938)	(0.2118)
	[-0.0316, -0.0043]	[-0.0285, -0.0124]	[-0.0110, 0.0020]	[-0.0058, 0.0011]
Controls	Yes	Yes	Yes	Yes
Applicant FE	Yes	No	Yes	No
Technology FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	37,371	42,834	37,333	42,797
Panel B: 2SLS	Rad	Rad	Ori	Ori
Lag PMR$_{elec}$	-0.0252	-0.0161	-0.0016	0.0030
	(0.0000)	(0.0074)	(0.7574)	(0.3468)
	[-0.0431, -0.0173]	[-0.0235, -0.0062]	[-0.0275, 0.0073]	[-0.0025, 0.0107]
Controls	Yes	Yes	Yes	Yes
Applicant FE	Yes	No	Yes	No
Technology FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
F-stat first stage	41.77	53.72	41.73	53.75
First Stage Bootstrap p-value	0.0006	0.0014	0.0006	0.0014
Observations	37,371	42,834	37,333	42,797

Notes: All specifications also control for the use of policies to promote clean energy innovation, GDP per capita, the price of crude oil imports, the number of applicant, the knowledge stock computed at the applicant level and its squared value. In Panel A, I report wild bootstrap p-values in parentheses and wild bootstrap 95% confidence intervals in square brackets, generated using boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters). In Panel B, regulation in telecommunication is used as instrument for regulation in the electricity sector and inference is based on a wild bootstrapped version of the Anderson-Rubin obtained using the boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters).

Table 1.5: OLS and 2SLS estimates of Model (4) - Dirty patents

	(1)	(2)	(3)	(4)
	Dirty	Dirty	Dirty	Dirty
Panel A: OLS	Rad	Rad	Ori	Ori
Lag PMR$_{elec}$	-0.0022	-0.0036	0.0010	-0.0041
	(0.6384)	(0.5276)	(0.8675)	(0.3356)
	[-0.0145, 0.0092]	[-0.0180, 0.0104]	[-0.0113, 0.0106]	[-0.0163, 0.0050]
Controls	Yes	Yes	Yes	Yes
Applicant FE	Yes	No	Yes	No
Technology FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	46,434	51,143	45,621	50,265
Panel B: 2SLS	Rad	Rad	Ori	Ori
Lag PMR$_{elec}$	-0.0004	-0.0011	0.0069	0.0038
	(0.9417)	(0.8730)	(0.5841)	(0.6320)
	[-0.0081, 0.0203]	[-0.0128, 0.0270]	[-0.0155, 0.0321]	[-0.0185, 0.0211]
Controls	Yes	Yes	Yes	Yes
Applicant FE	Yes	No	Yes	No
Technology FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
F-stat first stage	49.07	58.33	49.07	58.35
First Stage Bootstrap p-value	0.0047	0.0045	0.0047	0.0045
Observations	46434	51,143	45621	50,265

Notes: All specifications also control for the use of policies to promote clean energy innovation, GDP per capita, the price of crude oil imports, the number of applicant, the knowledge stock computed at the applicant level and its squared value. In Panel A, I report wild bootstrap p-values in parentheses and wild bootstrap 95% confidence intervals in square brackets, generated using boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters). In Panel B, regulation in telecommunication is used as instrument for regulation in the electricity sector and inference is based on a wild bootstrapped version of the Anderson-Rubin obtained using the boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters).

Table 1.6: OLS and 2SLS estimates of Model (4) - Heterogeneity by technology class

	(1)	(2)	(3)	(4)	(5)	(6)
	$Renewables$	$OtherTechnologies$	$Storage$	$Renewables$	$OtherTechnologies$	$Storage$
Panel A	Rad	Rad	Rad	Rad	Rad	Rad
PMR$_{elec,c,t-1}$,	0.0030	-0.0321	-0.0176	-0.0049	-0.0227	-0.0206
	(0.7964)	(0.0317)	(0.1411)	(0.4944)	(0.0555)	(0.0229)
	[-0.0205, 0.0226]	[-0.0544, -0.0045]	[-0.0361, 0.0102]	[-0.0250, 0.01]	[-0.0375, 0.0006]	[-0.0334, -0.0075]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Applicant FE	Yes	Yes	Yes	No	No	No
Country FE	Yes	Yes	Yes	Yes	Yes	
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,701	4,843	14,909	19,723	6,565	16,441
Panel B	Rad	Rad	Rad	Rad	Rad	Rad
PMR$_{elec,c,t-1}$,	0.0021	-0.02744	-0.0393	-0.0074	-0.0078	-0.0251
	(0.8530)	(0.3428)	(0.0204)	(0.4877)	(0.8288)	(0.0033)
	[-0.0429, 0.0281]	[-0.0841, 0.0332]	[-0.0992, -0.0040]	[-0.0519, 0.0165]	[-0.0402, 0.0439]	[-0.0337, -0.0165]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Applicant FE	Yes	Yes	Yes	No	No	No
Country FE	Yes	Yes	Yes	Yes	Yes	
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
F-stat first stage	31.39	26.53	46.16	47.42	34.54	63.58
First Stage Bootstrap p-value	0.007	0.0016	0.0001	0.0041	0.0018	0.0006
Observations	16,701	4,843	14,909	19,723	6,565	16,441

Notes: All specifications also control for the use of policies to promote clean energy innovation, GDP per capita, the price of crude oil imports, the number of applicant, the knowledge stock computed at the applicant level and its squared value. In Panel A, I report wild bootstrap p-values in parentheses and wild bootstrap 95% confidence intervals in square brackets, generated using boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters). In Panel B, regulation in telecommunication is used as instrument for regulation in the electricity sector and inference is based on a wild bootstrapped version of the Anderson-Rubin obtained using the boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters).

1.6 Discussion and conclusions

Clean energy technologies are expected to play a vital role in enabling us to achieve the three key objectives of energy security, economic development and environmental sustainability (IEA, 2023). It is therefore particularly important today to understand what factors can contribute to reduce path dependence in polluting technologies and steer technological change towards cleaner alternatives in the electricity sector. The challenge of significantly increasing energy supply while reducing greenhouse gas emissions will require profound changes in electricity supply, which in turn will require substantial R&D spending and the development of new technological solutions (IEA, 2023, 2020). As a result of the recent energy price crisis, changes to the design and regulation of the electricity sector have been discussed in many countries around the world, and some of the ideas on the table have the potential to significantly disrupt market competition (ACER, 2022). The findings presented in this paper shed light on the role that competition in the electricity sector can play in reducing path dependence on dirty technologies and promoting the development of clean-energy innovation.

After the reform of the electricity market, we see a decrease in dirty patents accompanied by an increase in clean energy patents. This increase is driven by patents in electricity storage, renewable energy and combustion technologies with mitigation potential. In addition, clean energy patents developed in a more competitive electricity market rely more on knowledge spillovers from other technological fields. The use of knowledge coming from outside technological fields is expected to benefit the development of clean energy innovations as green patents have been shown to require more diverse knowledge inputs and to rely more heavily on knowledge spillovers (Barbieri et al., 2023, Barbieri et al., 2020; De Marchi, 2012). Furthermore, these kind of spillovers can signal the development of more radical and exploratory innovations (Shane, 2001; Trajtenberg et al., 1997) and can reduce the need for R&D expenditures (IEA, 2020).

The analysis has some limitations that could be addressed in further research. The United States, which is a major developer of clean-energy technologies, could not be included in the sample. This was partly because the latest version of the OECD PMR database (2018) did not include data for the US at the time of writing (Vitale et al., OECD, 2020). More importantly however, a similar analysis for the US would need to be conducted at the state level rather than the federal level, and would therefore require a state-varying indicator to measure electricity regulation. Using an indicator calculated at the federal level would hide the heterogeneous regulatory environment faced by inventors in the different states of the US. This heterogeneity is significant because the US has never enacted a binding federal law to restructure the electricity market, leaving these decisions to the individual states (Joskow, 2008). A similar analysis for the US at the state level is therefore left to future research. Finally, no ex-post indicator of patent quality was included in the analysis. In future research, I plan to investigate whether and how electricity liberalisation changes the ex-post characteristics of clean-energy patents, such as the number of citations they receive or how these citations are distributed across different technological fields. This will allow for a more comprehensive picture of the impact of electricity liberalisation on the characteristics of clean-energy innovation.

References

Abadie, A., Athey, S., Imbens, G. W., Wooldridge, J. M. (2023). When should you adjust standard errors for clustering?. The Quarterly Journal of Economics, 138(1), 1-35. DOI: https://doi.org/10.1093/qje/qjac038

Acemoglu, D., Akcigit, U., Celik, M. A. (2022). Radical and incremental innovation: The roles of firms, managers, and innovators. American Economic Journal: Macroeconomics, 14(3), 199-249. DOI: https://doi.org/10.1257/mac.20170410

Acemoglu, D., Aghion, P., Bursztyn, L., Hemous, D. (2012). The environment and directed technical change. American economic review, 102(1), 131-166. DOI: https://doi.org/10.1257/aer.102.1.131

Agency for the Cooperation of Energy Regulators (ACER) - "ACER's Final Assessment of the EU Wholesale Electricity Market Design" - (2022). Link: https://rb.gy/jzbif6

Aghion, P., Bénabou, R., Martin, R., Roulet, A. (2023). Environmental preferences and technological choices: Is market competition clean or dirty?. American Economic Review: Insights, 5(1), 1-19. DOI: https://doi.org/10.1257/aeri.20210014

Aghion, P., Dechezleprêtre, A., Hemous, D., Martin, R., Van Reenen, J. (2016). Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry. Journal of Political Economy, 124(1), 1-51. DOI: https://doi.org/10.1086/684581

Ahuja, G., Morris Lampert, C. (2001). Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions. Strategic management journal, 22(6-7), 521-543. DOI: https://doi.org/10.1002/smj.176

Akcigit, U., Kerr, W. R. (2018). Growth through heterogeneous innovations. Journal of Political Economy, 126(4), 1374-1443. DOI: https://doi.org/10.1086/697901

Anderson, T. W., Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. The Annals of mathematical statistics, 20(1), 46-63. DOI: https://doi.org/10.1214/aoms/1177730090

Andrews, I., Stock, J. H., Sun, L. (2019). Weak instruments in instrumental variables regression: Theory and practice. Annual Review of Economics, 11, 727-753. DOI: https://doi.org/10.1146/annurev-economics-080218-025643

Arthur, W. B. (2007). The structure of invention. Research policy, 36(2), 274-287. DOI: https://doi.org/10.1016/j.respol.2006.11.005

Barbieri, N., Marzucchi, A., Rizzo, U. (2020). Knowledge sources and impacts on subsequent inventions: Do green technologies differ from non-green ones?. Research Policy, 49(2), 103901. DOI: https://doi.org/10.1016/j.respol.2019.103901

Batlle, C., Ocaña, C. (2013). Electricity regulation: principles and institutions. In Regulation of the power sector (pp. 125-150). Springer, London. Link: https://link.springer.com/book/10.1007/978-1-4471-5034-3

Batlle, C., Schittekatte, T., Knittel, C. R. (2022). Power price crisis in the EU 2.0+: Desperate times call for desperate measures. Available at SSRN 4074014. DOI: http://dx.doi.org/10.2139/ssrn.4074014

Baudry, M., Dumont, B. (2006). Comparing firms' triadic patent applications across countries: Is there a gap in terms of RD effort or a gap in terms of performances?. Research Policy, 35(2), 324-342. DOI: https://doi.org/10.1016/j.respol.2005.12.004

Calel, R., Dechezleprêtre, A. (2016). Environmental policy and directed technological change: evidence from the European carbon market. Review of economics and statistics, 98(1), 173-191. DOI: https://doi.org/10.1162/REST_a_00470

Cambini, C., Caviggioli, F., Scellato, G. (2016). Innovation and market regulation: evidence from the European electricity industry. Industry and Innovation, 23(8), 734-752. https://doi.org/10.1080/13662716.2016.1206464

Carley, S. (2009). State renewable energy electricity policies: An empirical evaluation of effectiveness. Energy policy, 37(8), 3071-3081. DOI: https://doi.org/10.1016/j.enpol.2009.03.062

Conner, K. R. (1988). Strategies for product cannibalism. Strategic Management Journal, 9(S1), 9-26. DOI: https://doi.org/10.1002/smj.4250090704

Conti, C., Mancusi, M. L., Sanna-Randaccio, F., Sestini, R., Verdolini, E. (2018). Transition towards a green economy in Europe: Innovation and knowledge integration in the renewable energy sector. Research Policy, 47(10), 1996-2009.

Correia, S. (2015). Singletons, cluster-robust standard errors and fixed effects: A bad mix. Technical Note, Duke University, 7. Link: http://scorreia.com/research/singletons.pdf

Correia, S. (2016). reghdfe: Estimating linear models with multi-way fixed effects. In 2016 Stata Conference (No. 24). Stata Users Group. Link: http://fmwww.bc.edu/repec/chic2016/chicago16_correia.pdf

Dahlin, K. B., Behrens, D. M. (2005). When is an invention really radical?: Defining and measuring technological radicalness. research policy, 34(5), 717-737.

Davidson, R., MacKinnon, J. G. (2010). Wild bootstrap tests for IV regression. Journal of Business Economic Statistics, 28(1), 128-144. DOI: https://doi.org/10.1198/jbes.2009.07221

Defeuilley, C. (2019). Energy transition and the future (s) of the electricity sector. Utilities Policy, 57, 97-105. DOI: https://doi.org/10.1016/j.jup.2019.03.002

De Marchi, V. (2012). Environmental innovation and RD cooperation: Empirical evidence from Spanish manufacturing firms. Research policy, 41(3), 614-623. DOI: https://doi.org/10.1016/j.respol.2011.10.002

De Noni, I., Ganzaroli, A., Pilotti, L. (2021). Spawning exaptive opportunities in European regions: The missing link in the smart specialization framework. Research Policy, 50(6), 104265. DOI: https://doi.org/10.1016/j.respol.2021.104265

Delmas, M. A., Montes-Sancho, M. J. (2011). US state policies for renewable energy: Context and effectiveness. Energy Policy, 39(5), 2273-2288. DOI: https://doi.org/10.1016/j.enpol.2011.01.034

Dooley, J. J. (1998). Unintended consequences: energy RD in a deregulated energy market. Energy Policy, 26(7), 547-555. DOI: https://doi.org/10.1016/S0301-4215(97)00166-3

Dolphin, G., Pollitt, M. (2020). Identifying Innovative Actors in the Electricicity Supply Industry Using Machine Learning: An Application to UK Patent Data. EPRG Working Paper 2004 - Cambridge Working Paper in Economics. Link: https://www.jstor.org/stable/pdf/resrep30421.pdf

Hall, B. H. (2005). Measuring the Returns to RD: The Depreciation Problem. Annales d'Économie et de Statistique, 79/80, 341–381. DOI: https://doi.org/10.2307/20777581

Hall, B. H., Jaffe, A. B., Trajtenberg, M. (2001). The NBER patent citation data file: Lessons, insights and methodological tools. NBER Working Paper Series. DOI: https://doi.org/10.3386/w8498

Hall, P. (2013). The bootstrap and Edgeworth expansion. Springer Science Business Media. Link: https://people.eecs.berkeley.edu/~jordan/sail/readings/edgeworth.pdf

Hausman, J. A. (1978). Specification tests in econometrics. Econometrica: Journal of the econometric society, 1251-1271. DOI: https://doi.org/10.2307/1913827

Higham, K., De Rassenfosse, G., Jaffe, A. B. (2021). Patent quality: Towards a systematic framework for analysis and measurement. Research Policy, 50(4), 104215. DOI: https://doi.org/10.1016/j.respol.2021.104215

Høj, J., Galasso, V., Nicoletti, G., Dang, T. T. (2006). The political economy of structural reform: empirical evidence from OECD countries. DOI: https://doi.org/10.1787/18151973

Hughes, T.P. (1987). The evolution of large technological systems. In: Bijker, W., Hughes, T.P., Pinch, T. (Eds.), The Social Construction of Technological Systems. Cambridge, MA, pp. 51–82. Link https://rb.gy/gecjk6

Hussinger, K., Pacher, S. (2019). Information ambiguity, patents and the market value of innovative assets. Research policy, 48(3), 665-675. DOI: https://doi.org/10.1016/j.respol.2018.10.022

IEA (2023), Energy Technology Perspectives 2023, IEA, Paris https://www.iea.org/reports/energy-technology-perspectives-2023, License: CC BY 4.0

IEA (2022a), Global Hydrogen Review 2022, IEA, Paris https://www.iea.org/reports/global-hydrogen-review-2022, License: CC BY 4.0

IEA (2022b), World Energy Outlook 2022, IEA, Paris https://www.iea.org/reports/world-energy-outlook-2022?language=it, License: CC BY 4.0 (report); CC BY NC SA 4.0 (Annex A)
IEA (2020), Energy Technology Perspectives 2020, IEA, Paris https://www.iea.org/reports/energy-technology-perspectives-2020, License: CC BY 4.0

IEA (2019a), Nuclear Power in a Clean Energy System, IEA, Paris https://www.iea.org/reports/nuclear-power-in-a-clean-energy-system, License: CC BY 4.0

IEA (2019b), The Future of Hydrogen, IEA, Paris https://www.iea.org/reports/the-future-of-hydrogen, License: CC BY 4.0

Jacobsson, S., Bergek, A. (2004). Transforming the energy sector: the evolution of technological systems in renewable energy technology. Industrial and corporate change, 13(5), 815-849. DOI: https://doi.org/10.1093/icc/dth032

Jamasb, T., Pollitt, M. (2008). Liberalisation and RD in network industries: The case of the electricity industry. Research Policy, 37(6-7), 995-1008. DOI: https://doi.org/10.1016/j.respol.2008.04.010

Jamasb, T., Pollitt, M. G. (2011). Electricity sector liberalisation and innovation: An analysis of the UK's patenting activities. Research Policy, 40(2), 309-324. DOI: https://doi.org/10.1016/j.respol.2010.10.010

Joerges, B. (1996). Large technical systems and the discourse of complexity. Complex technical systems, 55-72. Link: https://rb.gy/h6pop1

Joskow, P. L. (2008). Lessons learned from electricity market liberalization. The Energy Journal, 29(Special Issue 2). DOI: DOI:10.5547/ISSN0195-6574-EJ-Vol29-NoSI2-3

Kafouros, M., Aliyev, M., Krammer, S. M. (2021). Do firms profit from patent litigation? The contingent roles of diversification and intangible assets. Research Policy, 50(6), 104263. DOI: https://doi.org/10.1016/j.respol.2021.104263

Keane, M., Neal, T. (2023). Instrument strength in IV estimation and inference: A guide to theory and practice. Journal of Econometrics. DOI: https://doi.org/10.1016/j.jeconom.2022.12.009

Kruse, T., Dechezleprêtre, A., Saffar, R., Robert, L. (2022). Measuring environmental policy stringency in OECD countries: An update of the OECD composite EPS indicator. OECD Economics Department Working Papers. DOI: https://doi.org/10.1787/90ab82e8-en

Klepper, S. (1996). Entry, exit, growth, and innovation over the product life cycle. The American economic review, 562-583. Link https://www.jstor.org/stable/2118212

Kline, P., Santos, A. (2012). A score based approach to wild bootstrap inference. Journal of Econometric Methods, 1(1), 23-41. DOI: https://doi.org/10.1515/2156-6674.1006

Lal, A., Lockhart, M., Xu, Y., Zu, Z. (2023). How much should we trust instrumental variable estimates in political science? Practical advice based on over 60 replicated studies. arXiv preprint arXiv:2303.11399. DOI: https://doi.org/10.48550/arXiv.2303.11399

Lee, D. S., McCrary, J., Moreira, M. J., Porter, J. (2022). Valid t-ratio Inference for IV. American Economic Review, 112(10), 3260-3290. DOI: https://doi.org/10.1257/aer.20211063

Li, D., Heimeriks, G., Alkemade, F. (2020). The emergence of renewable energy technologies at country level: relatedness, international knowledge spillovers and domestic energy markets. Industry and Innovation, 27(9), 991-1013. DOI: https://doi.org/10.1080/13662716.2020.1713734

Mays, J., Craig, M. T., Kiesling, L., Macey, J. C., Shaffer, B., Shu, H. (2022). Private risk and social resilience in liberalized electricity markets. Joule, 6(2), 369-380. DOI: https://doi.org/10.1016/j.joule.2022.01.004

Marino, M., Parrotta, P., Valletta, G. (2019). Electricity (de) regulation and innovation. Research policy, 48(3), 748-758. DOI: https://doi.org/10.1016/j.respol.2018.11.005

Markard, J., Bento, N., Kittner, N., Nuñez-Jimenez, A. (2020). Destined for decline? Examining nuclear energy from a technological innovation systems perspective. Energy Research Social Science, 67, 101512. DOI: https://doi.org/10.1016/j.erss.2020.101512

Markard, J., Truffer, B. (2006). Innovation processes in large technical systems: Market liberalization as a driver for radical change?. Research policy, 35(5), 609-625. DOI: https://doi.org/10.1016/j.respol.2006.02.008

Negro, S. O., Alkemade, F., Hekkert, M. P. (2012). Why does renewable energy diffuse so slowly? A review of innovation system problems. Renewable and sustainable energy reviews, 16(6), 3836-3846. DOI: https://doi.org/10.1016/j.rser.2012.03.043

Nelson, R. R. (1985). An evolutionary theory of economic change. harvard university press. Link: https://www.hup.harvard.edu/catalog.php?isbn=9780674272286

Nesta, L., Vona, F., Nicolli, F. (2014). Environmental policies, competition and innovation in renewable energy. Journal of Environmental Economics and Management, 67(3), 396-411. DOI: https://doi.org/10.1016/j.jeem.2014.01.001

Newbery, D. (2010). A nuclear future? UK Government policy and the role of the market. Economic Affairs, 30(2), 21-27. DOI: https://doi.org/10.1111/j.1468-0270.2010.02001.x

Nicolini, M., Tavoni, M. (2017). Are renewable energy subsidies effective? Evidence from Europe. Renewable and Sustainable Energy Reviews, 74, 412-423. DOI: https://doi.org/10.1016/j.rser.2016.12.032

Nicolli, F., Vona, F. (2016). Heterogeneous policies, heterogeneous technologies: The case of renewable energy. Energy Economics, 56, 190-204. DOI: https://doi.org/10.1016/j.eneco.2016.03.007

Nicolli, F., Vona, F. (2019). Energy market liberalization and renewable energy policies in OECD countries. Energy Policy, 128, 853-867. DOI: https://doi.org/10.1016/j.enpol.2019.01.018

OECD (2015), Emerging technologies firm dynamics: The implications of green growth, OECD, Paris. Link: https://issue.com/oecd.publishing/docs/ggsd_2015_issuepaper2_finalpdf

Olsson, O. (2000). Knowledge as a set in idea space: An epistemological view on growth. Journal of economic growth, 5(3), 253-275. DOI: https://doi.org/10.1023/A:1009829601155

Pollitt, M. G. (2012). The role of policy in energy transitions: Lessons from the energy liberalisation era. Energy policy, 50, 128-137. DOI: https://doi.org/10.1016/j.enpol.2012.03.004

Popp, David, Pless, Jacquelyn, Haščič, Ivan and Johnstone, Nick. ”4. Innovation and Entrepreneurship in the Energy Sector”. The Role of Innovation and Entrepreneurship in Economic Growth, edited by Michael J Andrews, Aaron Chatterji, Josh Lerner and Scott Stern, Chicago: University of Chicago Press, 2022, pp. 175-248. DOI: https://doi.org/10.7208/chicago/9780226810645-006

Popp, D. (2019). Environmental Policy and Innovation: A Decade of Research. International Review of Environmental and Resource Economics, 13(3-4), 265-337. DOI: https://doi.org/10.1561/101.00000111

Porter, C., Serra, D. (2020). Gender differences in the choice of major: The importance of female role models. American Economic Journal: Applied Economics, 12(3), 226-54. DOI: https://doi.org/10.1257/app.20180426

Quintana-García, C., Benavides-Velasco, C. A. (2008). Innovative competence, exploration and exploitation: The influence of technological diversification. Research policy, 37(3), 492-507. DOI: https://doi.org/10.1016/j.respol.2007.12.002

Reinganum, J. F. (1983). Uncertain innovation and the persistence of monopoly. The American Economic Review, 73(4), 741-748. Link: https://www.jstor.org/stable/1816571

Rizzo, U., Barbieri, N., Ramaciotti, L., Iannantuono, D. (2020). The division of labour between academia and industry for the generation of radical inventions. The Journal of Technology Transfer, 45(2), 393-413. DOI: https://doi.org/10.1007/s10961-018-9688-y

Roodman, D., Nielsen, M. Ø., MacKinnon, J. G., Webb, M. D. (2019). Fast and wild: Bootstrap inference in Stata using boottest. The Stata Journal, 19(1), 4-60. DOI: https://doi.org/10.1177/1536867X1983087

Rosenkopf, L., Nerkar, A. (2001). Beyond local search: boundary-spanning, exploration, and impact in the optical disk industry. Strategic management journal, 22(4), 287-306. DOI: https://doi.org/10.1002/smj.160

Sanyal, P., Ghosh, S. (2013). Product market competition and upstream innovation: evidence from the US electricity market deregulation. Review of Economics and Statistics, 95(1), 237-254. DOI: https://doi.org/10.1162/REST_a_00255

Sterlacchini, A. (2012). Energy RD in private and state-owned utilities: An analysis of the major world electric companies. Energy policy, 41, 494-506. DOI: https://doi.org/

10.1016/j.enpol.2011.11.010

Shane, S. (2001). Technological opportunities and new firm creation. Management science, 47(2), 205-220. DOI: https://doi.org/10.1287/mnsc.47.2.205.9837

Sapsalis, E., de la Potterie, B. V. P., Navon, R. (2006). Academic versus industry patenting: An in-depth analysis of what determines patent value. Research Policy, 35(10), 1631-1645. DOI: https://doi.org/10.1016/j.respol.2006.09.014

Schmoch, U. (2008). Concept of a technology classification for country comparisons. Final report to the world intellectual property organisation (wipo), WIPO. Link: http://www.world-intellectual-property-organization.com/edocs/mdocs/classifications/en/ipc_ce_41/ipc_ce_41_5-annex1.pdf

Squicciarini, M., Dernis, H., Criscuolo, C. (2013). Measuring patent quality: Indicators of technological and economic value. OECD Science, Technology and Industry Working Papers. DOI: https://doi.org/10.1787/5k4522wkw1r8-en

Trajtenberg, M., Henderson, R., Jaffe, A. (1997). University versus corporate patents: A window on the basicness of invention. Economics of Innovation and new technology, 5(1), 19-50. DOI: https://doi.org/10.1080/10438599700000006

Verhoeven, D., Bakker, J., Veugelers, R. (2016). Measuring technological novelty with patent-based indicators. Research policy, 45(3), 707-723. DOI: https://doi.org/10.1016/j.respol.2015.11.010

Vitale, C., Bitetti, R., Wanner, I., Danitz, E., Moiso, C. (2020). The 2018 edition of the OECD PMR indicators and database: Methodological improvements and policy insights. OECD Economics Department Working Papers. DOI: https://doi.org/10.1787/2cfb622f-en

Wang, W., Zhang, Y. (2021). Wild bootstrap for instrumental variables regressions with weak and few clusters. arXiv preprint arXiv:2108.13707. DOI: https://doi.org/10.48550/arXiv.2108.13707

Wang, N., Mogi, G. (2017). Deregulation, market competition, and innovation of utilities: Evidence from Japanese electric sector. Energy Policy, 111, 403-413. DOI: https://doi.org/10.1016/j.enpol.2017.09.044

Winter, S. G. (1984). Schumpeterian competition in alternative technological regimes. Journal of Economic Behavior Organization, 5(3-4), 287-320. DOI: https://doi.org/10.1016/0167-2681(84)90004-0

Wooldridge, J. M. (2015). Control function methods in applied econometrics. Journal of Human Resources, 50(2), 420-445.

Young, A. (2022). Consistency without inference: Instrumental variables in practical application. European Economic Review, 147, 104112. DOI: https://doi.org/10.1016/j.euroecorev.2022.104112

1.7 Appendix A

A.1 Y02E technologies

Y02E / 10: Energy generation through renewable energy sources:

• Geothermal energy

• Hydro energy

• Energy from the sea, e.g. using wave energy or salinity gradient

• Solar thermal energy, e.g. solar towers

• Photovoltaic [PV] energy

• Thermal-PV hybrids

• Wind energy

Y02E / 20: Combustion technologies with mitigation potential:

• Heat utilisation in combustion or incineration of waste

• Combined heat and power generation [CHP]

• Combined cycle power plant [CCPP], or combined cycle gas turbine [CCGT]

• Technologies for a more efficient combustion or heat usage

• Direct CO2 mitigation

• Indirect CO2mitigation, i.e. by acting on non CO2directly related matters of the process, e.g. pre-heating or heat recovery

Y02E / 30: Energy generation of nuclear origin:

• Nuclear fusion reactors

• Nuclear fission reactors

Y02E / 40: Technologies for an efficient electrical power generation,transmission or distribution:

• Flexible AC transmission systems [FACTS]

• Active power filtering [APF]

• Reactive power compensation

• Arrangements for reducing harmonics

• Arrangements for eliminating or reducing asymmetry in polyphase networks

• Superconducting electric elements or equipment; Power systems integrating superconducting elements or equipment

• Smart grids as climate change mitigation technology in the energy generation sector

Y02E / 50: Technologies for the production of fuel of non-fossil origin:

• Biofuels, e.g. bio-diesel

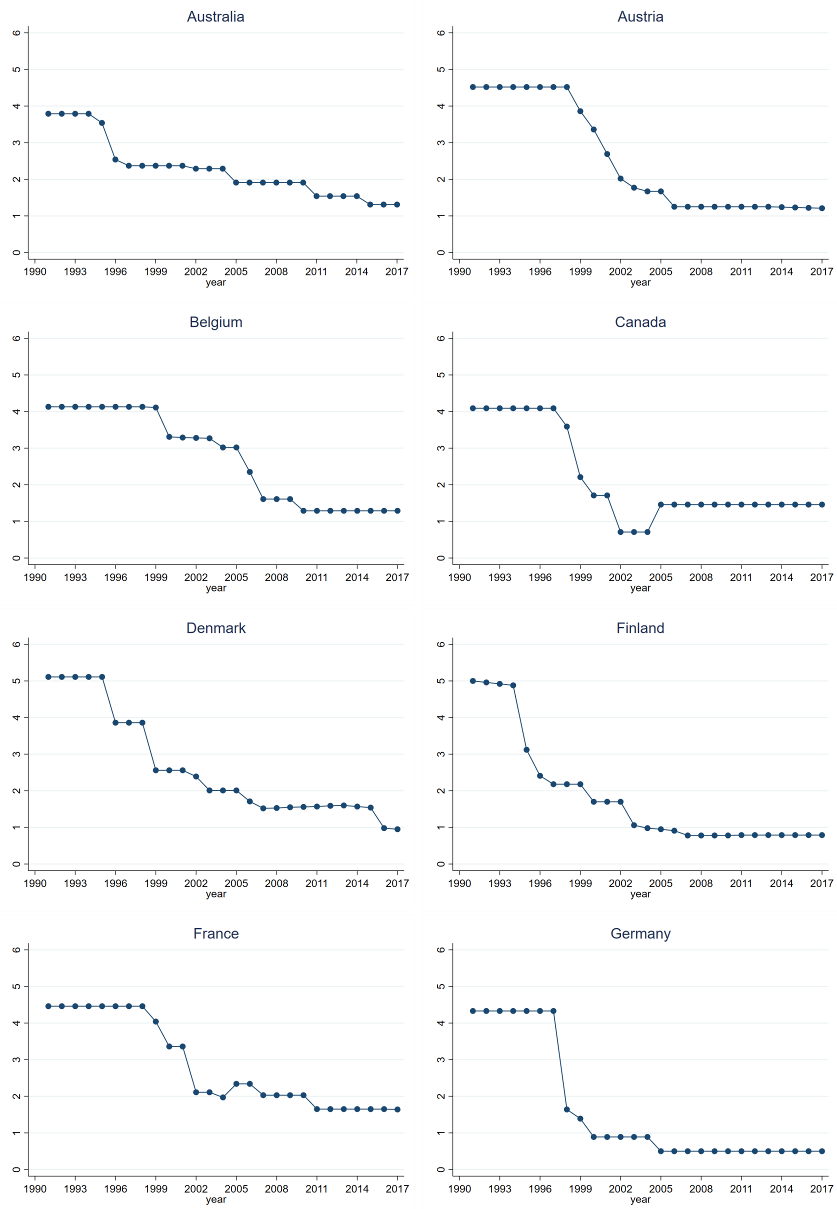
• Fuel from waste, e.g. synthetic alcohol or diesel

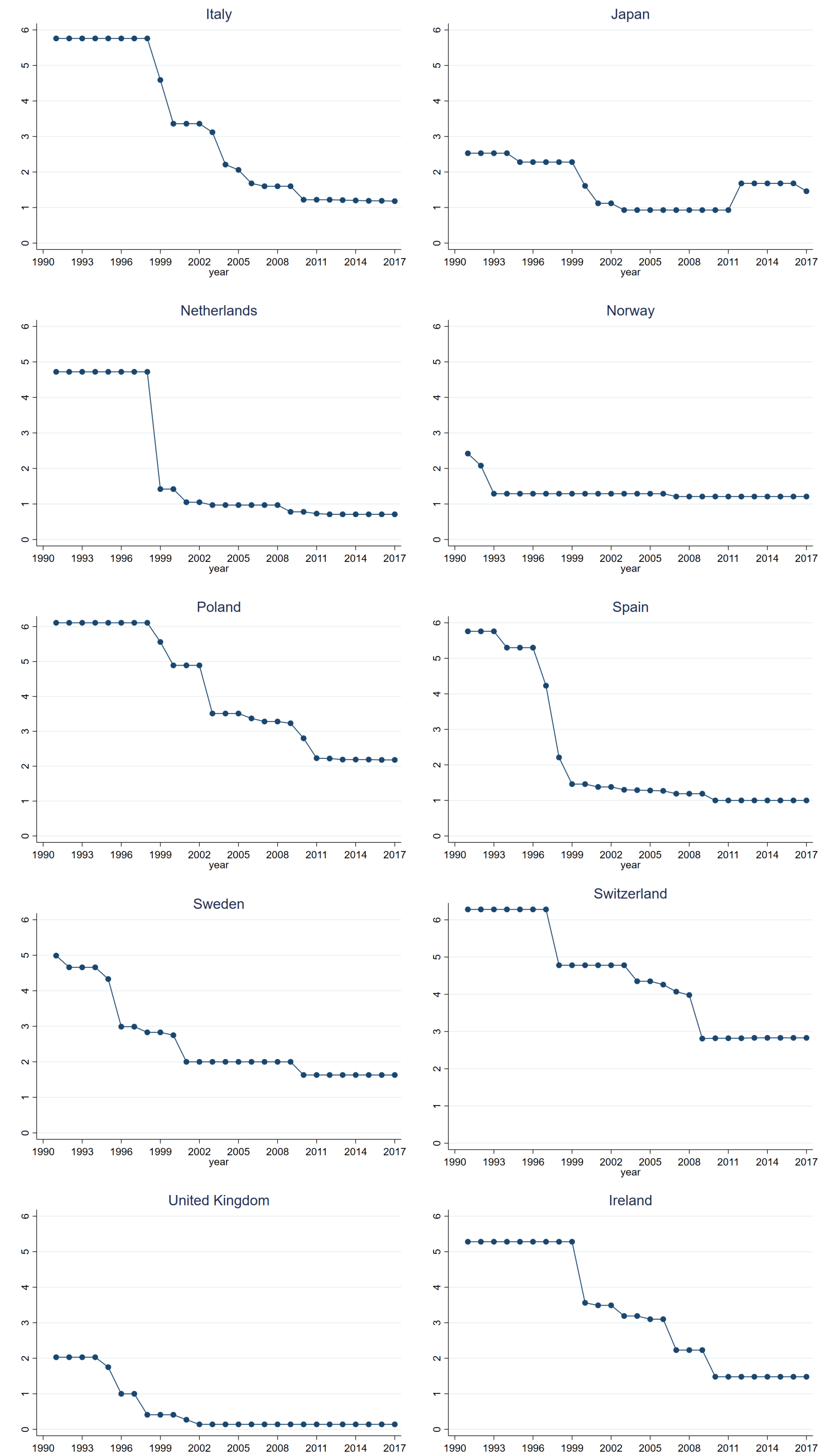
Y02E / 60: Enabling technologies; Technologies with a potential or indirect contribution to GHG emissions mitigation

• Energy storage using batteries
• Energy storage using capacitors
• Thermal energy storage
• Mechanical energy storage, e.g. flywheels or pressurised fluids
• Hydrogen technology
• Smart grids in the energy sector

Y02E / 70 :Other energy conversion or management systems reducing GHG emission

A.2 Evolution of the PMR_{elec} index by Country




Italy
Japan
Netherlands
Norway
Poland
Spain
Sweden
Switzerland
United Kingdom
Ireland
0
1
2
3
4
5
6
1990
1993
1996
1999
2002
2005
2008
2011
2014
2017
year

A.3 Fractional Probit Model

Table A.1 presents the results for the estimation of model (4) using a fractional probit model with no applicant fixed effects and relying on the score bootstrap for inference. These results confirm what we see in Table 1.4 in terms of sign, magnitude and statistical significance. In particular, the marginal effects in column (1) and (2) are -0.017 and -0.0022 respectively, both of which are very similar to the estimated coefficient in columns (2) and (4) of Table 1.4. Similarly, the marginal effects in columns (3) and (4) are -0.0031 and -0.0021 respectively. Once again these values are similar to what we see in columns (2) and (4) of Table 1.5.

Table A.1: Fractional Probit estimates of model (4)

	(1)	(2)	(3)	(4)
	Clean	Clean	Dirty	Dirty
	Rad	Ori	Rad	Ori
Lag PMR$_{elec}$	-0.0482	-0.0065	-0.0088	-0.0061
	(0.0011)	(0.1183)	(0.4871)	(0.6306)
Controls	Yes	Yes	Yes	Yes
Applicant FE	No	No	No	No
Technology FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	42,834	42,797	51,143	50,305

Notes: Fractional probit regressions. Score bootstrap cluster p-values in parentheses are generated using boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters).

A.4 Alternative patent family correction

As discussed in Section 1.3, different patent applications belonging to the same family may have different values for the same patent level indicator. Since we want to focus on patent families rather than patent applications in order to avoid double counting, this poses a problem. I deal with this problem by selecting only the patent with the highest value for the indicator of interest from each patent family (Barbieri et al., 2020; Verhoeven et al., 2016). Table A.2 shows that the results are robust to selecting the application with the lowest value of the indicator of interest from each family.

Table A.2: Patent Family correction based on the lowest value of the indicator of interest

	(1)	(2)	(3)	(4)
	Clean	Clean	Clean	Clean
Panel A: OLS estimates	Rad	Rad	Ori	Ori
Lag PMR_{elec}	-0.0170	-0.0171	-0.0007	-0.0006
	(0.0079)	(0.0002)	(0.7606)	(0.7435)
	[-0.0313, -0.0049]	[-0.0290, -0.0111]	[-0.0102, 0.0036]	[-0.0060, 0.0025]
Controls	Yes	Yes	Yes	Yes
Applicant FE	Yes	No	Yes	No
Technology FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	37,371	42,834	37,365	42,828
Panel B: 2SLS estimates (telecom)	Rad	Rad	Ori	Ori
Lag PMR_{elec}	-0.0210	-0.0125	0.0025	0.0059
	(0.0000)	(0.0409)	(0.6642)	(0.1005)
	[-0.0367, -0.0130]	[-0.0210, -0.0005]	[-0.0198, 0.0112]	[-0.0007, 0.0145]
Controls	Yes	Yes	Yes	Yes
Applicant FE	Yes	No	Yes	No
Other FEs	Yes	Yes	Yes	Yes
F-stat first stage	41.85	53.78	41.71	53.72
First Stage Bootstrap p-value	0.0006	0.0014	0.0006	0.0014
Observations	37,371	42,834	37,365	42,828

Notes: All specifications also control for the use of policies to promote clean energy innovation, GDP per capita, the price of crude oil imports, the number of applicant, the knowledge stock computed at the applicant level and its squared value. In Panel A, I report wild bootstrap p-values in parentheses and wild bootstrap 95% confidence intervals in square brackets, generated using boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters). In Panel B, regulation in telecommunication is used as instrument for regulation in the electricity sector and inference is based on a wild bootstrapped version of the Anderson-Rubin obtained using the boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters).

A.5 ”International” Applications

As discussed in Section 1.3, I exclude patents with inventors in more than one country from the patent-level analysis. The reason is that these patents are relatively rare and there is no obvious way to deal with them in this framework (as opposed, for example, to using the fractional count when we study the number of patents filed by country). Table A.3 shows the results when these ”international” patents are included and assigned to the most liberalised country among the countries of inventors. This allows to test whether the highest degree of electricity liberalisation to which a patent is exposed has an impact on its exploratory nature. The results are robust to this exercise.

Table A.3: ”International” patent applications

	(1)	(2)	(3)	(4)
	Clean	Clean	Clean	Clean
Panel A: OLS estimates	Rad	Rad	Ori	Ori
Lag PMR$_{elec}$	-0.0171	-0.0173	-0.0025	-0.0021
	(0.0430)	(0.0015)	(0.3209)	(0.1351)
	[-0.0329, -0.0010]	[-0.0280, -0.0115]	[-0.0125, 0.0037]	[-0.0062, 0.0009]
Controls	Yes	Yes	Yes	Yes
Applicant FE	Yes	No	Yes	No
Technology FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	38,676	44,228	38,302	43,850
Panel B: 2SLS estimates (telecom)	Rad	Rad	Ori	Ori
Lag PMR$_{elec}$	-0.0252	-0.0142	-0.0028	0.0020
	(0.0000)	(0.0338)	(0.6528)	(0.5193)
	[-0.0470, -0.0147]	[-0.0230, -0.0018]	[-0.0359, 0.0075]	[-0.0038, 0.0079]
Controls	Yes	Yes	Yes	Yes
Applicant FE	Yes	No	Yes	No
Other FEs	Yes	Yes	Yes	Yes
F-stat first stage	38.01	48.68	37.25	48.26
First Stage Bootstrap p-value	0.0006	0.0017	0.0008	0.0017
Observations	38,676	44,228	38,302	43,850

Notes: All specifications also control for the use of policies to promote clean energy innovation, GDP per capita, the price of crude oil imports, the number of applicant, the knowledge stock computed at the applicant level and its squared value. In Panel A, I report wild bootstrap p-values in parentheses and wild bootstrap 95% confidence intervals in square brackets, generated using boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters). In Panel B, regulation in telecommunication is used as instrument for regulation in the electricity sector and inference is based on a wild bootstrapped version of the Anderson-Rubin obtained using the boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters).

A.6 Heterogeneity of the top 1% applicants for patent families filed

The top 1% of applicants in terms of patents filed account for about 50% of all patent families in the sample. Table A.4 tests whether the effect of electricity liberalisation is driven by these applicants and finds no evidence of this pattern. This test is performed by interacting the variable PMR_{elec} with a variable equal to one for applicants in the top 1% and zero otherwise. The coefficient on the interaction term is not significant in all specifications, indicating that the effect for the top 1% of highly innovative applicants is not significantly different from that for other applicants.

Table A.4: Results without the top 1% applicants for patent families filed

	(1)	(2)	(3)	(4)
	Clean	Clean	Clean	Clean
Panel A: OLS estimates of model (3)	Rad	Rad	Ori	Ori
Lag PMR$_{elec}$	-0.0184	-0.01816	-0.0020	-0.0003
	(0.0470)	(0.0020)	(0.4945)	(0.8258)
	[-0.0365, -0.0004]	[-0.0285, -0.0116]	[-0.0159, 0.0028]	[-0.0045, 0.0021]
Lag PMR$_{elec}$ # Big	0.0024	-0.0008	-0.0002	-0.006
	(0.5506)	(0.8799)	(0.9099)	(0.2372)
	[-0.0082, 0.0115]	[-0.0170, 0.0127]	[-0.0038, 0.0112]	[-0.0179, 0.0109]
Controls	Yes	Yes	Yes	Yes
Applicant FE	Yes	No	Yes	No
Other FEs	Yes	Yes	Yes	Yes
Observations	37,371	42,834	37,333	42,797
Panel B: OLS estimates of model (3)	Rad	Rad	Ori	Ori
Lag PMR$_{elec}$	-0.0287	-0.0168	-0.0024	0.0041
	(0.0060)	(0.0200)	(0.6917)	(0.2142)
	[-0.0539, -0.0151]	[-0.0306, -0.0054]	[-0.0301, 0.0069]	[-0.0052, 0.0146]
Lag PMR$_{elec}$ # Big	0.007	-0.0017	0.0014	-0.0118
	(0.1421)	(0.7808)	(0.7678)	(0.3634)
	[-0.0050, 0.0208]	[-0.0174, 0.0216]	[-0.0069, 0.0186]	[-0.0344, 0.0192]
Controls	Yes	Yes	Yes	Yes
Applicant FE	Yes	No	Yes	No
Other FEs	Yes	Yes	Yes	Yes
Observations	37,371	42,834	37,333	42,797

Notes: All specifications also control for the use of policies to promote clean energy innovation, GDP per capita, the price of crude oil imports, the number of applicant, the knowledge stock computed at the applicant level and its squared value. In Panel A, I report wild bootstrap p-values in parentheses and wild bootstrap 95% confidence intervals in square brackets, generated using boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters). In Panel B, regulation in telecommunication is used as instrument for regulation in the electricity sector and inference is based on a wild bootstrapped version of the Anderson-Rubin obtained using the boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters).

A.7 Sensitivity analysis: controlling for patent characteristics

Previous literature using the radicalness and originality indices as dependent variables often controls for patent characteristics that are intimately related to how these indicators are built. In particular, for the radicalness index these characteristics are the number of technological fields to which the patent is allocated (patent scope) and the number of backward citations (Barbieri et al. 2020; Sapsalis et al., 2006; Hall et al., 2001). Similarly, for the originality index the number of backward citations is often included as control variable (Barbieri et al. 2020; Hall et al., 2001).

In this setting the inclusion of these variables as control variables is potentially problematic as they are likely to be ”bad controls” (Cinelli et al., 2022). That being said, in this section I test the sensitivity of the main results presented in the paper to the inclusion of this variables in order to gauge how much their inclusion would change the results obtained with the main specification. I do this by adding the number of backward citations among control variables in model (4) when the dependent variable is the originality index, and by adding both the number of backward citations and the patent scope when the dependent variable is the radicalness index. Figures A.1 and A.2 show the estimated effect of electricity liberalisation in model (4) both with and without controlling for these patent-level variables. The specifications used in these figures are the following:

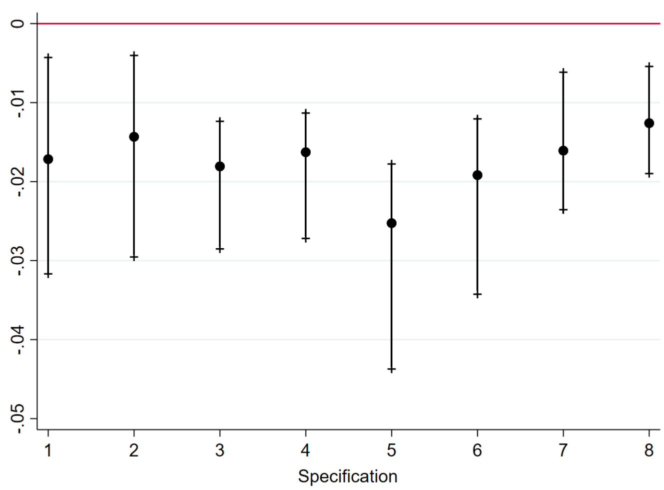
- Specification 1: OLS estimates of Model (4), with applicant fixed effect, as reported in Table 1.4
- Specification 2: OLS estimates of Model (4), with applicant fixed effect and with patent-level control variables
- Specification 3: OLS estimates of Model (4), without applicant fixed effect, as reported in Table 1.4
- Specification 4: OLS estimates of Model (4) without applicant fixed effects and with patent-level control variables
- Specification 5: 2SLS estimates of Model (4), with applicant fixed effect, as reported in Table 1.5
- Specification 6: 2SLS estimates of Model (4), with applicant fixed effect and with patent-level control variables

- Specification 7: 2SLS estimates of Model (4), without applicant fixed effect, as reported in Table 1.5

- Specification 8: 2SLS estimates of Model (4), without applicant fixed effect and with patent-level control variables

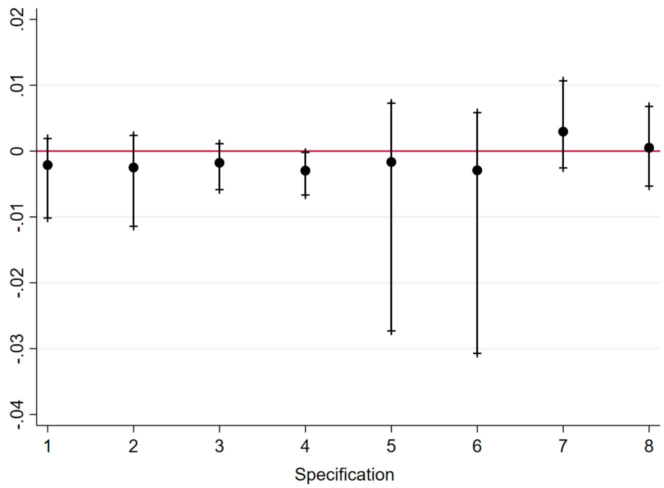
When included in the regressions, patent-level control variables are always significant and of the expected sign. However, their inclusion does not change meaningfully the estimated coefficient for PMR_elec and its significance level. The only exception is the coefficient on Specification 4 for the originality index, which becomes significant at the 5% level with the inclusion of backward citations among the control variables. That being said, this section shows that the results and conclusions of the analysis would not change even including patent-level control variables among the regressors.

Figure A.1: Sensitivity to patent-level control variables



Notes: Coefficient and confidence interval for PMR_elec in the estimates of specification 1 through 8 (see text). All specifications control for the use of policies to promote clean energy innovation, GDP per capita, the price of crude oil imports, the number of applicant, the knowledge stock computed at the applicant level and its squared value. Specifications 2,4,5 and 8 also control for patent-level characteristics. For specifications 1 through 4 inference is based on wild bootstrap (Roodman et al., 2019), standard errors are obtained using the boottest command in Stata (Roodman et al., 2019) and clustered at the country level (18 clusters). For specifications 5 through 8 inference, regulation in telecommunication is used as instrument for regulation in the electricity sector and inference is based on a wild bootstrapped version of the Anderson-Rubin obtained using the boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters).

Figure A.2: Sensitivity to patent-level control variables



Notes: Coefficient and confidence interval for PMR_elec in the estimate of specification 1 through 8 (see text). All specifications control for the use of policies to promote clean energy innovation, GDP per capita, the price of crude oil imports, the number of applicant, the knowledge stock computed at the applicant level and its squared value. Specifications 2,4,5 and 8 also control for patent-level characteristics. For specifications 1 through 4 inference is based on wild bootstrap (Roodman et al., 2019), standard errors are obtained using the boottest command in Stata (Roodman et al., 2019) and clustered at the country level (18 clusters). For specifications 5 through 8 inference, regulation in telecommunication is used as instrument for regulation in the electricity sector and inference is based on a wild bootstrapped version of the Anderson-Rubin obtained using the boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters).

A.8 Results with Patents in Hydrogen and Nuclear related technologies

Tables A.5 and A.6 shows the results when hydrogen and nuclear patents are included in the sample. In column (1) of Table A.5 we see that the effect of electricity liberalisation has the expected sign, but is no longer significant. The results in column (2) and (3) focus only on nuclear and hydrogen patents and shed light on why this is so. The relationship between electricity liberalisation and the number of patents filed in these technological fields is far from being statistically significant and the estimated coefficient has the opposite sign compared to the results of Tables 1.3 and 1.4. Note that the number of observations in column (2) is lower because in the estimation countries with no patents in nuclear energy technologies for all the time span of the analysis are dropped.

On the contrary, the results shown in Table A.6 are remarkably similar to those presented in Table 1.4. Finally, Table A.7 presents the results when focusing on hydrogen related technologies. Results focusing only on nuclear patents are not reported due to the low number of patents in this technologies and thus the low number of observations for running the regression.

Table A.5: Results with Patents in Hydrogen and Nuclear related technologies - model (3)

	(1)	(2)	(3)
Panel A	$Count_{ct}^{AllY02E}$	$Count_{ct}^{Nuclear}$	$Count_{ct}^{Hydrogen}$
$\text{PMR}_{elec,c,t-1,}$	-0.0568	0.1342	0.0469
	(0.0412)	(0.0933)	(0.1077)
Controls	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	464	428	464
Panel B	$Count_{ct}^{AllY02E}$	$Count_{ct}^{Nuclear}$	$Count_{ct}^{Hydrogen}$
$\text{PMR}_{elec,c,t-1,}$	-0.2045	0.3141	0.1999
	(0.1484)	(0.4089)	(0.4915)
Control Function	0.1782	-0.2223	-0.2223
	(0.1532)	(0.4220)	(0.4220)
Controls	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
F-stat first stage	14.70	10.38	14.70
First Stage Bootstrap p-value	0.0024	0.0092	0.0024
Observations	464	428	464

Notes: Notes: All specifications also control for the use of policies to promote clean energy innovation, GDP per capita, the price of crude oil imports, and the country's knowledge stocks in dirty and clean energy technologies. In Panel B regulation in the telecommunications sector is used as instrument for regulation in the electricity sector using a CF approach. Standard errors are clustered at the country level and reported in parentheses (18 clusters). In Panel B, if the control function is significant in the first-stage regression, I adjust the standard errors based on 999 bootstrap replications and report them in italics and in square brackets

Table A.6: Results with Patents in Hydrogen and Nuclear related technologies - model (4)

	(1)	(2)	(3)	(4)
	All Y02E	All Y02E	All Y02E	All Y02E
Panel A: OLS	Rad	Rad	Ori	Ori
Lag PMR$_{elec}$	-0.0172	-0.0170	-0.0008	-0.0020
	(0.0004)	(0.0001)	(0.6568)	(0.4006)
	[-0.0280, -0.0104]	[-0.0226, -0.0142]	[-0.0083, 0.0023]	[-0.0064, 0.0047]
Controls	Yes	Yes	Yes	Yes
Applicant FE	Yes	No	Yes	No
Technology FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	49,669	55,747	49,628	55,705
Panel B: 2SLS	All Y02E	All Y02E	All Y02E	All Y02E
Lag PMR$_{elec}$	-0.0225	-0.0139	0.0014	0.0039
	(0.0258)	(0.0000)	(0.3613)	(0.4610)
	[-0.0317, -0.0081]	[-0.0180, -0.0075]	[0.7092]	[-0.0053, 0.0250]
Controls	Yes	Yes	Yes	Yes
Applicant FE	Yes	No	Yes	No
Technology FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
F-stat first stage	37.98	46.08	37.97	46.09
First Stage Bootstrap p-value	0.0028	0.0026	0.0028	0.0026
Observations	49,669	55,747	49,628	55,705

Notes: All specifications also control for the use of policies to promote clean energy innovation, GDP per capita, the price of crude oil imports, the number of applicant, the knowledge stock computed at the applicant level and its squared value. In Panel A, I report wild bootstrap p-values in parentheses and wild bootstrap 95% confidence intervals in square brackets, generated using boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters). In Panel B, regulation in telecommunication is used as instrument for regulation in the electricity sector and inference is based on a wild bootstrapped version of the Anderson-Rubin obtained using the boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters).

Table A.7: Results for Patents in Hydrogen related technologies - model (4), by technological field

	(1)	(2)	(3)	(4)
	$Hydrogen$	$Hydrogen$	$Hydeogen$	$Hydeogen$
Panel A	Rad	Rad	Ori	Ori
PMR$_{elec,c,t-1,}$	-0.0066	-0.0118	0.0061	-0.0001
	(0.2645)	(0.0028)	(0.4263)	(0.9644)
	[-0.0393, 0.0033]	[-0.0248, -0.0068]	[-0.0203, 0.0140]	[-0.0189, 0.0099]
Controls	Yes	Yes	Yes	Yes
Applicant FE	Yes	Yes	No	No
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	10,004	11,016	10,000	11,012
Panel B	Rad	Rad	Ori	Ori
Lag PMR$_{elec}$	-0.0023	-0.0082	0.0231	0.0178
	(0.8321)	(0.1280)	(0.0219)	(0.1010)
	[-0.0313, 0.0112]	[-0.0655, 0.00186]	[0.0043, 0.0428]	[-0.0035, 0.0603]
Controls	Yes	Yes	Yes	Yes
Applicant FE	Yes	No	Yes	No
Technology FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
F-stat first stage	30.18	24.96	30.16	24.96
First Stage Bootstrap p-value	0.0138	0.0109	0.0138	0.0109
Observations	10,004	11,016	10,004	11,016

Notes: All specifications also control for the use of policies to promote clean energy innovation, GDP per capita, the price of crude oil imports, the number of applicant, the knowledge stock computed at the applicant level and its squared value. In Panel A, I report wild bootstrap p-values in parentheses and wild bootstrap 95% confidence intervals in square brackets, generated using boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters). In Panel B, regulation in telecommunication is used as instrument for regulation in the electricity sector and inference is based on a wild bootstrapped version of the Anderson-Rubin obtained using the boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters).

A.9 Complete Result Tables for main results

Table A.8 shows all estimated coefficients omitted from Panel A in Table 1.2. The results when using the CF approach are similar to those in Table A.8 and are therefore not reported. Similarly, Table A.9 shows all estimated coefficients omitted from Panel A in Table 1.4. Again, the results in the 2SLS estimates are very similar to those presented here, so they are not reported. In Table A.8, EPS_{tech} is negatively correlated with the number of dirty patents and positively correlated with the number of clean patents, as expected. The estimated effect of GDP per capita is not significant in either regression, while the price of crude oil imports is negatively correlated with the number of dirty innovations. Finally, the country's knowledge stocks in clean energy patents is positively correlated with the number of clean energy patents, while the opposite is true for the knowledge stock in dirty energy patents. In Table A.9 we see that country-level control variables other than PMR_{elec} are not significantly correlated with the radicalness or the originality of clean energy technologies. In two specification out of four (columns 2 an 3) we see the inverted-U relationship between the applicant knowledge stock and the indicators of interested hypothesised in Section 1.4.2.

Table A.8: Complete result table for Table 1.2

	(1)	(2)
Panel A	$Count_{ct}^{Dirty}$	$Count_{ct}^{Clean}$
PMR$_{elec,c,t-1,}$	0.1469***	-0.0999***
	(0.0439)	(0.0336)
EPS$_{tech,c,t-1,}$	-0.0588***	0.08199**
	(0.0219)	(0.0336)
GDP$_{pc;c,t-1,}$	0.0006	-0.0029
	(0.0039)	(0.0050)
Oil Price Imports$_{c,t-1,}$	-1.55*	-0.43
	(0.82)	(0.84)
Knowledge Stock$_{;c,t-1,}^{clean}$	-0.0038	0.014**
	(0.013)	(0.0063)
Knowledge Stock$_{;c,t-1,}^{dirty}$	0.0045	-0.036***
	(0.0148)	(0.0097)
Controls	Yes	Yes
Country FE	Yes	Yes
Year FE	Yes	Yes
Observations	464	464

Notes: Full result table from Table 1.2, Panel A. Standard errors are clustered at the country level and reported in parentheses (18 clusters).

Table A.9: Complete result table for Table 1.4

	(1)	(2)	(3)	(4)
	Rad	Rad	Ori	Ori
Lag PMR_{elec}	-0.0171	-0.0181	-0.0021	-0.0018
	(0.0167)	(0.0005)	(0.2938)	(0.2118)
	[-0.0316, -0.0043]	[-0.0285, -0.0124]	[-0.0120, 0.0020]	[-0.0058, 0.0011]
Lag EPS_{tech}	-0.0014	-0.0004	-0.0017	-0.0010
	(0.6271)	(0.9041)	(0.7135)	(0.6685)
	[-0.0070, 0.0049]	[-0.0075, 0.0064]	[-0.0072, 0.0072]	[-0.0057, 0.0061]
Lag GDP_{pc}	0.0006	0.0005	0.0002	0.0002
	(0.4414)	(0.4420)	(0.7134)	(0.5691)
	[-0.0009, 0.0017]	[-0.0007, 0.0013]	[-0.0010, 0.0013]	[-0.0006, 0.0008]
Lag Oil price (imports)	0.1896	0.1742	0.0286	0.0233
	(0.1049)	(0.2556)	(0.7431)	(0.8595)
	[-0.0516, 0.3528]	[-0.0996, 0.4794]	[-0.259, 0.1617]	[-0.3173, 0.2724]
Number Applicants	-0.0036	-0.0026	0.00002	0.0019
	(0.6836)	(0.7420)	(0.9988)	(0.6077)
	[-0.0619, 0.0345]	[-0.0232, 0.0361]	[-0.1107, 0.0307	[-0.0052, 0.0169]
Lag of Knowledge Stock_{clean}	-0.0230	-0.2068	-0.0639	-0.1871
	(0.7590)	(0.0173)	(0.0953)	(0.1141)
	[-0.3225, 0.2619]	[-0.3062, -0.0647]	[-0.1899, 0.0557]	[-0.3374, 0.1115]
Lag of Knowledge Stock^2_{clean}	-0.0103	0.2795	0.0703	0.2399
	(0.8315)	(0.0522)	(0.0775)	(0.1165)
	[-0.1525, 0.2179]	[-0.0064, 0.5011]	[-0.0482, 0.2038]	[-0.3689, 0.5881]
Applicant FE	Yes	No	Yes	No
Technology FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	37,371	42,834	37,333	42,797

Notes: Full result table from Table 1.4, Panel A. I report wild bootstrap p-values in parentheses and wild bootstrap 95% confidence intervals in square brackets, generated using boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters).

A.10 Share of Clean vs Dirty Patents

The home bias effect is a well known issue that affects studies based on patents from a unique patent office. This issue arise because inventors in some of the countries in the sample are generally more likely to file patent applications in the office considered. For instance, in the case at hand we can imagine that inventors in European countries will be more likely to file patents at the EPO than inventors in Japan or Canada. Luckily this issue can be fully dealt with in the estimation process (see for instance Conti et al., 2018). In model 1.3, I dealt with the home bias effect by including country fixed effects and by controlling for the country-level patent stock in clean and dirty technologies. By doing this, the proposed specification controls for the fact that inventors in some countries will tend prioritize filing patents at the EPO. This being the case, the relatively high number of patent applications in the sample from European countries does not affect the estimates. An even more direct way to address this issue is by using the the share of green patents as dependent variable, as this measure clearly abstract even more from the total number of applications by a country at the EPO (Conti et al., 2018). As expected, the results using the share of clean patents as dependent variable fully confirm the ones from the main analysis (see table A.10).

Table A.10: Share of Clean vs Dirty patents

	(1)	(2)
Panel A	$Share_{ct}^{Clean}$	$Share_{ct}^{Clean}$
PMR$_{elec,c,t-1,}$	-0.0346**	-0.0829**
	(0.0138)	(0.0345)
Control Function		0.0595
		(0.0374)
Controls	Yes	Yes
Country FE	Yes	Yes
Year FE	Yes	Yes
F-stat first stage		14.70
First Stage Bootstrap p-value		0.001
Observations	462	462

Notes: OLS regression. All specifications also control for the use of policies to promote clean energy innovation, GDP per capita and the price of crude oil imports. In Panel B regulation in the telecommunications sector is used as instrument for regulation in the electricity sector using a CF approach. Standard errors are clustered at the country level and reported in parentheses (18 clusters). In the first stage regressions of Panel B the control function is never significant, hence standard errors are not further corrected using a bootstrap procedure (Wooldridge, 2015)

A.11 Higher Lags of Policy Variables

Table A.11 and A.12 show the result of the base estimates when I use the third lag of the policy variables (PMR_{elec}, the use of policies to promote clean energy innovation, GDP per capita and the price of crude oil imports). The main difference is the change in the coefficient in the specification where the radicalness index is the dependent variable and we include applicants fixed effects (column 1 in table A.12). This coefficient is now significant only at the 10% level. However, the other results and the overall conclusions of the paper are robust to this exercise.

Table A.11: Results Using the third lag of policy variables - Patent Level Analysis

	(1)	(2)
Panel A	$Count_{ct}^{Dirty}$	$Count_{ct}^{Clean}$
PMR$_{elec,c,t-3,}$	0.1303***	-0.1262757***
	(0.0316)	(0.0373)
Controls	Yes	Yes
Country FE	Yes	Yes
Year FE	Yes	Yes
Observations	421	421

Notes: All specifications also control for the use of policies to promote clean energy innovation, GDP per capita, the price of crude oil imports, and the country's knowledge stocks in dirty and clean energy technologies. For the first three of these variables, as well as for the PMR index, in these estimates I use the third lag. Standard errors are clustered at the country level and reported in parentheses (18 clusters).

Table A.12: Results Using the third lag of policy variables - Patent Level Analysis

	(1)	(2)	(3)	(4)
	Clean	Clean	Clean	Clean
Panel A: OLS	Rad	Rad	Ori	Ori
Third Lag PMR$_{elec}$	-0.0144	-0.0144	-0.003	-0.0016
	(0.0941)	(0.0097)	(0.1158)	(0.2672)
	[-0.0250 0.0044]	[-0.0246, -0.0064]	[-0.0060, 0.0017]	[-0.0064, 0.0010]
Controls	Yes	Yes	Yes	Yes
Applicant FE	Yes	No	Yes	No
Technology FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	36,849	42,180	36,811	41,142

Notes: All specifications also control for the use of policies to promote clean energy innovation, GDP per capita, the price of crude oil imports, the number of applicant, the knowledge stock computed at the applicant level and its squared value. For the first three of these variables, as well as for the PMR index, in these estimates I use third lag. I report wild bootstrap p-values in parentheses and wild bootstrap 95% confidence intervals in square brackets, generated using boottest command in Stata (Roodman et al., 2019) for standard errors clustered at the country level (18 clusters).

Chapter 2

Fighting for Feedstock: Food-Fuel competition in France

Abstract

To date, there have been few empirical studies examining the relationship between biofuel support policies and food security in European countries. In this paper, we study the impact of bioethanol support in France on wheat usage and prices. To this end, we leverage a sudden and substantial shock to French bioethanol production capacity in 2003 and apply the synthetic control method. On the one hand, we find that the shock led to a significant increase in the amount of wheat used outside the food sector, resulting in a significant decrease in the share of wheat used in the food sector. On the other hand, we find no evidence of higher wheat prices in the post-treatment period. If, as economic theory predicts, higher demand for wheat in the fuel sector had an upward effect on wheat prices, this effect was likely overshadowed by the high volatility of wheat prices worldwide due to the global food crisis in 2007/2008.

Keywords: Biofuels, Food security,Climate Change, Energy markets. **Jel Codes**: Q16, Q42, Q18, Q01

2.1 Introduction

The World Energy Outlook (2022) reports that 90% of today's liquid biofuels come from food crops such as corn, wheat and soybeans (i.e., "conventional biofuels"). This creates a "food-fuel competition" where biofuel production competes with food production for agricultural feedstock, impacting food security (Austin et al., 2022; Muscat et al., 2020; Popp et al., 2016; Koizumi, 2015; Searchinger et al., 2015). This competition occurs when increased demand for agricultural commodities triggered by biofuel support policies leads to higher food prices and to the diversion of crops from food to biofuel production.

The debate around conventional biofuels in Europe has intensified with the "Fit for 55" plan and the Russian invasion of Ukraine (USDA, 2022). [1] Historically, NGOs have opposed the production of conventional biofuels due to their impact on food security and land use (Pilgrim et al., 2010; IFEU, 2023), while advocates argue for their environmental benefits (E-pure, 2022; E3-modelling, 2021).

To address the potential problems of conventional biofuels, Europe is shifting towards advanced biofuels (e.g, biofuels from non-food crops or waste) [2], aiming for a 14% share of renewable energy in transport by 2030, with advanced biofuels accounting for at least 3.5% of transport energy (WOE, 2022; EU Renewable Energy Directive, 2018). However, the development of advanced biofuels requires long-term policies, as they are not yet mature and face technological challenges to compete with conventional biodiesel and bioethanol (Cadillo-Benalcazar et al, 2021, Millinger at al, 2017).

Our analysis sheds light on the European debate on conventional biofuels. The existing literature that studies the effect of biofuel support polices focuses on the projection of expected effects obtained through simulation approaches, general or partial equilibrium models and biophysical models (Lark et al., 2022, Khanna et al., 2021, Muscat et al., 2020, Gardebroek et al., 2017, Efroymson et al., 2016). Econometric studies can therefore be a valuable addition to this literature as they help us understand the realized impact of past biofuel support

[1] Examples of this debate can often also be found in popular newspapers, see for example: 1)https://www.lemonde.fr/en/economy/article/2022/07/24/biofuel-mania-takes-hold-of-france_5991250_19.html; 2)https://www.nytimes.com/2023/06/25/opinion/letters/biofuels-environment.html; 3)https://www.argusmedia.com/en/news/2332243-germany-suggests-phase-out-of-crop-biofuels-by-2030 4)https://www.theguardian.com/environment/2022/jul/13/halt-use-of-biofuels-to-ease-food-crisis-says-green-group

[2] See also: https://energy.ec.europa.eu/topics/renewable-energy/bioenergy/biofuels_en

policies on variables such as cropland or feedstock prices (Austin, 2022; Lark et al., 2022, Newes et al 2022). Several recent works attempt to address this gap in the literature, some examples are Boly and Sanou, 2022; Lark et al., 2022; Sant'Anna, 2021; Ifft et al., 2019; Li, et al., 2019 ; Gardebroek et al., 2017; Wright et al., 2017. However, econometric evidence for European countries is lacking (Gardebroek et al 2017). This paper fills this gap in the literature by examining the impact of higher bioethanol production capacity on cereal use and prices in a European country using a method well suited for estimating a causal relationship (Abadie 2021, Abadie et al 2010). In particular, we investigate how the support for bioethanol in France has affected wheat use and prices by examining the causal effect of the sudden increase in bioethanol production capacity in France in 2003 on these variables. To this end, we use the synthetic control method (SCM) to estimate what would have happened to the outcome variables of interest if this shock had not occurred (Abadie et al., 2010; Abadie and Gardeazabal, 2003).

We selected France and wheat as our study's focal points for several reasons. Firstly, France has emerged as the largest bioethanol producer in Europe since the 2003 shock to bioethanol production capacity, a position it maintains today (USDAb 2022, Sorda et al 2010). This makes it an ideal case study to investigate the impact of bioethanol production support on cereal usage and prices in Europe.

Moreover, France is the largest wheat producer among EU Member States and the European Union is the world's top wheat producer (Mohanty and Swaine et al. 2019). Consequently, factors that affect wheat supply in France can have significant implications for food security at both the European and global levels.

Finally, according to FAOSTAT data, wheat dominates cereal production in France with 36,559.45 thousand tonnes produced in 2021. In the decade before the shock in bioethanol production capacity (1992-2002), wheat accounted for 74% of all cereals used outside the food sector use in France, 46% of all cereal consumption for animal feed and 84% of cereals consumed directly by humans. [3] These figures suggests that if the shock to bioethanol production capacity had an effect on the use of agricultural feedstock in France, we expect this effect to be particularly visible for wheat. Note also that focusing on one cereal allows us to establish a more direct link between a change in use and the expected change in price, as we do not need to rely on a price index built as an average of prices for different cereals. In Appendix B.3 we show that while, as expected, the effect of the shock is particularly visible when we focus on wheat, the main conclusions we draw from the analysis remain valid when

[3] Source: https://www.fao.org/faostat/en/#home

we examine the use of cereals as a broader category.

On the one hand, our findings show a significant shift in wheat usage in France after 2003, driven by the increase in bioethanol production capacity. By 2012, 28kg per capita more wheat was used outside the food sector, causing a 11.5% drop in the share of domestic wheat supply used in the food sector food. On the other hand, we see no notable impact on wheat prices compared to the counterfactual scenario. These findings are consistent with Gardebroek et al. (2017), who study the effect of biodiesel support on rapeseed prices in Germany and France from 2000 to 2015. According to economic theory, all else being equal, one would expect the additional demand for wheat in the fuel sector to put upward pressure on wheat prices (Gardebroek et al. 2017, Baffes and Haniotis, 2016; Koizumi, 2015, De Gorter et al., 2015 Busse et al., 2012). However, in Section 2.6 we propose that the effect of the shock on wheat prices may have been relatively small compared to the post-treatment volatility of wheat prices, making it difficult to discern using the SCM (Abadie, 2021). In particular, if there was an effect on wheat prices, it was likely overshadowed by the high volatility of cereal prices worldwide as a result of the 2007/2008 global food crisis (Tadasse 2016, Mittal 2009). Refer to Section 6 for a more comprehensive discussion of this matter.

From these results, we conclude that support for bioethanol in France was not a significant driver of wheat prices during the time span of the analysis. Any impact that the 2003 shock to bioethanol production capacity may have had on wheat prices was relatively small compared to other factors that influenced wheat prices over the same period. However, the substantial change in wheat usage in the post-treatment period suggests that the expansion of conventional biofuels could eventually exert significant upward pressure on agricultural commodity prices if future biofuel policies do not prioritise food security. In the following section, Section 2.2, the paper discusses the background and the relevant literature. Section 2.3 describes the data and the empirical strategy. Section 2.4 presents the main results, while Section 2.5 assesses their robustness. Finally, Section 2.6 discusses the results and concludes.

2.2 Background

Conventional biofuels compete directly with food production for agricultural feestock, land use and solar radiation (Amigues and Moreaux, 2019). Hence, economic theory predicts a positive effect of biofuel support policies on crop prices and land use (Gardebroek et al. 2017, Baffes and Haniotis, 2016; Koizumi, 2015, De Gorter et al., 2015 Busse et al., 2012). These policies are expected to lead to an increase in the demand for agricultural feedstock, which will put upward pressure on the prices of agricultural products, thus providing an incentive for producers to increase agricultural production (Austin et al., 2022, Popp et al., 2016, Searchinger et al., 2015). The new equilibrium of the market, and thus the answer to the question of how much agricultural prices and supply will ultimately change, depends on the elasticity of supply with respect to prices (Koizumi, 2015).

The literature studying the effects of biofuel support policies on agricultural prices, land use and agricultural production mainly relies on simulation approaches, general or partial equilibrium models and biophysical models (Lark et al. 2022, Khanna et al. 2021, Muscat et al. 2020, Gardebroek et al. 2017, Efroymson et al., 2016). Complementing this literature, there is a growing body of work that focuses on the realised effects of previous biofuel policies implemented in the early 2000s. These studies attempt to isolate a causal relationship between an observed effect and a treatment, controlling for confounding factors (Austin 2022). Section 2.2.1 discusses this literature, focusing on empirical studies that address the relationship between biofuel production, land-use and food security. Section 2.2.2 describes in more detail the European and French context during the period of analysis.

2.2.1 Empirical Studies on Biofuel, Land Use, and Food Security

Most of the recent literature that empirically examines the realized impact of biofuel support policies is focused on the Renewable Fuel Standard (RFS) in the US (some examples are Lark et al. 2022; Ifft et al 2019; Li et al 2019; Wright et al. 2017; Carter et al., 2016). This policy was introduced in 2005 and significantly tightened in 2007 to require the blending of an increasing share of biofuels with transportation fuels.

Lark et al. (2022) examine the effects of the RFS through a combination of econometric analysis and biophysical modelling and conclude that it increased US corn prices by 30% and prices of other crops by 20%. As a result, the area used to grow corn in the US increased by 8.7% and the total cropland area increased by 2.4% between 2008 and 2016, leading to a substantial increase in annual nationwide fertiliser use and significant land-use changes. Because of these changes, they estimate that the carbon intensity of corn bioethanol produced under the RFS is no less than that of gasoline and likely at least 24% higher. A different conclusion is reached by Newes et al. (2022), who suggest that the role usually attributed to RFS in supporting US biofuel production may be overstated and bioethanol competitiveness alone can explain much of the increase in US production between 2002 and 2019.

Wright et al. (2017) use the implementation of the RFS to study the relationship between land-use and biofuels in more detail. They find that nearly 4.2 million acres of arable non-cropland have been converted to crops within 100 miles of bioethanol refineries, with the rate of conversion of grassland to cropland being positively correlated with proximity to a refinery. On the other hand, Li, Miao, and Khanna (2019) examine the determinants of changes in corn acreage and total acreage in the United States between 2003 and 2014, and their results suggest that both corn acreage and total acreage are relatively inelastic to changes in bioethanol capacity at nearby locations and crop prices. Similarly, Ifft et al. (2019) use data from 10 Midwestern states (Illinois, Indiana, Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota, and Wisconsin) between 1999 and 2014 and find that proximity to an bioethanol plant is not a significant predictor of land exiting from the U.S. Federal Conservation Reserve programme. Their results contrast with Secchi et al. 2011, who estimated that a 1% increase in corn price leads to a 0.7% to 1.5% decrease in land enrolment in the U.S. Federal Conservation Reserve Programme.

The meta-analysis by Austin et al. (2022) summarises the various papers that attempt to estimate the impact of the RFS on land-use. According to this review, the empirical estimates of the impact of biofuel production on cropland in the US ranges from 0.38 to 0.66 million acres per billion-gallon of increase in biofuel.

Empirical analyses of the effects of biofuel support policies outside the US are less common. Sant'Anna (2021) examines the effect of biofuel production on land use and deforestation in Brazil from 2004 to 2013 and finds that 92% of new bioethanol came from expanding cropland and only 8% from increased yields.

Boly and Sanou (2022) use the synthetic control method to estimate the impact of biofuel production on food security in Mexico and Indonesia over the period 2000-2013. They find that biofuel production had a negative impact on food security in Mexico, while this was not the case in Indonesia. The authors suggest that this difference might be explained by the different feedstocks used to produce biofuels in the two countries. Indonesia primarily uses palm oil and jatropha, which are not directly used for food, while Mexico relies mainly on maize. Their findings confirm the importance of taking a country-specific approach when studying the effect of biofuel support policies, as country-level heterogeneity can result in similar policies having different different outcomes in different countries (Muscat et al., 2020). In light of this, in Section 2.6 we discuss some trends specific to France, particularly in relation to the production of livestock products, which may have played a role in determining the outcomes we observe. Looking at the European Union, Gardebroek et al. (2017) show that, contrary to what economic theory suggests, the increase in biodiesel production in France and Germany in the period 2000-2015 was not correlated with higher rapeseed prices and acreage. In the same paper, the authors also point to the need for more econometric studies that empirically examine the impact of past biofuel policies in the EU, a need that to the best of our knowledge has not been met yet.

2.2.2 The European and French legislative background

When discussing the EU approach to promoting biofuels, it is important to distinguish between policies implemented at EU level and those implemented at country level. The former are meant to establish a general legislative framework and set general targets (Gardebroek et al 2017). These targets are usually expressed as a percentage of biofuels to be blended with conventional road transport fuels (Sorda et al 2010), i.e. they refer to biofuel consumption rather than production. Within the framework set by the EU, individual countries have the autonomy to decide how to achieve the targets and this has led to significant differences in the extent to which different EU Member States support biofuel production (Gardebroek et al., 2017; Sorda et al., 2010; Guide and Jacquet, 2008). In this section, we focus on the policies implemented by the European Union and France in the early 2000s; the period of interest for this study.

In 2003, the EU implemented a new reform of the Common Agricultural Policy (hereafter CAP) and two directives aimed at promoting the production and consumption of biofuels. The reform of the CAP led to the introduction of the Energy Crop Scheme, a subsidy of €45 per hectare devoted to the cultivation of energy crops (Cadillo-Benalcazar et al., 2021; Council Regulation (EC) No 1782/2003). The first of the two directives, the so-called Biofuels Directive (2003/30/ EC), set a target of 2% biofuel blending with road transport fuels to be achieved by 2005, followed by a higher target of 5.75% for 2010 (Cadillo-Benalcazar et al., 2021; Gardebroek et al., 2017; Berti and Levidov, 2010; Sorda et al., 2010; Directive 2003/30/ EC). However, it is worth noting that these targets were not binding and are generally considered as an ineffective measure (Cadillo-Benalcazar et al., 2021, Cansino et al. 2012; Sorda et al., 2010; Kutas et al., 2007). The second directive was the Energy Tax Directive (2003/96/ EC), which allowed individual Member States to introduce tax reductions or exemptions for biofuels in order to meet the targets set in the 2003 Biofuels Directive (Cadillo-Benalcazar et al., 2021, Cansino et al. 2012, Sorda et al., 2010; Directive 2003/96/ EC).

Taking advantage of this framework, France and Germany took the strongest stance in the support of biofuels production and became the two leading biofuel producers in the European Union (Sorda et al., 2010, Kutas et al., 2007). In particular, France quickly became, and still is, the first country in the European Union for bioethanol production. By 2010, French bioethanol production capacity was about 1092 thousand tonnes per year; 37% higher than that of Germany (797 tonnes per year) and 235% higher than that of Spain (464 tonnes per year), respectively the second and third countries for bioethanol production according to EUROSTAT data. [4]

In 2005, France moved the European target of 5.75% set by the Biofuels Directive from 2010 to 2008. At the same time, a more ambitious target of 7% was set for 2010 (Gardebroek et al. 2017, Guide and Jacquet 2008, Kutas et al., 2007). To achieve this more ambitious target, the French government also introduced two instruments that provided strong economic incentives for biofuel production and greatly reduced business risks for biofuel producers (Gardebroek et al. 2017, Wiesenthal et al., 2009, Guide and Jacquet 2008, Kutas et al., 2007).

First, an amendment to the Taxe Générale sur les Activités Polluantes (TGAP, General Tax on Polluting Activities) introduced an incentive tax on biofuel blending. Under this policy, fuel traders had to pay a surcharge in addition to the excise tax (Gardebroek et al 2017, Guide and Jacquet 2008). This surcharge was relatively expensive, but could be avoided by blending biofuels with conventional fuels (Gardebroek et al 2017, Kutas et al, 2007). Second, the French government introduced a system of production quotas. These quotas were allocated through public tenders and biofuels produced under this system benefited from tax reductions, which were gradually reduced until 2010 to meet targets set at the national level (Sorda et al., 2010, Kutas et al., 2007). In 2005, the reduction for bioethanol was €0.38/l, while in 2009 it was €0.15/l (Sorda et al. 2010). Given these measures, while the targets set by the French government were not binding, they are generally considered as akin to a biofuel mandate (Gardenbroek et al., 2017; Kutas et al., 2007; Wiesenthal et al., 2009).). These measures led to a sharp increase in French bioethanol production capacity in the early 2000s, see Figure 2.1. Charles et al (2013) also estimates a strong increase in bioethanol investment and production capacity in France after 2005.

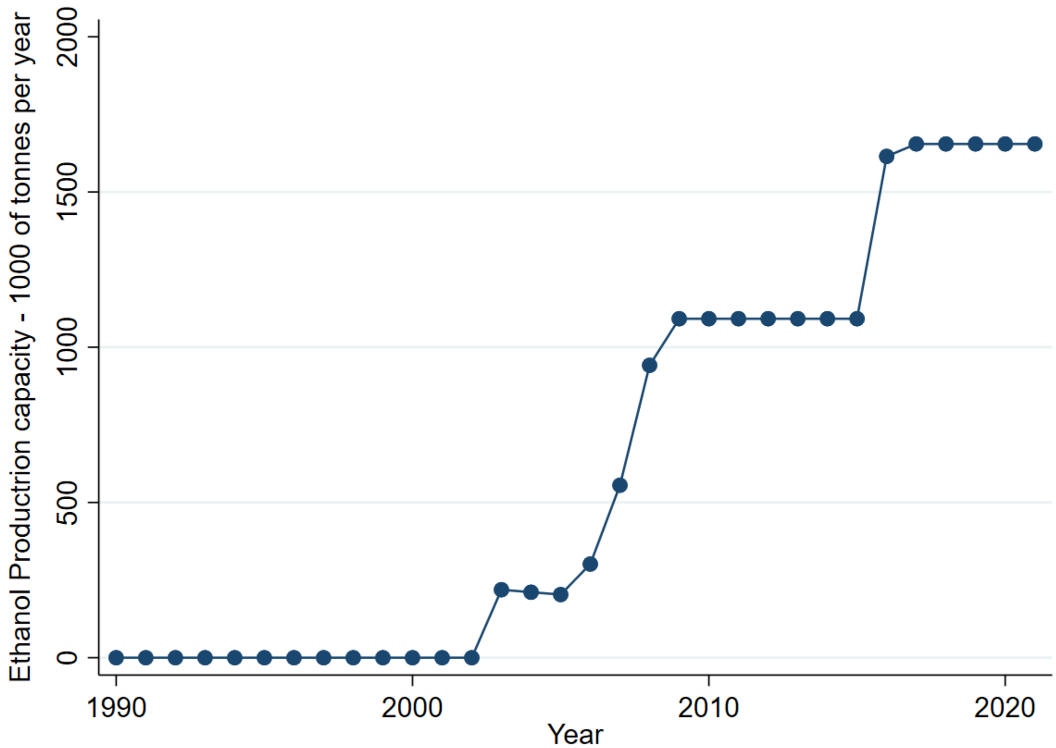
2.3 Empirical Strategy

2.3.1 The Synthetic Control Method

We use the synthetic control method (SCM) developed in Abadie and Gardeazabal (2003) and Abadie et al. (2010, 2015) to estimate the causal impact of the 2003 shock to biofuel production capacity in France on wheat use and price. The SCM creates a counterfactual

[4] https://ec.europa.eu/eurostat/cache/metadata/en/nrg_inf_lbpc_esms.htm

Figure 2.1: Bioethanol Production Capacity - France



Notes:Bioethanol production capacity in France, measured in 1000 of tonnes per year. Source: EUROSTAT, https://ec.europa.eu/eurostat/cache/metadata/en/nrg_inf_lbpc_esms.htm

for the treated unit (the so-called ”synthetic control”) using a weighted average of untreated units or units whose treatment exposure differs significantly from the treated unit (Abadie et al., 2010). This group is often referred to as the ”donor pool”. This method is particularly attractive when we observe only a few aggregated units, as an optimally chosen combination of units from the donor pool can often reproduce the characteristics of the treated unit much better than any of the untreated units on their own (Abadie et al 2021, Cunningham, 2021). The formalisation of the SCM briefly discussed in the remainder of this section follows Abadie et al (2021) and Cunningham (2021).

Let $J + 1$ be the number of countries we observe over a time span of T periods, where T_0 indicates the year of the intervention. Without loss of generality, we assume that the index for the treated unit is $j = 1$. The countries in the donor pool are therefore indexed by $j = 2, ..., J + 1$. In this particular application, the treated country is France, the treatment is the shock to bioethanol production capacity, T_0 is equal to 2003, the time span of the analysis starts in 1970 and ends in 2012 and the countries in the donor pool are those discussed in

Section 2.3.6. For each country, we observe the outcome of interest Y_{jt} and a set of other variables that help predict this outcome (X_{1j},...,X_{kj}). These variables, called predictors or matching variables, are discussed in detail in Section 2.3.5. Following standard notation, we can define X_{1j},...,X_{kj} as the k x 1 vectors of predictors for units $j = 1, ..., J + 1$ and X_0 = [X_2,...,X_{J+1}] as the kxj matrix that collects the predictors for the units in the donor pool. Since we observe Y_{1t}, the problem of estimating the effect of the treatment boils down to estimating what Y_{1t} would have been without the shock; following standard notation, we call this counterfactual Y_{1t}^N. The SCM estimates Y_{1t}^N using a weighted average of the units in the donor pool, as described in equation (1).

$$\hat{Y}_{it}^N = \Sigma_{j=2}^{J+1} w_j Y_{jt} \tag{2.1}$$

Where w_j = w_2 , ... , w_{J+1} is a $J1$ vector of weights. We then use Y_{1t}^N to estimate the effect of the treatment on the treated unit (Π_{it}) as Π_{it} = Y_{1t} - Y_{1t}^N. To select the optimal set of weights $W*j$, Abadie and Gardeazabal (2003) and Abadie, Diamond and Hainmueller (2010) propose to choose w_2 , ... , w_{J+1} such that the synthetic control best resembles the pre-intervention values of Y_{1t} and X_{11},...,X_{k1}. In other words: $W * j$ is chosen to be the set of weights that minimizes equation 2.2.

$$\|X_1 - X_0 W\| = (\Sigma_{h=1}^{k} v_h (X_{h1} - w_2 X_{h2} - ..... - w_j X_{hj})^2)^{1/2} \tag{2.2}$$

Subject to the constraints that the weights are non-negative and sum to one. The positive constant v_h reflects the importance of each of the different matching variables. As a general rule, $v_1, \ldots, v_k$ should reflect the predictive values of covariates and, in practice, the set of $v_1, \ldots, v_k$ that minimizes the mean squared error is often the choice (Abadie 2021, Cunningham 2021). Cunningham (2021) defines the SCM as a picture-intensive estimator because both the results of the analysis and its robustness are often presented using images and graphs. In particular, the results of the analysis are presented in the form of a graph in which we compare the evolution of the outcome of interest for the treated unit and the synthetic control. If the shock had a significant effect on the outcome variable, we expect to see two lines that closely follow each other in the pre-treatment period and diverge after the treatment.

2.3.2 Inference and Robustness Checks in the SCM

The use of the SCM in empirical research is usually accompanied by a series of tests that provide inference for the results and assess their robustness. The remainder of this section briefly introduces these tests. See Abadie (2021) for a more comprehensive discussion about these tests and their interpretation.

2.3.2.1 In-Space Placebo Test (Permutation test)

The in-space placebo test is performed by iteratively applying the synthetic control method to each unit in the donor pool. If the treatment is effective, we should expect the estimated effect for the treated unit to be extreme compared to what we estimate for other countries in the donor pool (Cunningham 2021, Abadie et al 2010). To carry out this comparison Abadie, Diamond, and Hainmueller (2010) propose to compare for each unit the ratio between the post- and pre-intervention fit, i.e. post- and pre-intervention root mean squared error. If this ratio is particularly high for the treated unit compared to the units in the donor pool, this is evidence that the estimated effect for the treated unit is indeed extreme (Abadie, 2021; Cunningham, 2021).

2.3.2.2 In-Time Placebo (Backdating)

The idea of the in-time placebo test is to artificially anticipate the treatment by backdating it. The aim of this test is twofold: 1) we want to assess whether the synthetic control estimator is able to closely track the trajectory of the outcome variable for the treated unit in the years between the fake and the actual intervention and 2) we want to see if the estimates still show an effect around the year of the actual intervention. The emergence of these patterns lends credibility to the results (Abadie 2021, Abadie, Diamond and Hainmueller 2015).

2.3.2.3 Leave-one-out Test

The leave-one-out test iteratively removes units in the donor pool that received a positive weight and that were therefore used to build the synthetic control, re-running the estimation each time (Abadie et al. 2010). This allows us to test whether the estimated effect depends on the inclusion of a particular unit in the donor pool (Abadie, 2021).

2.3.4 Data

In the analysis that follows we focus on three outcome variables: 1) the amount of wheat used for non-food purposes, measured in kilogrammes per person, 2) the share of domestic wheat supply used in the food sector and 3) wheat prices, measured in US Dollars per tonnes.

Data on the production, domestic supply, use and price of wheat by country and year come from FAOSTAT database.[5] Domestic supply is defined as wheat production minus exports plus imports and stocks. It therefore measures the amount of an agricultural commodity that remains in the country after adjusting for trade and stock fluctuations. The possible uses of the domestic supply of each agricultural commodity are categorized into three categories: human consumption, animal feed and non-food-related uses. The per capita amount of wheat allocated to the last of these categories is our first outcome variable of interest. The second outcome variable of interest, i.e. the share of domestic wheat supply used in the food sector, is calculated as the sum of wheat used for human consumption and animal feed divided by the total domestic wheat supply in a given country and year. Table 2.1 presents descriptive statistics for each of these variables. Wheat domestic supply and its uses are observed for the whole period under study, i.e. from 1980 to 2012 [6] while for wheat prices we have data only from 1993 onwards.[7]

2.3.5 Predictors

Following common practice in SCM applications, in all our specifications we include as predictors pre-treatment lags of the outcome variable, as this has been shown to help control for unobserved confounding factors (Abadie 2021, Chunningham 2021, Ferman, Pinto and Possebom 2020, Abadie, Diamond and Hainmueller 2010). The specific pre-treatment lags included are shown in Tables 2.3, 2.5 and 2.7. When the outcome variable is the amount of wheat used outside the food sector, we also use the share of wheat used in the food sector as matching variable, and vice versa. We do this under the assumption that how wheat is used before the shock can affect how wheat use will change in response to the increase in bioethanol production capacity. Following the smae rationale, when the outcome variable is the price of wheat we again include as predictor the share of wheat used in the food sector. It is important to acknowledge that our current model specifications are just one approach among several possible alternatives. Recent developments in the literature have drawn attention to the issue of specification searching in empirical applications employing the SCM. This problem can arise when the researcher has some control over the results through the selection of the predictors

[5]https://www.fao.org/faostat/en/#home

[6]The analysis stops in 2012 because in 2013 the FAO changed the methodology used to compute Food Balances. See https://www.fao.org/faostat/en/#data/FBSH/metadata

[7]FAOSTAT data for wheat usage start in 1960. We decided to restrict our sample and start from 1980 as 23 years of pre-treatment period seems an appropriate number. Indeed, adding also the years between 1960 and 1980 does not change the results of the analysis.

used for matching (Ferman et al., 2020). The minimisation algorithm used by the SCM makes specification searching more difficult, but the problem remains (Ferman et al., 2020; Cunningham 2021). As suggested by Ferman et al. (2020) we address this issue in Appendix B.1 by showing that our results are robust across a wide number of different specifications.

Table 2.1: Descriptive Statistics

Variable	Observations	Mean	Std. Deviation	Min	Max
Wheat, non-food uses, Kg per capita	330	8.77	12.28	0	46.40
Share of wheat used in the food sector	330	0.957	0.056	0.793	1
Wheat producer prices, USD per tonnes	168	193.91	73	91.7	440.7

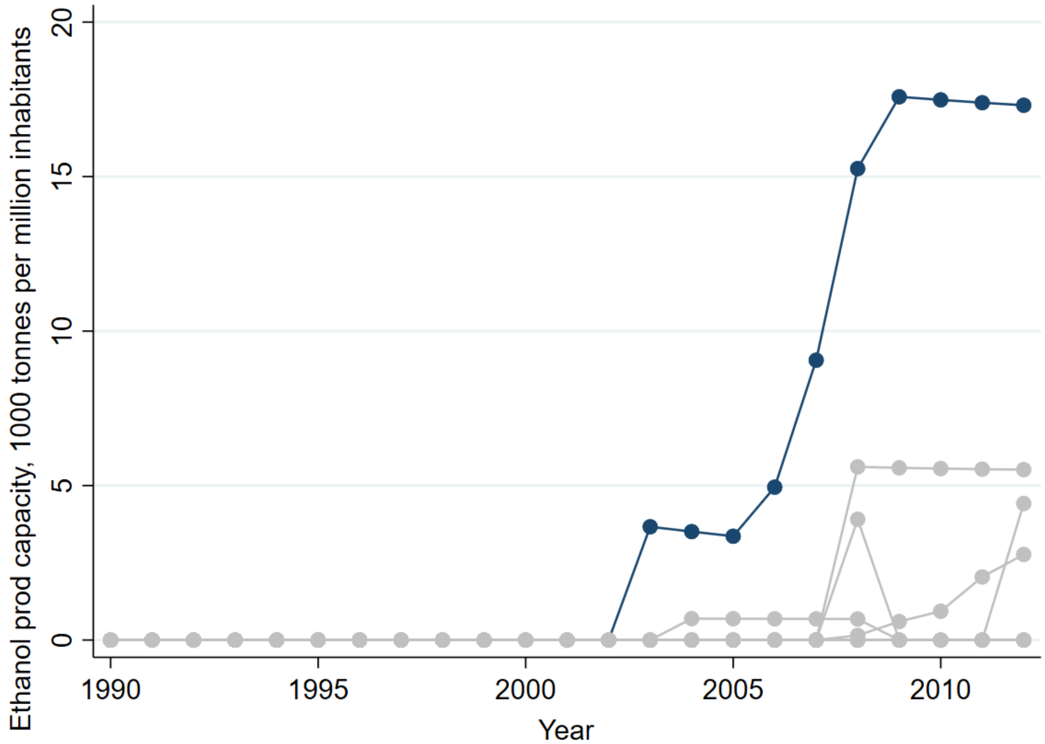
Notes: Descriptive statistics for the main variables in the analysis

2.3.6 Countries in the Donor Pool

Countries in the Donor Pool are selected starting from those in the European Union. However, as is common in the literature, we exclude from the donor pool countries that have experienced a shock in bioethanol production capacity that, although smaller compared to what we observe in France, can still be considered significant (either in absolute terms or relative to their size), see Abadie 2021 and Abadie et al 2010. The reason for doing this is that a comparative analysis can only be carried out when the treatment only applies to a subset of the units under study or if the exposure to such a treatment differs significantly between the units in the analysis (Abadie et al., 2010). In particular, we romove from the donor pool Austria, Hungary, Germany, Poland, Spain and Sweden. However, in Appendix B.2 we show that our robust to the use of a larger donor pool, which includes all these countries. The remaining EU countries for which we have complete data over the period of interest are those included in the donor pool. For our first two outcome variables these countries are: Bulgaria, Denmark, Finland, Greece, Ireland, Italy, Netherland, Portugal and Romania. Since we do not have data about wheat prices for Bulgaria and Romania we can not include them in the donor pool when the outcome variable is the price of wheat. Figure 2.2 compares the shock to bioethanol production capacity per million inhabitants in France and in countries in the donor pool, showing that this shock was substantially higher for our treated unit. [8]

[8] Since France is the country with the highest population among all of these, the difference in the shock is even bigger when we do not measure it in per capita terms.

Figure 2.2: Bioethanol Production Capacity - France and countries in the Donor Pool



Notes: Biothanol production capacity is measured in thousands of tonnes per year per million inhabitants. The blue line represents France. Other countries included in the graph are Bulgaria, Denmark, Finland, Greece, Ireland, Italy, Netherlands, Portugal and Romania. Source: EUROSTAT, https://ec.europa.eu/eurostat/cache/metadata/en/nrg_inf_lbpc_esms.htm

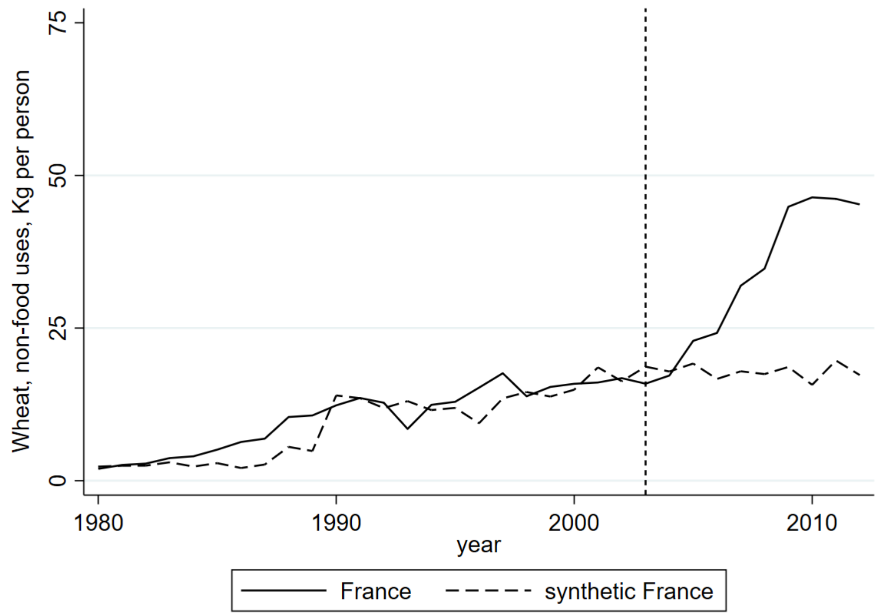
2.4 Empirical Analysis

2.4.1 Wheat used for non-food-related purposes and share of wheat domestic supply going to the food sector

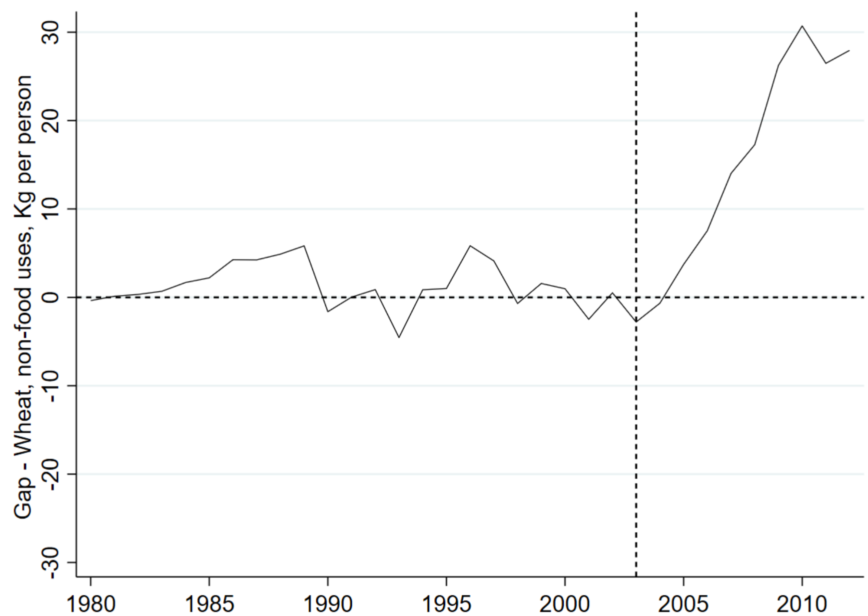
Figure 2.3 shows the results of our analysis when the outcome variable is the amount of wheat used outside the food sector, measured in kilograms per capita. Table 2.2 and 2.3 show the countries used to generate synthetic France and compare the pre-treatment value of the matching variables for France, synthetic France and the simple average of the countries in the donor pool. Synthetic France follows France well throughout the pre-treatment period and approximates the pre-treatment values of the matching variables much better than the average of the other countries in the donor pool. After 2003, the amount of wheat used for non-food purposes in France starts to increase significantly, while this is not the case in the counterfactual scenario. Note that the outcome variable in France follows a very similar pattern to bioethanol production capacity in Figure 2.1: it increases from 2003 to 2009 and then stabilises. This result is evidence that the shock to bioethanol production capacity led to more wheat being used outside the food sector. To be more precise, an additional 28 kg of wheat per capita was used outside the food sector in 2012 compared to the counterfactual scenario. To better understand the magnitude of this effect in relation to France wheat domestic supply, we examine the effect of the shock on the share of wheat used in the food sector. The results of this exercise are shown in Figure 2.4, while tables 2.4 and 2.5 show how synthetic France is built and compare the pre-treatment value of the matching variables for France, Synthetic France and a simple average of the units in the donor pool. Again, France and synthetic France move together until 2003 and then start to diverge as the share of wheat used in the food declines in France but remains stable in the counterfactual scenario. In 2012, due to the shock in bioethanol production, the share of domestic wheat supply used in the food sector was 11.5 percentage points lower that it would have been otherwise. In the next section we investigate whether or not these changes had an effect on wheat domestic supply or prices.

Figure 2.3

Panel A: Wheat used outside the food sector



Panel B: Difference between France and Synthetic France in Panel A



Notes: The outcome variable is the amount of wheat used outside the food sector, measured in Kg per capita. The vertical line is set in 2003. Panel A compares the evolution of the outcome variable in France and in Synthetic France. Panel B plots the gap between the two lines in Panel A.

Table 2.2

Country	Weight	Country	Weight
Bulgaria	0	Italy	0
Denmark	0	Netherlands	0.274
Finland	0	Portugal	0.505
Greece	0.02	Romania	0.201
Ireland	0		

Notes: Country weights in synthetic France for the results presented in Figure 2.3

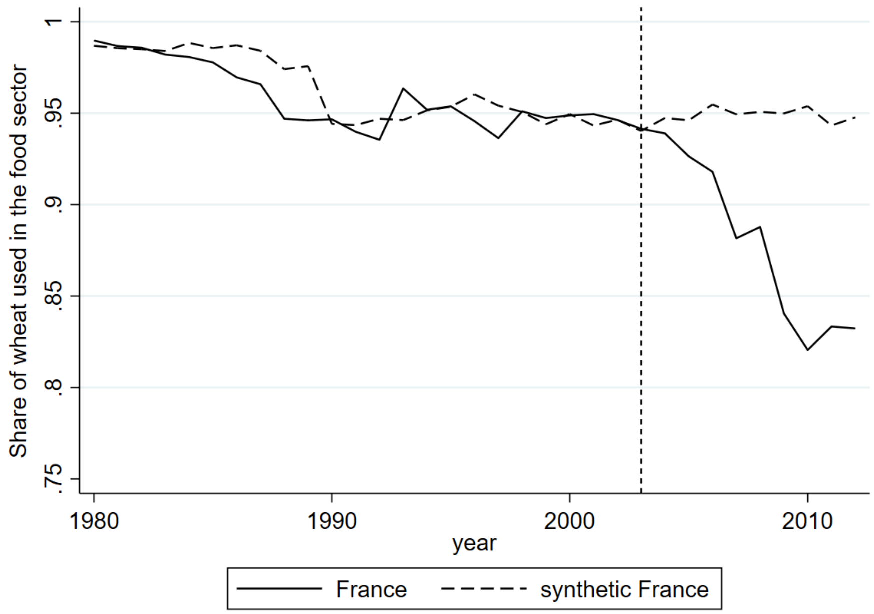
Table 2.3

Variable	France	Synthetic France	Avg. Donor Pool
Wheat, non-food uses, Kg per person (2002)	16.79	16.28	10.71
Wheat, non-food uses, Kg per person (1995)	12.92	11.92	8.10
Wheat, non-food uses, Kg per person (1990)	12.34	13.96	8.11
Wheat, non-food uses, Kg per person (1980)	1.94	2.30	3.63
Share of wheat used in the food sector	0.975	0.969	0.959

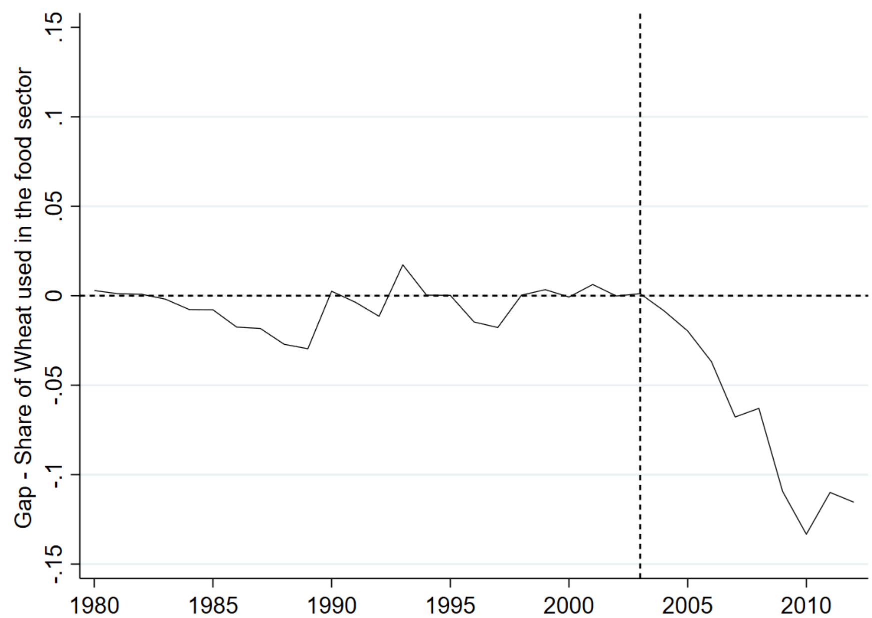
Notes: Predictor values for France, Synthetic France and the simple average of all countries in the donor pool. This specification is the one used to obtain the results presented in Figure 2.3

Figure 2.4

Panel A – Share of wheat used in the food sector



Panel B – Difference between France and Synthetic France in Panel A



Notes: The outcome variable is the share of wheat used in the food sector. The vertical line is set in 2003. Panel A compares the evolution of the outcome variable in France and in Synthetic France. Panel B plots the gap between the two lines in Panel A.

Table 2.4

Country	Weight	Country	Weight
Bulgaria	0	Italy	0
Denmark	0.577	Netherlands	0.109
Finland	0	Portugal	0
Greece	0.145	Romania	0.169
Ireland	0		

Notes: Country weights in synthetic France for the results presented in Figure 2.4

Table 2.5

Variable	France	Synthetic France	Avg. Donor Pool
Share of wheat used in the food sector (2002)	0.946	0.946	0.947
Share of wheat used in the food sector (1995)	0.954	0.953	0.953
Share of wheat used in the food sector (1990)	0.946	0.944	0.956
Share of wheat used in the food sector (1980)	0.989	0.987	0.974
Wheat, non-food uses, Kg per person	5.96	4.81	6.50

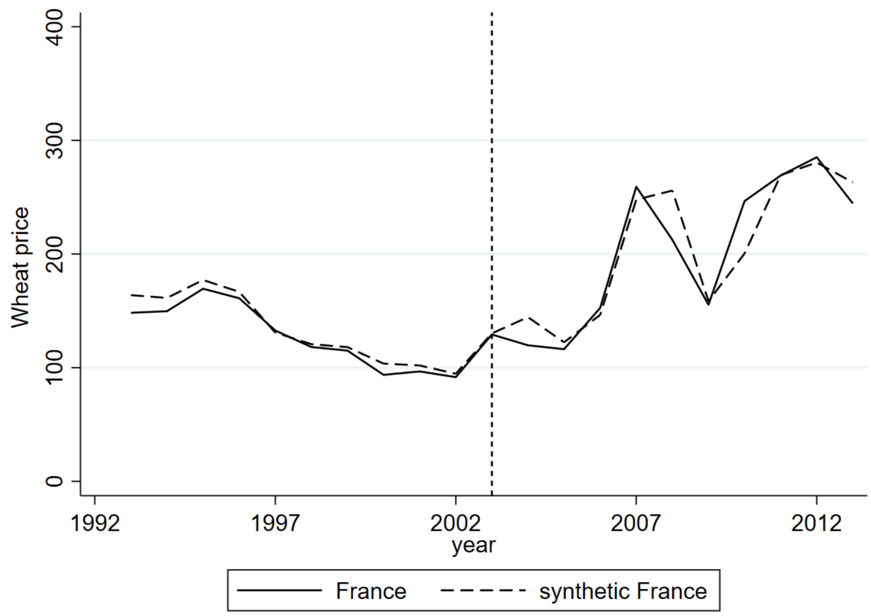
Notes: Predictor values for France, Synthetic France and the simple average of all countries in the donor pool. This specification is the one used to obtain the results presented in Figure 2.4.

2.4.2 Wheat Prices

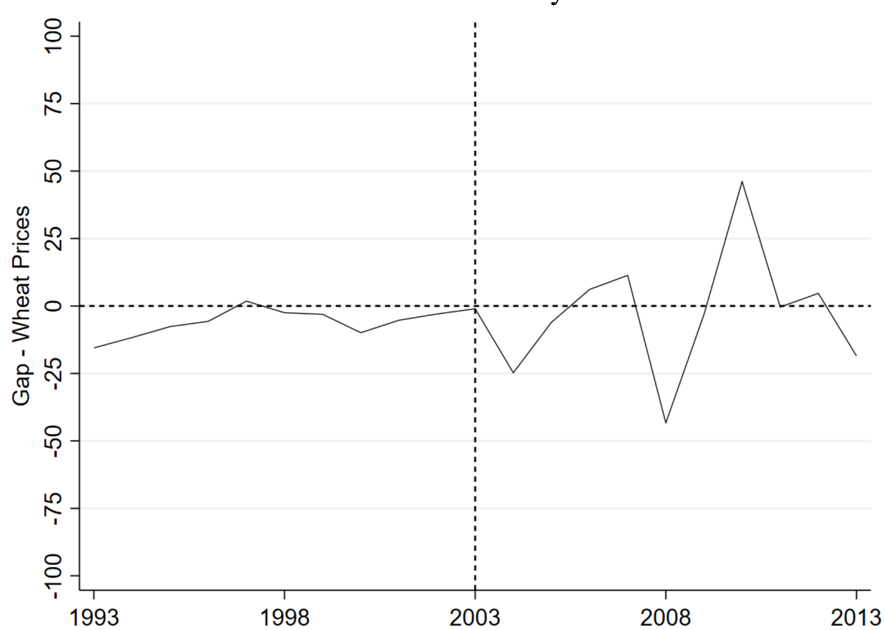
As discussed in Section 2.2, economic theory suggests that, other things being equal, an expansion of biofuel production capacity will increase demand for agricultural feedstock. This in turn is expected to trigger a competition between food and fuel that should put upward pressure on agricultural prices and thus supply. The results presented in the previous section substantiate the initial aspect of this narrative, demonstrating that the increase in bioethanol production capacity led to higher demand for wheat in non-food sector. We now proceed to investigate the latter aspect, which pertains to the influence of this shock on wheat prices. The results of the analysis are shown in Figure 2.5, whereas Tables 2.6 and 2.7 show how synthetic France is built and compare the values of the matching variables for France, synthetic France and the simple average of units in the donor pool. As we can see, there is no significant difference in wheat price between France and synthetic France, neither before nor after the shock. This suggests that the shock to bioethanol production capacity observed in France in 2003 was not a significant driver of wheat prices in the country in the post-treatment period. However, for a proper interpretation of this result it is important to note that, during the post-treatment period, wheat price increases and becomes more volatile in both France and synthetic France. This is due to the global food crisis that hit the world in those same years, leading to a dramatic increase in both the level and the volatility of international food prices (Tadasse 2016, Mittal 2009). This is important to consider, as the synthetic control estimator encounters challenges when estimating relatively ”small” effects, i.e., effects of magnitude similar or lower to the volatility of the outcome variable. See Section 2.6 for a more detailed explanation of how this relates to the interpretation of the results in presented in this section. An alternative explanation for this patter could be an high correlation between wheat prices in European countries that may lead to an almost identical prices of wheat. In this regard, Appendix B.4 shows that while the price of wheat for country included in the analysis is clearly correlated, there are large differences in the price of wheat across countries. This high correlation may however play a role in the results presented in Figure 2.5 and needs to be kept i mind in their interpretation. Finally. Appendix B.5 presents evidence suggesting that no significant change happened in France in the share of arable land cultivated with wheat after the shock, a results which is in line with the lack of effect on wheat prices that we observed in this section.

Figure 2.5

Panel A – Wheat prices



Panel B – Difference between France and Synthetic France in Panel A



Notes: The outcome variable is the price of wheat. The vertical line is set in 2003. Panel A compares the evolution of the outcome variable in France and in Synthetic France. Panel B plots the gap between the two lines in Panel A.

Table 2.6

Country	Weight	Country	Weight
Denmark	0.322	Netherlands	0.665
Finland	0	Portugal	0.013
Greece	0	Romania	0
Italy	0		

Notes: Country weights in synthetic France for the results presented in Figure 2.5

Table 2.7

Variable	France	Synthetic France	Avg. Donor Pool
Wheat prices, dollars (2002)	91.7	94.65	118.73
Wheat prices, dollars (1995)	169.5	177.13	198.71
Wheat prices, dollars (1993)	148.3	163.84	231.82
Share of wheat used in the food sector	0.949	0.916	0.96

Notes: Predictor values for France, Synthetic France and the simple average of all countries in the donor pool. This specification is the one used to obtain the results presented in Figure 2.5.

2.5 Robustness Checks

This section presents the results of the robustness tests described in Section 2.3. The placebo-in-space test and the placebo-in-time test are used to assess the statistical significance and robustness of the results presented in section 2.4.1. As no impact on wheat prices was found, these tests are not performed for the results in Section 2.4.2. The leave-one-out test is instead presented for all our outcome variables to show that the pattern identified in our main results remains robust for all them when we iteratively remove units from the donor pool.

2.5.1 Placebo in Space

Figure 2.6 shows the ratio of pre-treatment RSME to post-treatment RSME for France and all countries in the donor pool. This ratio is particularly high for France, which means that the estimated effect for our treated unit is extreme compared to what we estimate for other units in the donor pool, providing evidence of its statistical significance (Abadie, 2021, Abadie et al., 2010).

2.5.2 Placebo in Time (backdating)

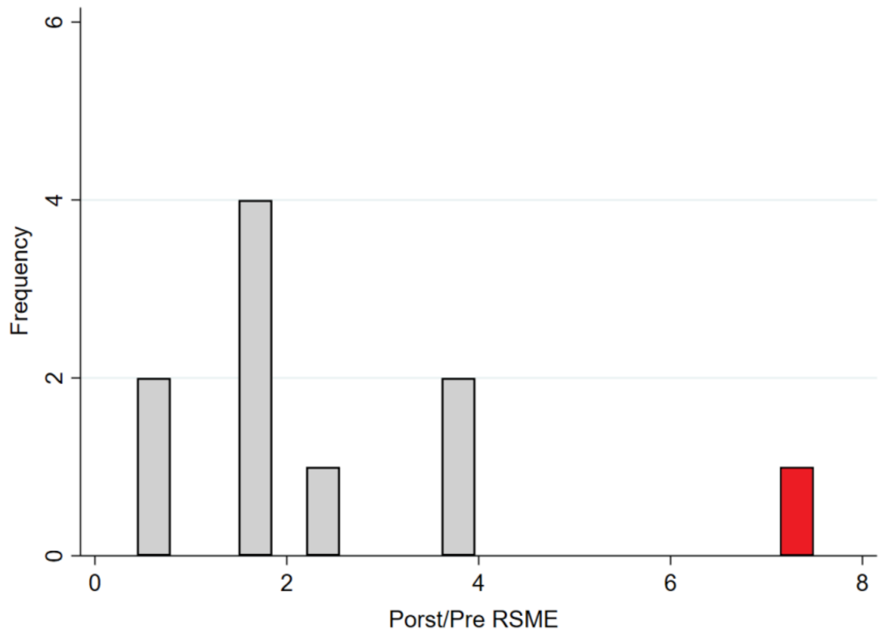
We perform the placebo-in-time test by backdating the shock to bioethanol production capacity to 1995. Figure 2.7 show the results of this exercise. In both cases, synthetic France keeps following France after the “fake” intervention and they only start diverging around 2003, providing evidence in favour of the robustness in of our results (Abadie et al., 2021).

2.5.3 Leave one out Test

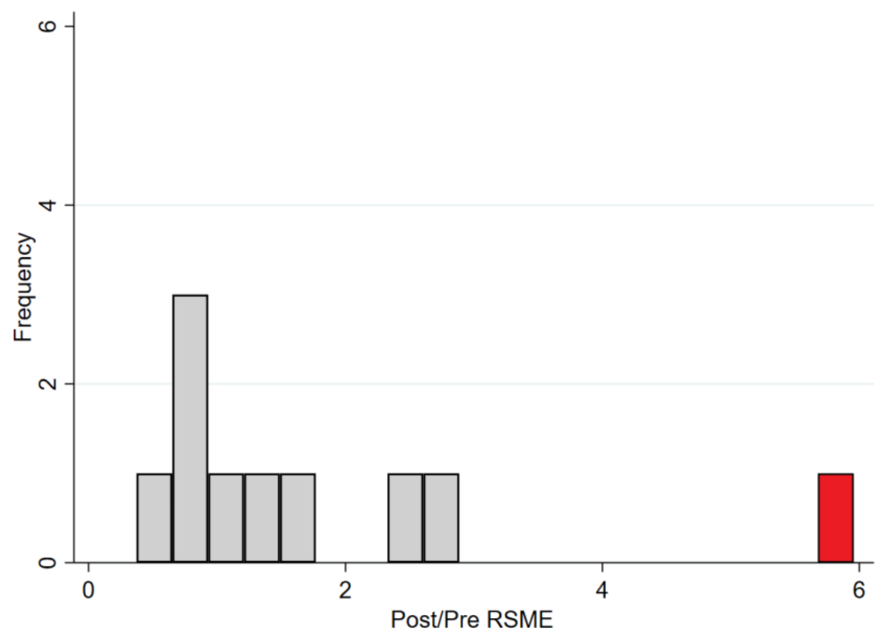
The results of the leave-one-out test show that the pattern observed in Sections 2.4.1. and 2.4.2. is not affected by iteratively removing units from the donor pool (see Figure 2.8). This means that our results are not dependent on any particular unit being in the donor pool.

Figure 2.6: Placebo in Space

Panel A – Post to Pre RMSE ratio - Wheat, non-food uses, (Kg per capita)



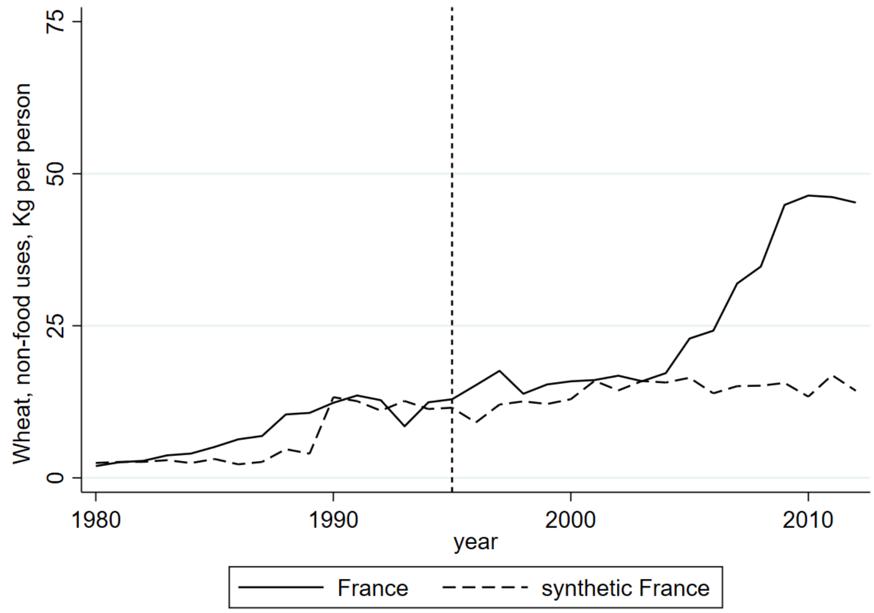
Panel B – Post to Pre RMSE ratio - Share of Wheat used in the food sector



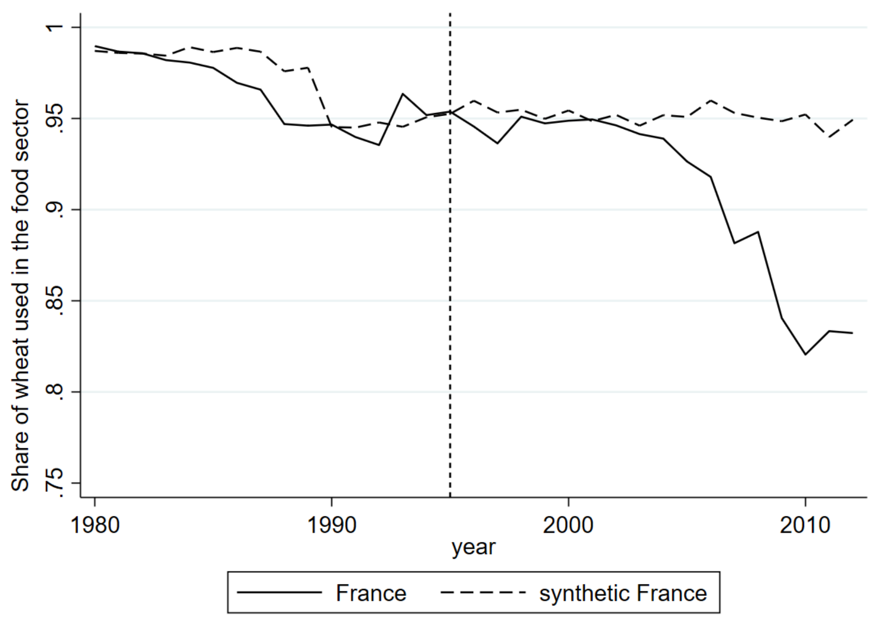
Notes: Ratio of the Root Mean Squared Error in the post- and pre-treatment periods. France in the red bar, other countries in the donor pool are represented in grey.

Figure 2.7: Placebo in Time (backdating)

Panel A – Wheat, non-food uses, Kg per person



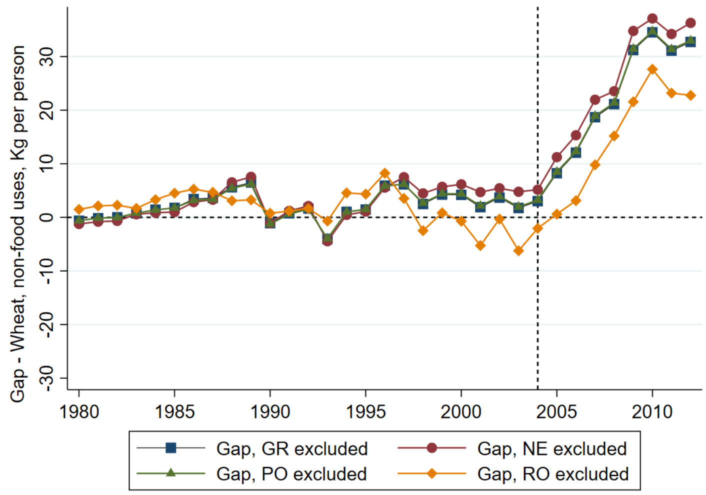
Panel B – Share of wheat used outside the food sector



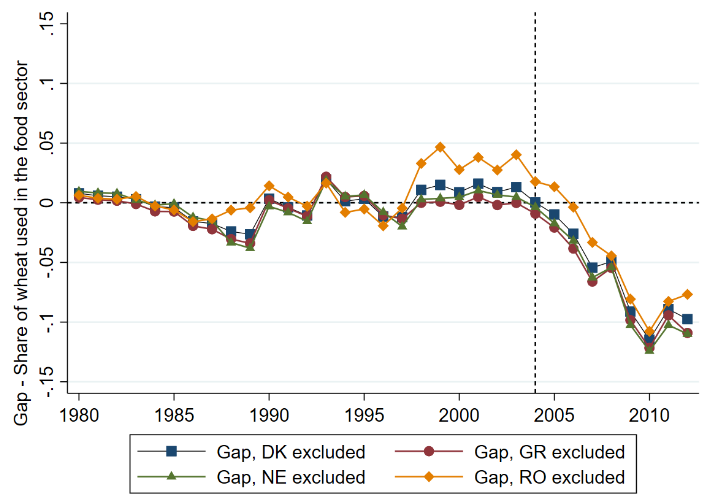
Notes: In Panel A the outcome variable is the wheat used outside the food sector, while in Panel B it is the share of wheat used outside the food sector. The vertical line is set in 1995, the year of the "fake" intervention

Figure 2.8: Leave one out test

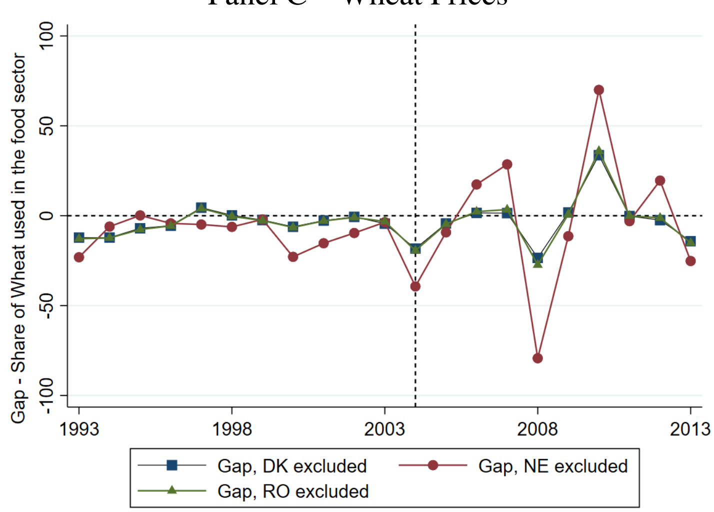
Panel A – Wheat, non-food uses, Kg per person



Panel B – Share of wheat used outside the food sector



Panel C – Wheat Prices



Notes:The test is performed by iteratively remove each country that obtained a positive weight in the results presented in Section 2.4 from the donor pool.

2.6 Discussion

We show that the shock to France bioethanol production capacity observed in 2003 had a strong and statistically significant impact on the amount of wheat used outside the food sector. As a result of this shock, an additional 28 kg of wheat per capita were used for non-food related purposes in 2012, leading to a decrease of 11.5 percentage points in the share of wheat used in the food sector.

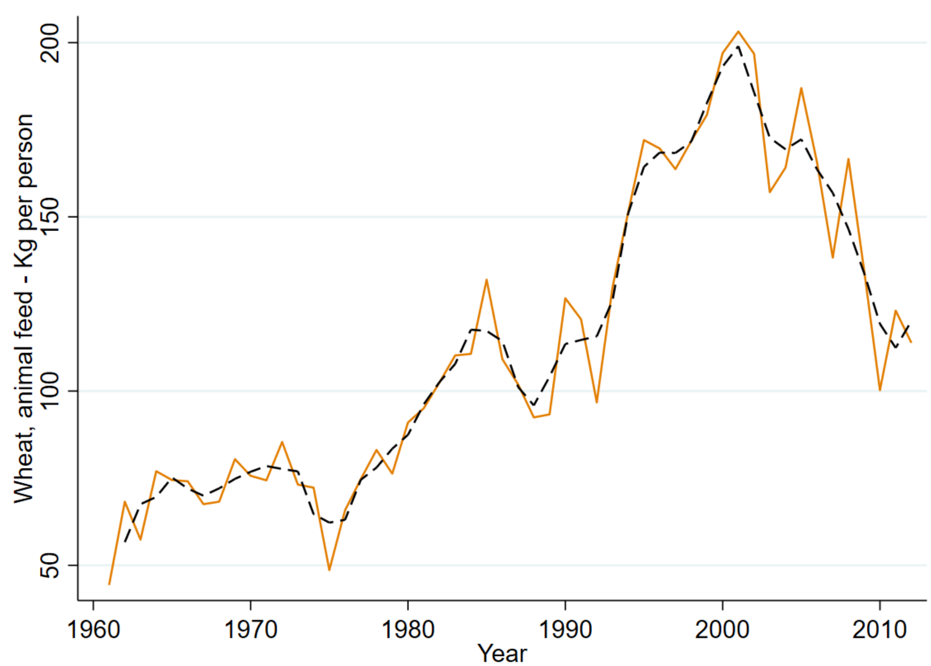
One factor that may have contributed to such a strong response is the decline in wheat used as animal feed that we observe in France over the same period; see Figure 2.9. Between 2000 and 2010, the amount of wheat used as animal feed in the country declined by 100 kg per capita. This decline is not surprising considering that France lost half of its market share in animal products between 2000 and 2016 (Cheptea and Huchet, 2020) and that its production of animal products declined rapidly in the early 2000s after having peaked in the late 1990s; see Figure 2.10. The proposed analysis cannot establish a causal link between the decline in the use of wheat as animal feed and the strong increase in the amount of wheat used outside the food sector in response to the 2003 shock. Similarly, we cannot say whether or not the strong support for bioethanol in France discussed in Section 2.2 is to some extent related to this decline and to the loss of market share in animal products. Nevertheless, both hypotheses seem plausible and we plan to investigate them in future research. In the context of this work, we believe these trends are important to mention as they may have implications for the external validity of our results; in the sense that a similar treatment in a country without these conditions could have a different effect.

According to economic theory, all else being equal, the additional demand of wheat to be used in the fuel sector is expected to put upward pressure on wheat price (Gardebroek et al. 2017, Baffes and Haniotis, 2016; Koizumi, 2015, De Gorter et al., 2015 Busse et al., 2012). This is not reflected in our results, as we found no impact of the shock on the price of wheat. It would however be wrong to interpret our results as conclusive evidence that the 2003 shock to bioethanol production capacity had no impact on wheat price in France. A first reason for this is the increase in the volatility of wheat price in the post-treatment period due to the 2007/2008 global food crisis. Focusing in particular on cereals, between 2005 and 2008 international maize prices almost tripled, international rice price grew by 170 percent and wheat prices increased by 127 percent (Tadasse 2016, Mittal 2009). This can be clearly seen in Figure 2.11, which plots the FAO Cereal index from 1990 to 2023. The index almost doubled in 2007 and then decreased substantially only to spike again in 2010. As we can see in Figure 2.5, wheat prices in France were not immune from this sudden burst in volatility. According to

FAOSTAT data, the price of wheat in France was 129.2 USD/tonnes in 2003, it spiked to 259.2 USD/tonnes by 2007, decreased to 212.5 USD/tonnes in 2008 and then again spiked to 246.6 USD/tonnes in 2010. These fluctuations are significantly larger than the expected impact of the shock to bioethanol production capacity on wheat prices. As pointed out by Abadie (2021), when using the SCM relatively "small" effects, i.e., effects of magnitude similar or lower to the volatility of the outcome, are difficult to detect. This being the case, a different and more granular approach might be needed in order to capture the effect of the shock on wheat prices and distinguish it from the effect of the global food crisis of 2007/2008. Nonetheless, the results outlined in Section 2.4.2 remain relevant. Looking at them, what can infer that if the shock had any effect on wheat prices, as economic theory would strongly suggest, this effect was comparatively low with respect to other drivers that were putting upward pressure on the price of wheat in France during the same time period.

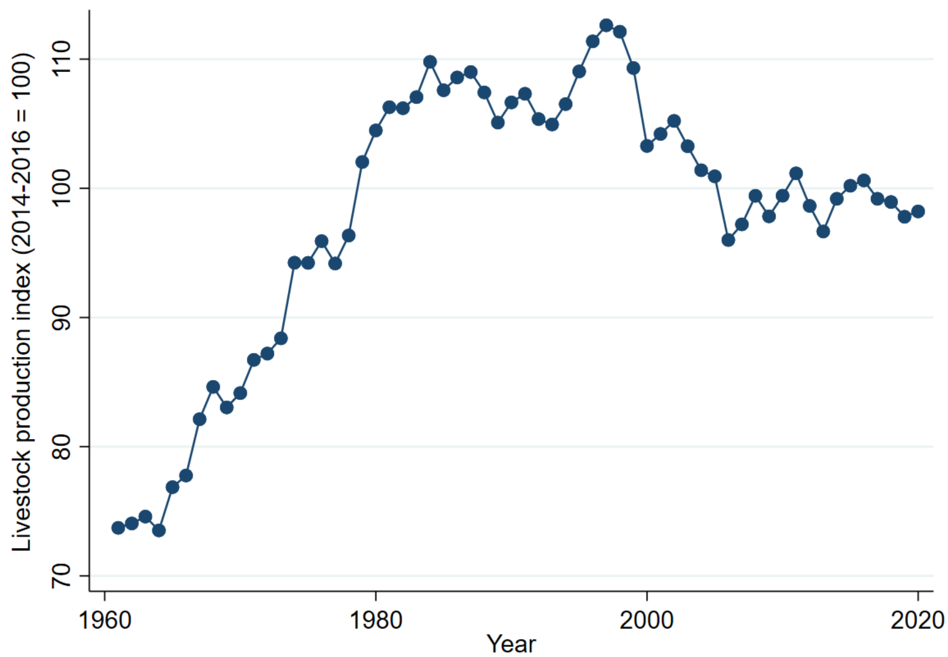
Another reason we cannot conclusively say that the 2003 shock to bioethanol production capacity had no impact on wheat price in France is the correlation in the price of wheat across different countries, which we highlight in Appendix A.4. This high correlation may contribute to hide the impact of the policy shock on wheat prices.

Figure 2.9: Wheat used as animal feed in France



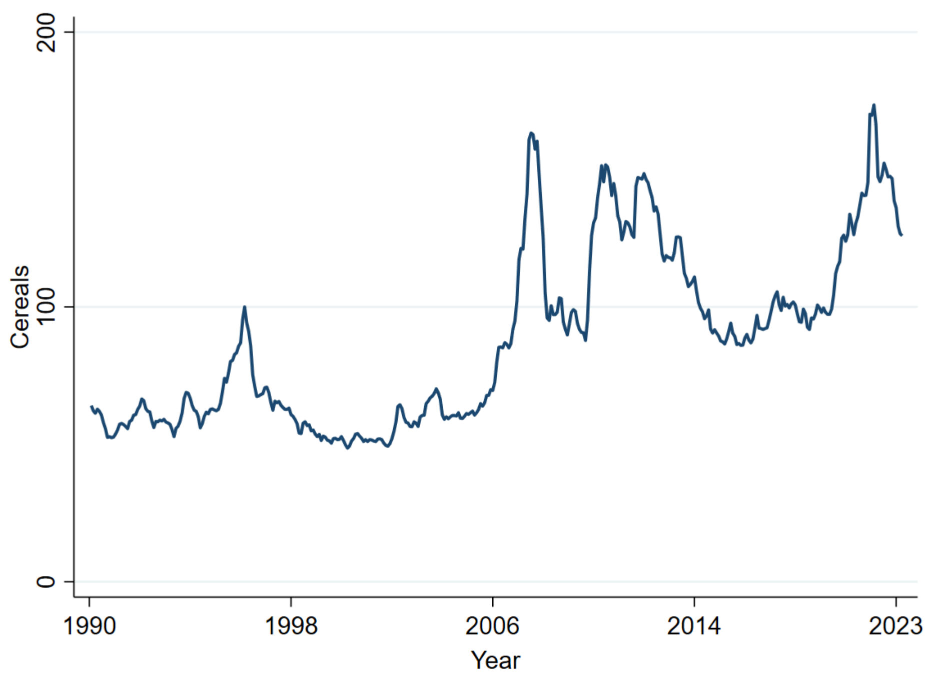
Notes: Source: FAOSTAT, https://www.fao.org/faostat/en/#home

Figure 2.10: Livestock production Index for France



Notes: Source: World Bank, `https://data.worldbank.org/`

Figure 2.11: FAO Cereal Price Index



Notes: Source: FAOSTAT, `https://www.fao.org/worldfoodsituation/foodpricesindex/en/`

2.7 Conclusions and Policy Implications

The popularity of conventional liquid biofuels made from food crops like maize, wheat, and soybeans has sparked a debate on the so called ”food-fuel competition” and its impact on the world’s food security. This competition is a result of the increased demand for agricultural products brought on by biofuel support policies, which can shift crops away from food production and boost food prices.

The European biofuels debate recently heated up amid the Russian invasion of Ukraine and efforts like the “Fit for 55” plan. Advanced biofuels are becoming more popular in Europe, which holds promise, but continued support and policies are needed to overcome technological issues and compete with traditional bioethanol and biodiesel.

By examining the effects of higher bioethanol production capacity on wheat use and price in France, our work contributes to this discussion and sheds light on the impact of supporting conventional biofuels on the price and use of agricultural feedstock. In particular, this study contributes to filling a gap in the existing literature by providing an empirical analysis focused on a European country.

We quantify the impact of the 2003 bioethanol production capacity shock in France by using the Synthetic Control Method. The analysis unveils that historical backing for bioethanol production in France was not a significant driver of wheat prices in the country, at least relative to other factors that were affecting the series in the same time period. This is relevant as the effect of biofuel support policies on food prices is often a prominent concern in the context of the food-fuel competition. At the same time, our results also demonstrate that in the case under study support for bioethanol has substantially increased the quantity and proportion of wheat used in the non-food sector. This stresses the importance of cautious policymaking in the development of future biofuel support strategies, as overreliance on convetional biofuels could potentially lead to relevant food-fuel competition within the European Union.

In future research, we plan to study in more detail whether and to what extent the decline in the production of livestock products and in the use of wheat as animal feed observed in France (see Figures 20.9 and 2.10) played a role in how domestic wheat supply responded to the 2003 shock in bioethanol production capacity. This will allow us to better understand the factors behind the support for traditional biofuels in the European Union.

References

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. Journal of Economic Literature, 59(2), 391-425. DOI:https://doi.org/10.1257/jel.20191450

Abadie, A., Diamond, A., Hainmueller, J. (2015). Comparative politics and the synthetic control method. American Journal of Political Science, 59(2), 495-510. DOI:https://doi.org/10.1111/ajps.12116

Abadie, A., Diamond, A., Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. Journal of the American statistical Association, 105(490), 493-505. DOI:https://doi.org/10.1198/jasa.2009.ap08746

Abadie, A., Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. American economic review, 93(1), 113-132. DOI: https://doi.org/10.1257/000282803321455188

Amigues, J. P., Moreaux, M. (2019). Competing land uses and fossil fuel, and optimal energy conversion rates during the transition toward a green economy under a pollution stock constraint. Journal of Environmental Economics and Management, 97, 92-115. DOI: https://doi.org/10.1016/j.jeem.2019.03.006

Austin, K. G., Jones, J. P. H., Clark, C. M. (2022). A review of domestic land use change attributable to US biofuel policy. Renewable and Sustainable Energy Reviews, 159, 112181. DOI:https://doi.org/10.1016/j.rser.2022.112181

Baffes, J., Haniotis, T. (2016). What explains agricultural price movements?. Journal of Agricultural Economics, 67(3), 706-721. DOI: https://doi.org/10.1111/1477-9552.12172

Berti, P., Levidow, L. (2014). Fuelling expectations: a policy-promise lock-in of UK biofuel policy. Energy Policy, 66, 135-143. DOI: https://doi.org/10.1016/j.enpol.2013.09.044

Boly, M., Sanou, A. (2022). Biofuels and food security: evidence from Indonesia and Mexico. Energy Policy, 163, 112834. DOI: https://doi.org/10.1016/j.enpol.2022.112834

Busse, S., Brümmer, B., Ihle, R. (2012). Price formation in the German biodiesel supply chain: A Markov-switching vector error-correction modeling approach. Agricultural Economics, 43(5), 545-560. DOI: https://doi.org/10.1111/j.1574-0862.2012.00602.x

Cadillo-Benalcazar, J. J., Bukkens, S. G., Ripa, M., Giampietro, M. (2021). Why does the European Union produce biofuels? Examining consistency and plausibility in prevailing narratives with quantitative storytelling. Energy Research Social Science, 71, 101810. DOI: https://doi.org/10.1016/j.erss.2020.101810

Cansino, J. M., del P Pablo-Romero, M., Román, R., Yñiguez, R. (2012). Promotion of biofuel consumption in the transport sector: An EU-27 perspective. Renewable and Sustainable Energy Reviews, 16(8), 6013-6021. DOI: https://doi.org/10.1016/j.rser.2012.06.011

Carter, C. A., Rausser, G. C., Smith, A. (2017). Commodity storage and the market effects of biofuel policies. American Journal of Agricultural Economics, 99(4), 1027-1055. DOI: https://doi.org/10.1093/ajae/aaw010

Cheptea, A., Huchet, M. (2020). The competitiveness of French exports of animal products: A comparative analysis. Économie et Prévision, (1), 89-116.

Charles, C., Gerasimchuk, I., Bridle, R., Moerenhout, T., Asmelash, E., Laan, T. (2013). Biofuels - At What Cost?: A review of costs and benefits of EU biofuel policies. Winnipeg, Manitoba, Canada: International Institute for Sustainable Development. Link: https://www.agrireseau.net/energie/documents/biofuels_subsidies_eu_review.pdf

Cunningham, S. (2021). Causal inference: The mixtape. Yale university press. Link: https://mixtape.scunning.com/

Directive 2003/30/EC of the European Parliament and of the Council of 8 May 2003 on the promotion of the use of biofuels or other renewable fuels for transport. Link: https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A32003L0030

Directive 2003/96/EC of 27 October 2003 restructuring the Community framework for the taxation of energy products and electricity Link: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32003L0096

Efroymson, R. A., Kline, K. L., Angelsen, A., Verburg, P. H., Dale, V. H., Langeveld, J. W., McBride, A. (2016). A causal analysis framework for land-use change and the potential role of bioenergy policy. Land Use Policy, 59, 516-527. DOI: https://doi.org/10.1016/j.landusepol.2016.09.009

Ferman, B., Pinto, C., Possebom, V. (2020). Cherry picking with synthetic controls. Journal of Policy Analysis and Management, 39(2), 510-532. DOI: https://doi.org/10.1002/pam.22206

De Gorter, H., Drabik, D., Just, D. R. (2015). The economics of biofuel policies: Impacts on price volatility in grain and oilseed markets. Springer.

Gardebroek, C., Reimer, J. J., Baller, L. (2017). The impact of biofuel policies on crop acreages in Germany and France. Journal of Agricultural Economics, 68(3), 839-860. DOI: https://doi.org/10.1111/1477-9552.12218

Guinde, L., Jacquet, F., Millet, G. (2008). Impacts of the French bio-fuel policy on the French arable crop sub-sector. DOI: https://doi.org/10.22004/ag.econ.43540

IEA. (2022). World energy outlook 2022. Paris, France: IEA. Link: https://www.iea.org/reports/world-energy-outlook-2022

IFEU (2023). Fehrenbach, H., Bürck, S., Wehrle, A - The Carbon and Food Opportunity Costs of Biofuels in the EU27 plus the UK. Ifft, J., Rajagopal, D., Weldzuis, R. (2019). Ethanol plant location and land use: A case study of crp and the ethanol mandate. Applied economic perspectives and policy, 41(1), 37-55. DOI: https://doi.org/10.1093/aepp/ppy007

Kalkuhl, M., Von Braun, J., Torero, M. (2016). Food price volatility and its implications for food security and policy (p. 626). Springer Nature. DOI: https://doi.org/10.1007/978-3-319-28201-5

Kaul, A., Klößner, S., Pfeifer, G., Schieler, M. (2022). Standard synthetic control methods: The case of using all preintervention outcomes together with covariates. Journal of Business Economic Statistics, 40(3), 1362-1376. DOI: https://doi.org/10.1080/07350015.2021.1930012

Khanna, M., Rajagopal, D., Zilberman, D. (2021). Lessons learned from US experience with biofuels: Comparing the hype with the evidence. Review of Environmental Economics and Policy, 15(1), 67-86. DOI: https://doi.org/10.1086/713026

Koizumi, T. (2015). Biofuels and food security. Renewable and Sustainable Energy Reviews, 52, 829-841. DOI: https://doi.org/10.1016/j.rser.2015.06.041

Kutas, G., Lindberg, C., Steenblik, R. (2007). Biofuels–at what Cost?: Government Support for Ethanol and Biodiesel in the European Union (pp. 14-25). Geneva: International Institute for Sustainable Development. Link: http://www.iisd.org/gsi/sites/default/files/subsidies_to_biofuels_in_the_eu_final.pdf

Lark, T. J., Hendricks, N. P., Smith, A., Pates, N., Spawn-Lee, S. A., Bougie, M., ... Gibbs, H. K. (2022). Environmental outcomes of the US renewable fuel standard. Proceedings of the National Academy of Sciences, 119(9), e2101084119. DOI: https://doi.org/10.1073/pnas.2101084119

Li, Y., Miao, R., Khanna, M. (2019). Effects of ethanol plant proximity and crop prices on land-use change in the United States. American Journal of Agricultural Economics, 101(2), 467-491. DOI: https://doi.org/10.1093/ajae/aay080

Millinger, M., Ponitka, J., Arendt, O., Thrän, D. (2017). Competitiveness of advanced and conventional biofuels: Results from least-cost modelling of biofuel competition in Germany. Energy Policy, 107, 394-402. DOI: https://doi.org/10.1016/j.enpol.2017.05.013

Mittal, A. (2009). The 2008 food price crisis: rethinking food security policies (pp. 1-40). New York: UN. Mohanty, S. K., Swain, M. R. (2019). Bioethanol production from corn and wheat: food, fuel, and future. In Bioethanol production from food crops (pp. 45-59). Academic Press. DOI: https://doi.org/10.1016/B978-0-12-813766-6.00003-5

Muscat, A., De Olde, E. M., de Boer, I. J., Ripoll-Bosch, R. (2020). The battle for biomass: a systematic review of food-feed-fuel competition. Global Food Security, 25, 100330. DOI: https://doi.org/10.1016/j.gfs.2019.100330

Newes, E., Clark, C. M., Vimmerstedt, L., Peterson, S., Burkholder, D., Korotney, D., Inman, D. (2022). Ethanol production in the United States: The roles of policy, price, and demand. Energy policy, 161, 112713. DOI: https://doi.org/10.1016/j.enpol.2021.112713

Pilgrim, S., Harvey, M. (2010). Battles over biofuels in Europe: NGOs and the politics of markets. Sociological Research Online, 15(3), 45-60. DOI: https://doi.org/10.5153/sro.2192

Popp, J., Harangi-Rákos, M., Gabnai, Z., Balogh, P., Antal, G., Bai, A. (2016). Biofuels and their co-products as livestock feed: global economic and environmental implications. Molecules, 21(3), 285. DOI: https://doi.org/10.3390/molecules21030285

Sant'Anna, M. (2021). How green is sugarcane ethanol?. Review of Economics and Statistics, 1-45. DOI: https://doi.org/10.1162/rest_a_01136

Searchinger, T., Edwards, R., Mulligan, D., Heimlich, R., Plevin, R. (2015). Do biofuel policies seek to cut emissions by cutting food?. Science, 347(6229), 1420-1422. DOI: https://doi.org/10.1126/science.1261221

Sorda, G., Banse, M., Kemfert, C. (2010). An overview of biofuel policies across the world. Energy policy, 38(11), 6977-6988. DOI: https://doi.org/10.1016/j.enpol.2010.06.066

The Council of the European Union, Council Regulation (EC) No 1782/2003 of 29 September 2003 establishing common rules for direct support schemes under the common agricultural policy and establishing certain support schemes for farmers and amending Regulations (EEC) No 2019/93, (2003). Link: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32003R1782

USDA (2022). European Union Grain and Feed Annual. Madrid, Spain: USDA. Link: https://www.fas.usda.gov/data/european-union-grain-and-feed-annual-2

USDAb (2022). European Union: Biofuels annual, The Hague, Netherlands: USDA. Link: https://www.fas.usda.gov/data/european-union-biofuels-annual-2

Wiesenthal, T., Leduc, G., Christidis, P., Schade, B., Pelkmans, L., Govaerts, L., Georgopoulos, P. (2009). Biofuel support policies in Europe: Lessons learnt for the long way ahead. Renewable and Sustainable Energy Reviews, 13(4), 789-800. DOI: https://doi.org/10.1016/j.rser.2008.01.011

Wright, C. K., Larson, B., Lark, T. J., Gibbs, H. K. (2017). Recent grassland losses are concentrated around US ethanol refineries. Environmental Research Letters, 12(4), 044001. DOI:10.1088/1748-9326/aa6446

2.8 Appendix B

B.1 Robustness to specificaton searching

The issue of specification searching in empirical applications using the synthetic control method arises when the researcher has control over the results by changing the set of predictors used in the analysis. As a robustness check against specification searching, Ferman Pinto and Posserborn (2021) propose to present results for different specifications that rely on all or a subset of pre-treatment lags of the outcome variable (Ferman Pinto and Posserborn, 2021). Kaul et al. (2021) show that by including all lags of the outcome as separate predictors we make other matching variables irrelevant and therefore advise against this in applications where we believe these covariates to be important.

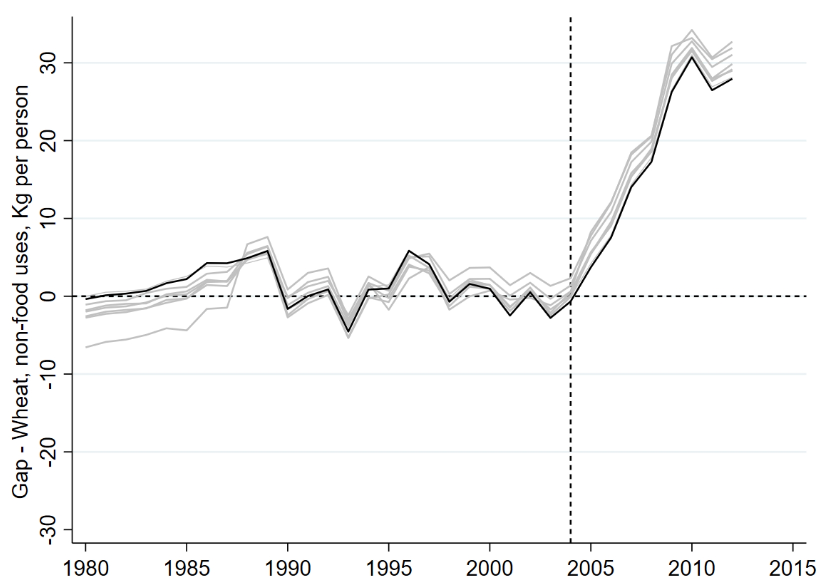
Borrowing from both these papers, we present results for the following 8 additional specifications additional specifications:

- **Specification A**: All pre-treatment lags of the outcome variables
- **Specification B**: Only even pre-treatment lags of the outcome variables
- **Specification C**: Only odd pre-treatment lags of the outcome variables
- **Specification D**: Only the first half of the pre-treatment lags of the outcome variables
- **Specification E**: The average of pre-treatment values of the outcome variables
- **Specification F**: The initial and final value of pre-treatment values of the outcome variable with the addition of wheat yields and the per capita domestic what supply as predictor variables. Furthermore, depending on the outcome variable we also add the share of wheat used in the food sector or the amount of wheat used outside the food sector following the same rationale as in the main analysis (see Section 2.3.5).
- **Specification G**: The initial and final value of pre-treatment values of the outcome variable with the addition of wheat yields, the per capita domestic what supply as predictor variables, the share of all cereals used outside the food sector, the share of all cereals used for direct human consumption, the share of all cereals used as animal feed and the terrain ruggedness index, i.e. an index measuring the amount of elevation difference at the country level

Specifications A to E are taken directly from Ferman et a., (2021), while in Specifications F and G we rely more heavily on covariates that might help predict the outcome variables of interest.

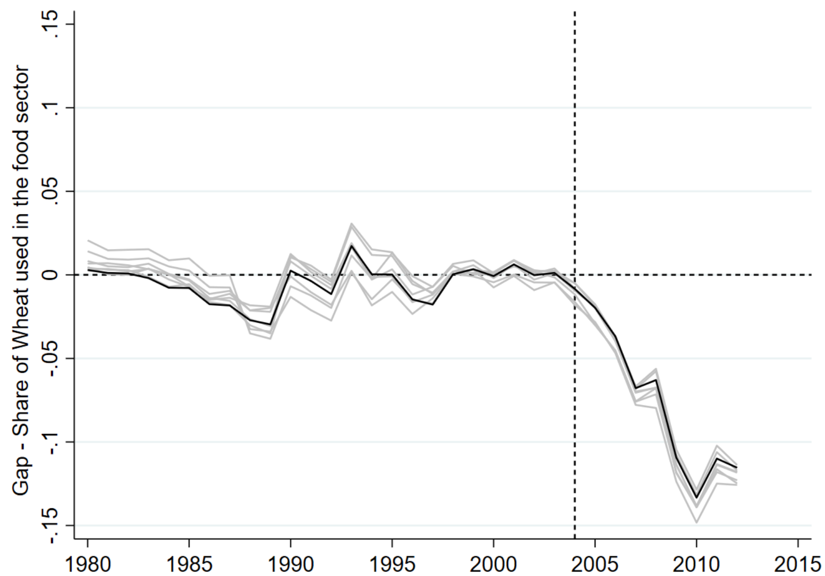
In Figures B.1, B.2 and B.3 we compare results from these additional specifications with results from the specifications used in the main analysis and show that different specifications result in similar estimates of the effect of interest.

Figure B.1: Specification Searching – Wheat, non-food uses, Kg per Person



Notes: The figure compares the gap between France and Synthetic France obtained using the main specification (black line) and specifications A to G described in the Appendix.

Figure B.2: Specification Searching – Share of wheat used outside the food sector



Notes:The figure compares the gap between France and Synthetic France obtained using the main specification (black line) and specifications A to G described in the Appendix.

Figure B.3: Specification Searching – Wheat prices

Notes: The figure compares the gap between France and Synthetic France obtained using the main specification (black line) and specifications A to G described in the Appendix.

B.2 Larger Donor Pool

In this Section we present the results including in the Donor Pool all countries available, i.e. we do not exclude from the sample countries that experienced a shock in bioethanol production capacity that, while smaller than the one observed in France, can be considered sizable. Interestingly the results using this larger donor are very similar to the one presented in the main analysis even if there are differences in the countries used to build Synthetic France.

Figure B.4: Wheat used outside the food sector

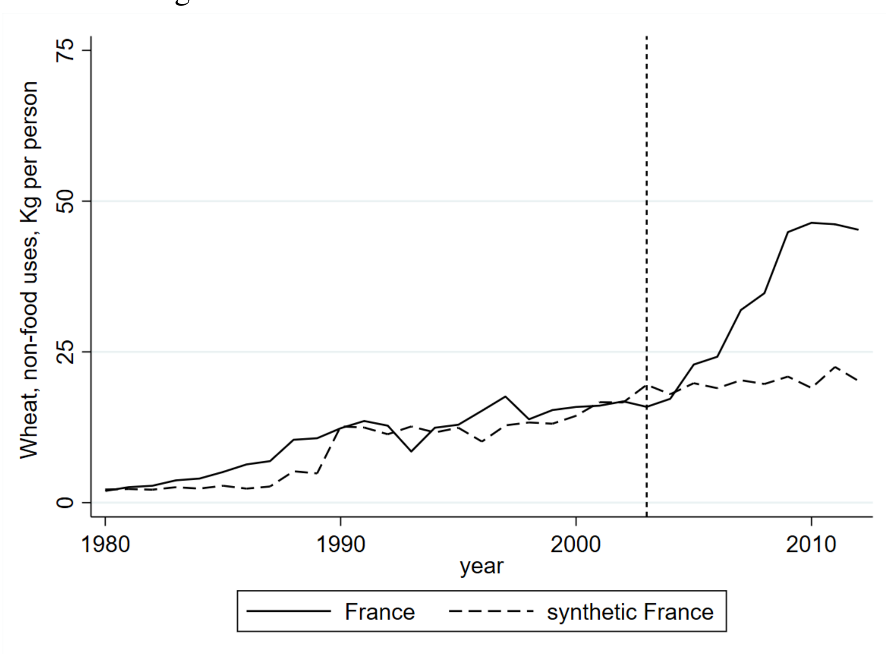


Table B.1

Country	Weight	Country	Weight
Austria	0	Ireland	0
Bulgaria	0	Italy	0
Cyprus	0	Netherlands	0.204
Finland	0	Poland	0.245
Denmark	0	Portugal	0
Germany	0.37	Romania	0.179
Greece	0	Spain	0
Hungary	0	Sweden	0

Notes: Country weights in synthetic France for the results presented in Figure B.4

Figure B.5: Share of wheat used in the Food Sector

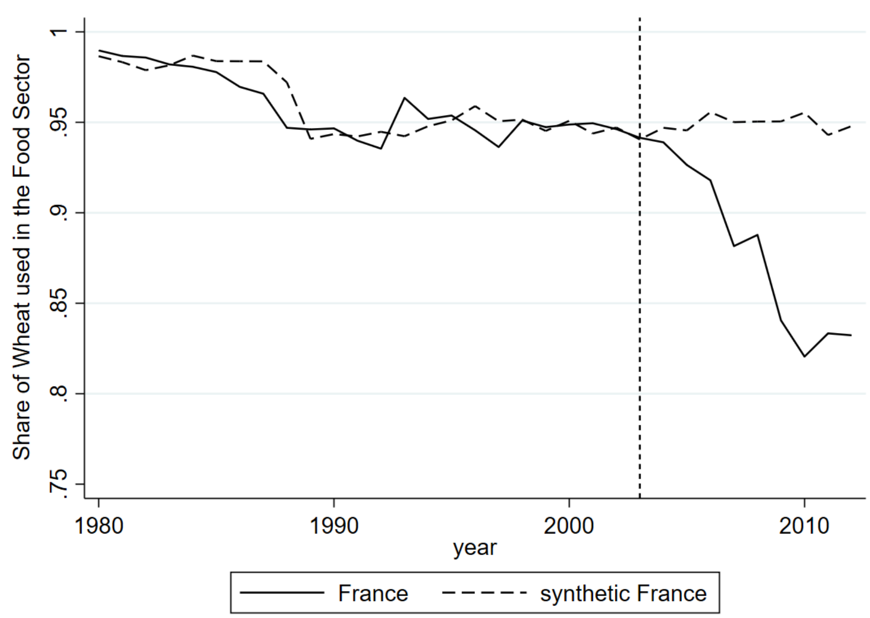


Table B.2

Country	Weight	Country	Weight
Austria	0	Ireland	0
Bulgaria	0	Italy	0
Cyprus	0	Netherlands	0.133
Finland	0	Poland	0.019
Denmark	0	Portugal	0
Germany	0.37	Romania	0.189
Greece	0.007	Spain	0
Hungary	0.651	Sweden	0

Notes: Country weights in synthetic France for the results presented in Figure B.5

Figure B.6: Wheat price

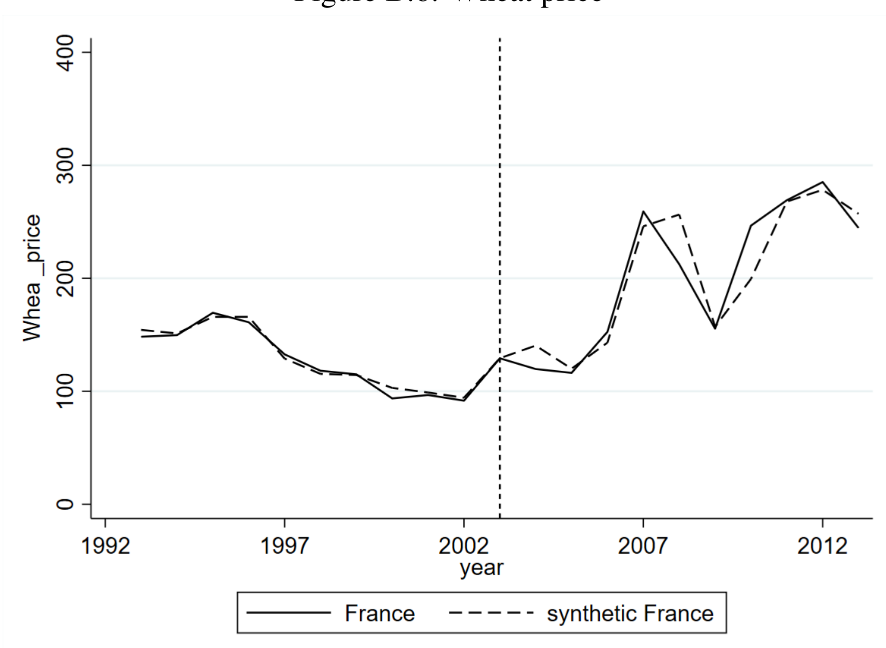


Table B.3

Country	Weight	Country	Weight
Austria	0	Italy	0
Cyprus	0	Netherlands	0.502
Finland	0	Poland	0
Denmark	0.301	Portugal	0
Germany	0	Romania	0
Greece	0	Spain	0
Hungary	0.126	Sweden	0.071

Notes: Country weights in synthetic France for the results presented in Figure B.6

B.3 Results - Cereals

In the main analysis, we investigate the effect of the 2003 shock to bioethanol production capacity to wheat use and prices. The reasons why we focus precisely on wheat are explained in the main text. This section shows that the main conclusions of the analysis hold even if we focus on cereals in general. The pattern of Synthetic France suggests that countries in the Donor Pool relied less on wheat and more on other cereals to produce bioethanol compared to France. A comparison of Figure B.7 with Figure 2.3 shows instead that, as expected, in France wheat was by far the cereal that reacted most to the shock in bioethanol production capacity. This was expected as in the decade before the shock (1992-2002) wheat accounted for 74% of all cereals used outside the food sector in the country (as discussed in Section 2.1) and in light of the discussion presented in Section 2.6.

Figure B.7: Cereals used outside the food sector

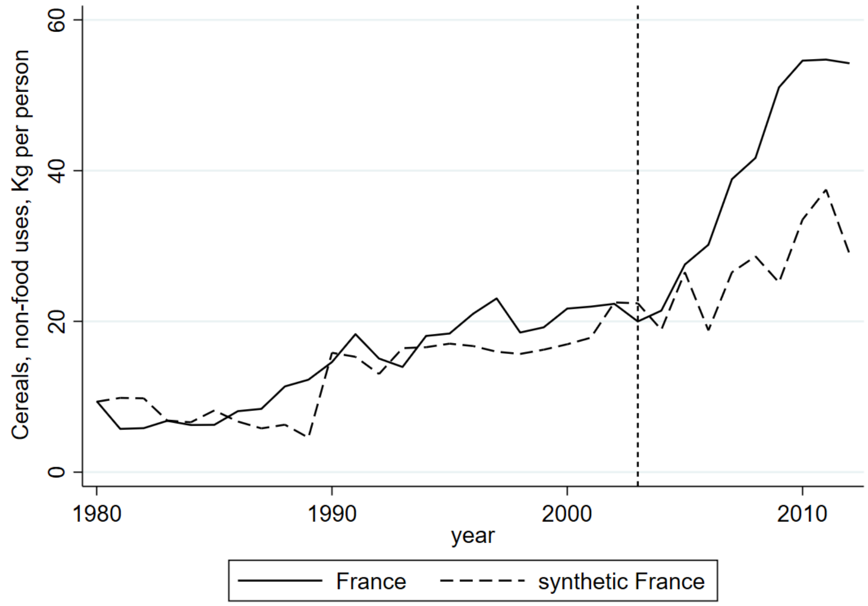


Table B.4

Country	Weight	Country	Weight
Bulgaria	0.371	Italy	0
Denmark	0	Netherlands	0
Finland	0.04	Portugal	0.256
Greece	0	Romania	0.333
Ireland	0		

Notes: Country weights in synthetic France for the results presented in Figure B.7

Table B.5

Variable	France	Synthetic France	Avg. Donor Pool
Cereals, non-food uses, Kg per person (2002)	22.33	22.51	18.60
Cereals, non-food uses, Kg per person (1995)	18.39	17.04	14.51
Cereals, non-food uses, Kg per person (1990)	14.61	15.83	14.71
Cereals, non-food uses, Kg per person (1980)	6.27	8.19	15.46
Share of cereals used in the food sector	0.977	0.984	0.96

Notes: Predictor values for France, Synthetic France and the simple average of all countries in the donor pool. This specification is the one used to obtain the results presented in Figure b.7

Figure B.8: Share of Cereals used in the food sector

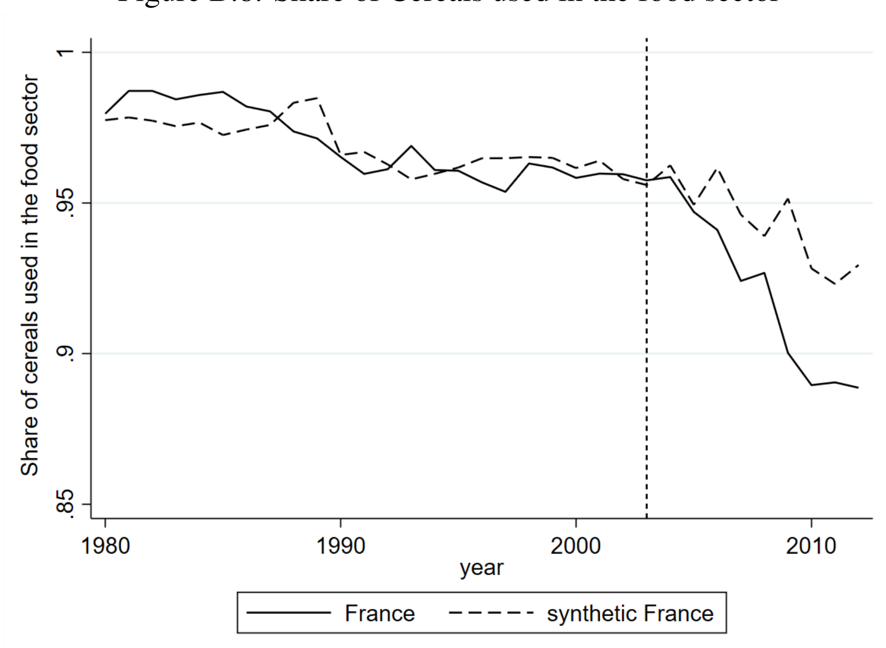


Table B.6

Country	Weight	Country	Weight
Bulgaria	0.371	Italy	0
Denmark	0	Netherlands	0
Finland	0.04	Portugal	0.256
Greece	0	Romania	0.333
Ireland	0		

Notes: Country weights in synthetic France for the results presented in Figure B.8

Table B.7

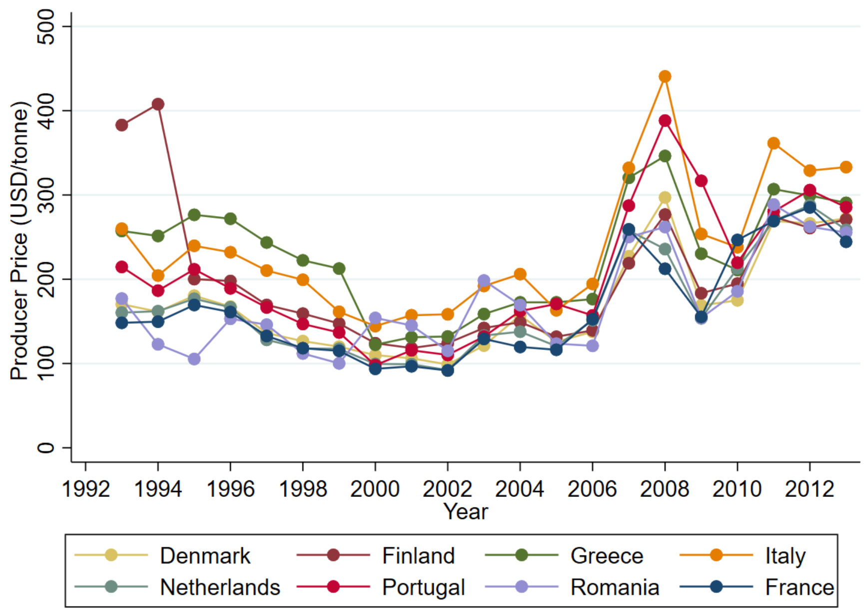
Variable	France	Synthetic France	Avg. Donor Pool
Wheat prices, dollars (2002)	91.7	94.65	118.73
Wheat prices, dollars (1995)	169.5	177.13	198.71
Wheat prices, dollars (1993)	148.3	163.84	231.82
Share of wheat used in the food sector	0.949	0.916	0.96

Notes: Predictor values for France, Synthetic France and the simple average of all countries in the donor pool. This specification is the one used to obtain the results presented in Figure B.8

B.4 Wheat Prices in European Countries

Figure B.9 shows the price of wheat for country included in the analysis in paragraph 2.4.2. The graph shows that while the price of wheat follows a similar trend for all countries in the sample, there are large differences in the price of wheat across countries.

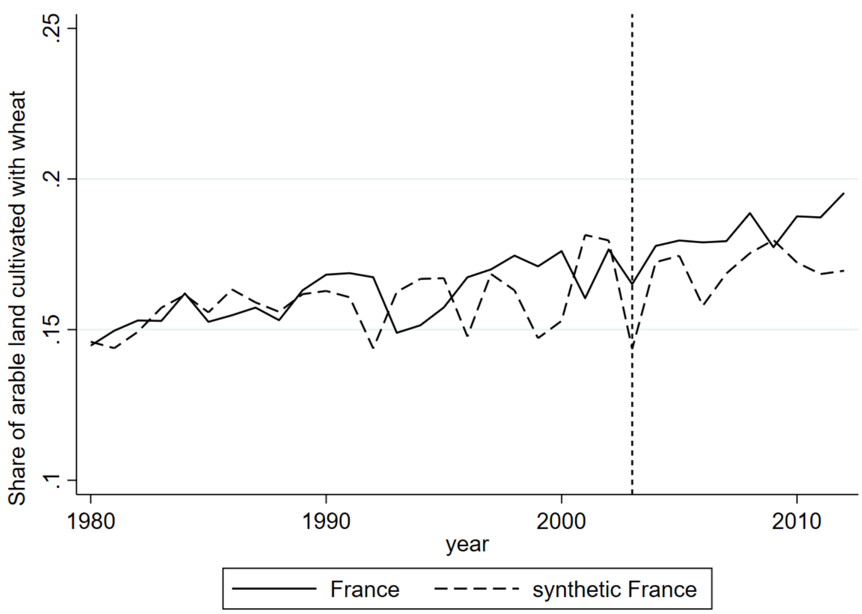
Figure B.9: Share of Cereals used in the food sector



B.5 Share of Arable Land cultivated with Wheat

In this Section, we briefly look at the evolution of the share of land allocated to wheat in France during the period of interest. Country level data on overall arable land come directly from FAOSTAT, while the arable land allocated to wheat has been estimated using data on wheat production and wheat yields, which are also contained in the FAOSTAT database. As we can see, the share of agricultural land used to grow wheat remained more or less stable between 1980 and 2012. A small positive difference between France and Synthetic France may seem to emerge in the post-treatment period, but further tests suggest this difference is far from being significant and not robust to different specifications. This result is in line with the lack of a statistically significant effect on wheat prices.

Figure B.10: Share of Arable Land cultivated with Wheat



Country	Weight	Country	Weight
Bulgaria	0.261	Italy	0.303
Denmark	0.101	Netherlands	0
Finland	0.087	Portugal	0
Greece	0	Romania	0.248
Ireland	0		

Notes: Country weights in synthetic France for the results presented in Figure B.10

Variable	France	Synthetic France	Avg. Donor Pool
Share wheat on arable land (2002)	0.176	0.180	0.123
Share wheat on arable land (1995)	0.157	0.167	0.115
Share wheat on arable land (1990)	0.168	0.163	0.113
Share wheat on arable land (1980)	0.145	0.146	0.097
Share of wheat used in the food sector	0.976	0.975	0.962
Wheat, non-food uses, Kg per person	5.97	4.90	6.50

Notes: Country weights in synthetic France for the results presented in Figure B.10