INSTITUTE
OF ECONOMICS

Scuola Superiore
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM
WORKING PAPER SERIES

Climate change and labour-saving technologies: the twin transition via patent texts

Tommaso Rughi [a]
Jacopo Staccioli [b,a]
Maria Enrica Virgillito [a,b]

[a] Institute of Economics and EMbeDS, Scuola Superiore Sant'Anna, Pisa, Italy.
[b] Department of Economic Policy, Università Cattolica del Sacro Cuore, Milano, Italy.

2023/11 **March 2023**
ISSN(ONLINE) 2284-0400

Climate change and labour-saving technologies: the twin transition via patent texts*

Tommaso Rughi†[1], Jacopo Staccioli[2,1], and Maria Enrica Virgillito[1,2]

[1]Institute of Economics and EMbeDS, Scuola Superiore Sant'Anna, Pisa (Italy)

[2]Department of Economic Policy, Università Cattolica del Sacro Cuore, Milano (Italy)

Abstract

This paper provides a direct understanding of the twin transition from the innovative activity domain. It starts with a technological mapping of the technological innovations characterised by both climate change mitigation/adaptation (green) and labour-saving attributes. To accomplish the task, we draw on the universe of patent grants in the USPTO since 1976 to 2021 reporting the Y02-Y04S tagging scheme and we identify those patents embedding an explicit labour-saving heuristic via a dependency parsing algorithm. We characterise their technological, sectoral and time evolution. Finally, after constructing an index of sectoral penetration of LS and non-LS green patents, we explore its impact on employment share growth at state level in the US. Our evidence shows that employment shares in sectors characterised by a higher exposure to LS (non-LS) technologies present an overall negative (positive) growth dynamics.

Keywords: Climate change mitigation technologies, Labour-saving technologies, Search heuristics, Natural Language Processing, Labour markets

JEL classification: C38, J24, O33, Q55

*We acknowledge comments and suggestions received by the participants to the SESTEF 2022 Conference, Paris-Saclay. We thank Stefan Thurner and the EcoFin group at the Complex Science Hub of Vienna as well as Frank Neffke, Florent Bédécarrats for the fruitful exchanges and suggestions.

†Corresponding author: tommaso.rughi@santannapisa.it

1 Introduction

An increasing consensus, which encompasses also international financial institutions such as the IMF (International Monetary Fund, 2022), is emerging on the urgency to tackle the climate crisis thorough the mitigation of global warming, in particular with a substantial reduction in greenhouse gas (GHG) emissions. The transition to a green economy is defined by UNEP as "[...] *low carbon, resource efficient and socially inclusive. In a green economy,* ***growth in employment*** *and income are driven by public and private investment into such economic activities, infrastructure and assets that allow reduced carbon emissions and pollution, enhanced energy and resource efficiency, and prevention of the loss of biodiversity and ecosystem services"*.[1] It therefore entails the effort of a plurality of actors, both private and public, in achieving a low or even null level of climate impacts in terms of greenhouse emissions.[2] Despite some serious limitations and drawbacks that the green economy and green growth paradigms encompass, highlighted by various research streams (Unmüßig et al., 2012; Van Vuuren et al., 2017; Hickel and Kallis, 2020; D'Alessandro et al., 2020), they still represent the core in terms of both reflections for the academic community and implementation for policy makers and business actors.

More recently, in order to achieve a sustained green growth, policy makers and particularly the European Union, are focusing on the so called *twin transition*, defined as the conjunction between the digital transition, aimed at increasing the overall productivity of the economy, and the effort to foster environmental processes and technologies to achieve climate sustainability. Such efforts are somehow even intertwined with the stated objective of promoting a *just transition*, according to which *"[...] A solid knowledge base is needed to interlink the digital and green transitions with the social dimension of the just transition and to ensure that 'no one is left behind' "* (Stefan et al., 2021).

The two transformations entail a common threat: the possibility of losing jobs at the cost of gaining in environmental sustainability on the one hand, and in productivity efficiency on the other. The costs of such transitions are going to be heterogeneous across sectors and countries, especially according to the identification of most exposed sectors and occupations. While the common understanding tends to identify the green trajectory as mainly labour augmenting (International Labour Office, 2018), it is still lacking a clear mapping of the underlying heuristics of innovators in the climate change domains in terms of labour-efficiency processes. It might be the case that environmental innovations also come with lower labour-input requirements, therefore challenging the common wisdom of the green transition as net job creator.

This paper intends to fill this gap providing a direct understanding of the twin transition from the innovative activity domain. It starts with a technological mapping of innovations characterised by both climate change mitigation/adaptation (green, thereafter) and labour-saving attributes. To accomplish this task, we draw on the universe of patent grants by the USPTO from 1976 to 2021 with at least one CPC

[1] UNEP.org (access 08/03/2023).

[2] The Green New Deal, presented by the European Commission on 11^{th} December 2019, represents one of the most ambitious public plans on this respect (GreenNewDeal.EU).

code of either Class `Y02` or Subclass `Y04S`, which refer to green technologies.[3] We identify those patents embedding an explicit labour-saving (LS thereafter) heuristic via a dependency parsing algorithm. Next, we characterise their technological, sectoral and time evolution. Finally, after constructing an index of sectoral penetration of LS and non-LS green patents, we explore its impact on employment share growth at state level in the US.

Our contribution puts forth a methodological advancement in studying and detecting the twin transition: in fact, we depart from the task-based approach, which only accounts for the occupational components and their inherent degree of "greenness", and we move to a method based on patent full-texts, able to construct a direct measure of technological penetration. Our methodological approach, which relies upon advanced semantic analysis and natural language processing (Montobbio et al., 2022b), allows us to investigate the inventors' heuristics embedded in green patents and detect the extent to which they incorporate a true LS trait and scope. Our method of analysis allows therefore to move from the technological domain to the labour market domain, constructing a multi-layer and integrated interface of analysis.

Our results detect, first, a rapid increase of LS heuristics in the majority of green technological domains considered, and, second, a negative significant impact upon employment shares growth in the sectors more exposed to the *use* of these technologies, therefore validating ex-post the penetration of such heuristics. In a nutshell, our findings challenge the common understanding of the "green transition" as only labour augmenting. Potentially, the capacity of the "green" segment as a net labour-absorber might be weaker than commonly expected. Direct policy interventions are therefore necessary beyond adaptation policies to "green skills" currently envisaged by institutions.

The paper is organised as follows. Section 2 discusses the extant literature, while section 3 presents the relevant data sources. Our methodology is outlined in section 4, where we describe the steps to identify LS heuristics in green related patents, including the novel use of the spaCy neural network model (Honnibal and Montani, 2017). After the identification of two sets of green related patents (either associated to LS heuristics or not), we present our results in section 5, which includes descriptive statistics emerging from our identification strategy (5.1) and the employment impacts related to the penetration of labour-saving heuristics in different industrial sectors (5.2). Our conclusions are presented in section 6.

2 Technologies, labour markets and the green transition: state of the art and open research questions

In order to study the twin transition, we mobilise two main research streams: the first studies the effects of technological changes upon labour markets, with specific attention to digital and automation technologies,

[3] In particular, `Y02` includes "Technologies or applications for mitigation or adaptation against climate change", while `Y04S` "Systems integrating technologies related to power network operation, communication or information technologies for improving the electrical power generation, transmission, distribution, management or usage, i.e. smart grids".

while the second studies the characteristics of green jobs.

Reflections and concerns about possible negative effects of technological change upon the labour market can be traced back to the dawn of the history of capitalism (Staccioli and Virgillito, 2021), where the vast introduction of capital machines, at the beginning of the First Industrial Revolution, generated awareness among workers of the possible pernicious impact on labour, with the Luddites movement representing a paradigmatic example (Nuvolari et al., 2002). The challenging relationship between technological change and labour persisted across the XX century (L. Barbieri et al., 2019), along with the adoption of the steam engine and later with the ICT revolution (Noble, 1986; Zuboff, 1988). In the past decade, those worries involved specifically the new technological trend dubbed Industry 4.0, spurring debates on the effects of automated processes and industrial robots upon employment (Brynjolfsson and McAfee, 2014; Frey and Osborne, 2017; Acemoglu and Restrepo, 2020). Results are however quite inconclusive and mostly depend upon the level of aggregation considered and on the type of technological proxy used in the study (Montobbio et al., 2022a).

The efforts to decarbonise economic products and processes in order to achieve better environmental sustainability have gained increasing traction among scholars as an object of study. Empirical attempts devoted both to analyse the characteristics and knowledge base of green technologies and the related labour market have been rising. Green technologies and their characteristics have been largely studied at the regional level (Tanner, 2014; Corradini, 2019; Quatraro and Scandura, 2019; Montresor and Quatraro, 2020; N. Barbieri et al., 2021; Santoalha et al., 2021; N. Barbieri et al., 2022), sectoral level, for instance in the automotive sector (Mazzei et al., 2022), and micro level (N. Barbieri et al., 2020).

To study green labour markets, numerous empirical methods concur (Bowen and Kuralbayeva, 2015), but analyses which draw upon the task-based approach, in line with the literature on inequality and technologies, are the most widespread and adopted (Dierdorff et al., 2009; Vona et al., 2018; Vona, 2021; International Monetary Fund, 2022; Curtis and Marinescu, 2022). Contributions are increasingly providing new evidence, especially from the O*NET-SOC database (Dierdorff et al., 2009). Vona et al. (2018) and Vona (2021) develop a method, based on task contents and their level of greenness, to measure and define green employment.

Adopting the same approach, the IMF has recently dedicated a chapter to the green transition (International Monetary Fund, 2022), documenting that the most green and most polluting-intensive jobs are concentrated in terms of workforce and sectors, even if environmental characteristics of jobs are widely dispersed both across and within sectors leaving scope for reallocation of workers. Second, green intensive occupations tend to be associated with high skills and urban workers, as opposed to brown occupations, therefore green jobs seem to show a higher degree of complexity. Third, transitioning from brown or neutral to green jobs seems less likely than languishing in similar types of occupations. Finally, environmental polices might prove effective in greening jobs, but only if well tailored.

Curtis and Marinescu (2022) move beyond the O*NET-SOC dataset and employ online vacancy data in the US collected by Burning Glass Technologies. In defining green jobs, the paper looks specifically at the open positions in wind and solar sectors between 2010 and 2019. Both sub-sectors exhibit substantial growth rates, especially since 2013. A relevant share of solar jobs (approximately one third) is in sales, while a similar share is scattered across installation and maintenance: coherently with these results, the most common industry for wind jobs is manufacturing (29%), while utilities play an important although less relevant role for both categories of green jobs (about 15-16%). With regards to the pay premium, it is higher for green jobs even when controlling for educational level, and for jobs that require lower education. Finally, in the US, green jobs tend to be localised in specific areas characterised by a high share of oil & gas sectoral employment. Other applications using BGT data are in Saussay et al. (2022), focussing on employment reallocation across so called low-carbon and high-carbon jobs, and the ensuing cost of transition for affected workers.

These recent studies are but an example of the growing efforts to map and better characterise the green transition, which still begs for higher quality data and further research (Vona, 2021). According to our reading, the main limitation of the task-based approach, together with deeper theoretical flaws discussed in Staccioli and Virgillito (2021) and especially in the identification of green jobs, relies on the fact that the degree of greenness can hardly be inferred by the content reported within O*NET task descriptions. The approach to green occupations is based on Dierdorff et al. (2009), who define the Green Economy program of O*NET that groups green jobs in:

- existing occupations that are expected to experience significant employment growth due to the greening of the economy (Green Demand);
- existing occupations that are expected to undergo significant changes in terms of task content (Green Enhanced Skills);
- new occupations that emerge as a response to specific needs of the green economy (Green Emerging).

Although this approach has potential in mapping green occupations, it essentially conceives ex-ante the green economy as a net creator of new jobs, because it is understood to be *a new growing sector*. What if, however, green applies not only to products/sectors, but also to processes? What if the greening of a given existing sector implies essentially efficiency-enhancing processes, reducing input absorption, and labour thereof? What if green technologies are not linked only to a new emerging economy but rather to more efficient green processes? And, what happens to sectoral employment if production processes become at the same time net saver of emissions and labour inputs? After all, to a larger extent, green processes are essentially productivity-enhancing processes, and as such they might incorporate a LS trait (Rosenberg, 1976; Dosi, 1988; Tunzelmann, 1995).

Adopting an evolutionary perspective on technological change and drawing upon Montobbio et al. (2022b), we want to assess the extent to which patents meant to mitigate climate change present an explicit LS content. Patents indeed represent a viable proxy of codified technological knowledge within firms and thus constitute a powerful tool to understand the rate and direction of innovative activities (Pavitt, 1985). The willingness to focus on LS heuristics comes from the possibility to challenge the very notion of green skills, since we question the existence of processes, and ensuing human skills, uniquely connected to the development of green products; in so doing, we focus on the greenness of processes, rather than products. Moreover our research question, differently from incumbent studies devoted to understanding the development of new occupations within sectors, concerns the extent to which existing efforts in developing green technologies are coupled with efforts in reducing labour inputs, via efficiency-enhancing processes. Should the twin transition present a limited capability in the development of labour-friendly products, unfolding especially towards labour-saving green processes, we shall argue that LS effects may prevail in the realisation of less polluting new processes, also requiring less manpower.

3 Data description

The technological dataset is represented by USPTO data. We first retrieved from PatentsView[4] all granted patents published between 1976 and 2021 which contain at least one CPC code of either Class `Y02` or Subclass `Y04S`, which are intended to encompass green technologies (Veefkind et al., 2012; Angelucci et al., 2018). A total of 475,597 patents are found in this step, whose temporal evolution is depicted in figure 1. Given this set, we will devise and apply a procedure to identify LS patents therein.

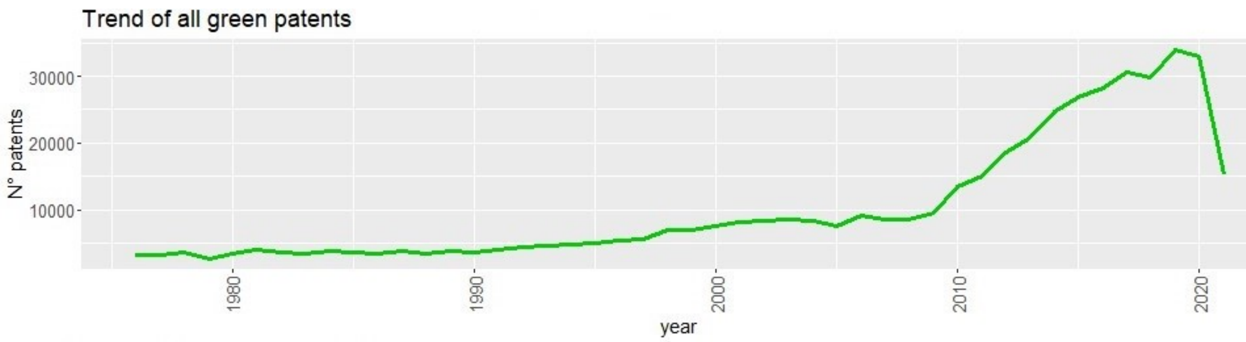


Figure 1: Number of green patents per year

The second dataset moves from technology, to sectors, to state level labour markets, in order to evaluate the sectoral (industrial) penetration of LS technologies and their employment impacts at state level in the US. In particular, we leverage upon:

- **IPC-NACE concordance table**: in order to match each single patent to a given industrial sector (NACE)[5], we adopt the concordance table provided by the European Patent Office, at 6-digits.[6]

[4]PatentsView.
[5]Nomenclature generale des Activites economiques dans les Communautes europeennes.
[6]NACE & IPC concordance.

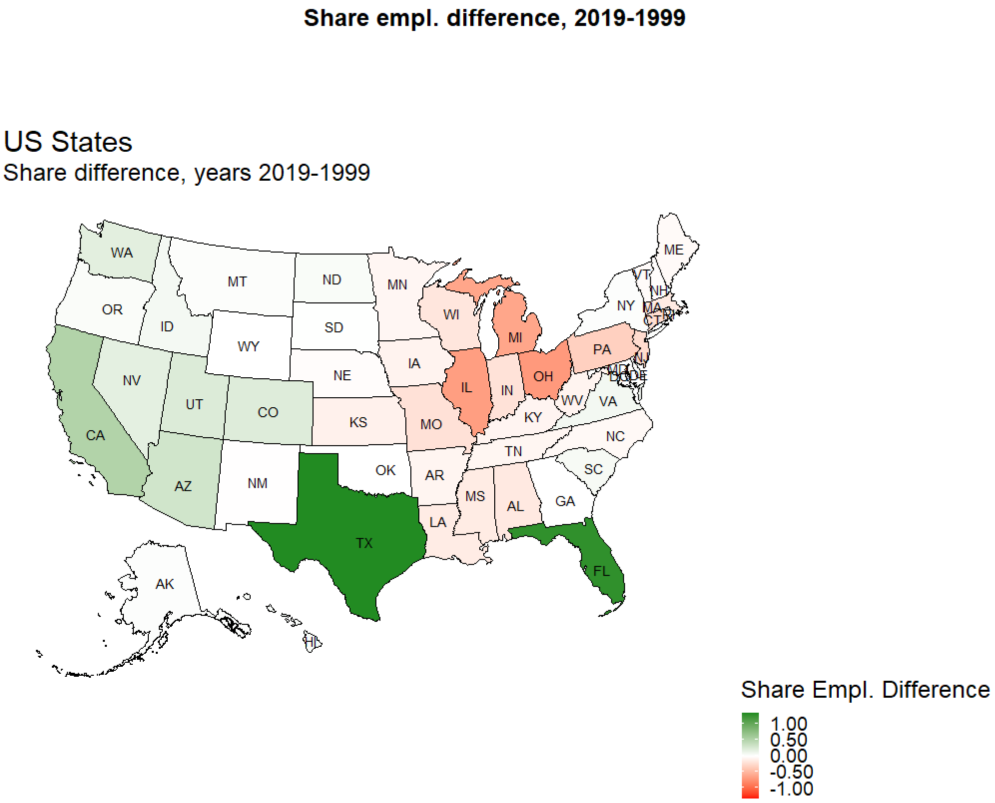


Figure 2: Employment share difference 2019 vs. 1999

- **Sectoral employment data (US)**: for sectoral employment data we adopt the Statistics for US Business (SUSB), collected and made available by the United States Census Bureau.[7] We retrieve state level data for three years, in particular 1999-2009-2019. Data are shown in figure 2, plotting employment share change over twenty years. Remarkable differences emerge already at this stage in terms of the geography of employment with net loosing and net gaining states. The map signals the inner structural change in terms of manufacturing (Rust Belt) versus the coastal and southern areas linked to both high end (California and Washington) and low end (Texas and Florida) services.

- **NAICS-NACE concordance table**: tables of concordance at 6-digits are made available by the European Commission. Last table conversion is available for 2017[8].

- **NAICS tables**: over time there were different rounds of NAICS codes. Tables are made available by the United States Census Bureau.[9].

In section B of the Appendix we provide more details upon the concordance tables and the overall strategy adopted to match the data.

[7]Details, descriptions and limitations of dataset can be found online at census.gov/methodology.
[8]NAICS-NACE concordance.
[9]Concordance tables link.

4 Methodology

In order to identify LS heuristics inside green patents, we delve into natural language processing techniques. Our approach relies on the textual analysis of the entire patent document text and constitutes an advancement with respect to contributions that solely focus on patent titles and abstracts (see e.g. Webb, 2020). Textual search is becoming increasingly applied in economics, while in other social sciences the sophistication and usage of NLP algorithm is spurring (Do et al., 2022). However, while some methodological improvements and empirical application to specific sectors have been provided (Hain et al., 2020; Hain et al., 2022), few contributions, at the best of our knowledge, are comparable to ours in terms of methodology. In particular Mann and Püttmann (2021) establish a training sample with manual validation and then they extend the identification through machine learning algorithm to classify the larger population of identified patents. Dechezleprêtre et al. (2019) construct a composite identification strategy which involves both patent classification and multiple keywords search.

Our empirical strategy entails first, the focus on semantic procedure rather than simple keywords search; second, the implementation of natural language processing techniques to validate ex-post identified patents to limit false positives. Third, the construction of alternative semantic constructs to enlarge the scope of identification of true positive. Therefore, with our multi-steps approach we are able to isolate specific LS heuristics inside green patents texts, representing a (conservative) picture of patents involved in the twin transition, because explicitly embedding LS traits.

In subsection 4.1 we briefly describe the approach that leads to the identification of potential LS green patents, combining the match of Y02-Y04S CPC patents (Veefkind et al., 2012; Angelucci et al., 2018) and the textual approach developed in Montobbio et al. (2022b). However, we face a deep methodological challenge that is the identification of true labour saving patents. Indeed, the validation procedure on which we leverage upon, described in subsection 4.2, represents an advancement and novelty in the analysis of patent texts, with only few exceptions in Meindl and Mendonça (2021), where the authors rely, as we do, on the spaCy library (Honnibal and Montani, 2017), applied to Industry 4.0 patents. In figure 3 we present the synthetic flow chart of our methodology.

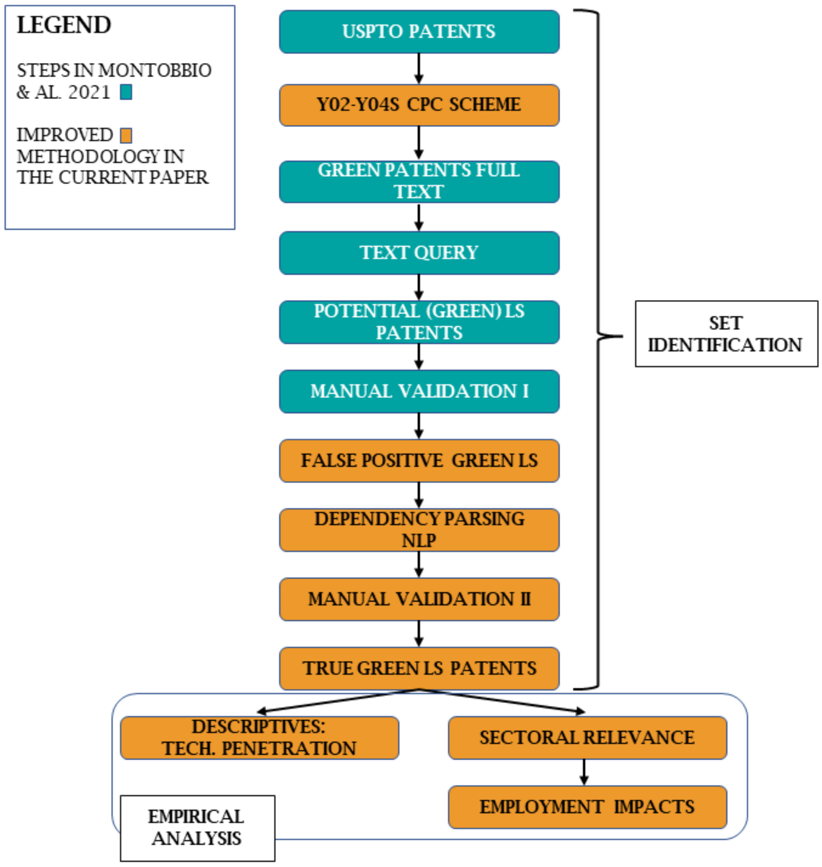


Figure 3: Flow chart of the methodology

4.1 Identification of the patent set: potential LS green patents

We firstly retrieved, from the USPTO Bulk Data Storage System, all the patents from 1976 to 2021 containing at least one CPC code of the Y02-Y04S category, that is all 475,597 green patents. Then, in order to analyse the potential LS effects embedded into these patents, we adopted the textual algorithm and procedure described in Montobbio et al. (2022b). While we refer the reader to the paper for a full description of the methodology, in figure 4 we show the structure of triplets to identify the LS traits. The algorithm entails to consider the full-text of each green patent, after tokenisation, removal of stop words and stemming.[10] The algorithm looks for the joint occurrence of a triplet, which differently from trigrams does not require word adjacency of words identified by the list, and assigns a flag as *potential LS* patent if one sentence contains at least one of the $k \times j \times m$ triplets.

$$\underbrace{\begin{bmatrix} \text{`pred.1'} \\ \text{`pred.2'} \\ \text{`...'} \\ \text{`pred.k'} \end{bmatrix}}_{\text{predicate}} \times \underbrace{\begin{bmatrix} \text{`obj.1'} \\ \text{`obj.2'} \\ \text{`...'} \\ \text{`obj.j'} \end{bmatrix}}_{\text{object}} \times \underbrace{\begin{bmatrix} \text{`attr.1'} \\ \text{`attr.2'} \\ \text{`...'} \\ \text{`attr.m'} \end{bmatrix}}_{\text{attribute}}$$

Figure 4: Structure of the *labour-saving* textual query

The preliminary steps of the identification strategy allow us to highlight 10430 patents as potential green LS patents. A first manual validation of a 10% sample of data shows however a high level of false positives. Two examples here below:

*"The sHASEGPs or a soluble **human** hyaluronidase domain thereof or pharmaceutically acceptable derivatives can be prepared with carriers that protect the soluble glycoprotein against rapid **elimin**ation from the body, such as **time** release formulations or coatings"* US9562223B2

[10] In order: tokenisation is obtained by means of a punctuation of regular expressions; the list of stop-words was specified by the nltk Python library; from the same Python package we implement an advanced version of Porter (1980).

*"[...] for **human** consumption, soybean cultivar can be used to produce edible protein ingredients which offer a healthier, **less expens**ive **replac**ement for animal protein in meats, as well as in dairy-type products" US8076545B2*

As it is possible to notice, despite the words highlighted belong broadly to the semantic area of LS technologies, the meaning of such phrases does not, showing two powerful examples of false LS green patents. Thus, in order to isolate the true positive green LS patents, we initially perform an exploration based on the exclusion of specific CPC associated to pharmaceutical and biotech technologies, also in line with Mann and Püttmann (2021) that classify most of chemical and pharmaceutical patents as non-automation patents. However, the distribution of false positives in our dataset does not show a strong concentration into specific technological classes, but rather a wide dispersion around CPC codes. Therefore, in order to correctly pinpoint true positives, we move to the analysis of the semantic structure of statements, applying a dependency parsing algorithms.

4.2 Identification of true LS green patents: dependency parsing analysis

Dependency parsing belongs to a family of grammar formalisms, whereby *"[...], phrasal constituents and phrase-structure rules do not play a direct role. Instead, the syntactic structure of a sentence is described solely in terms of the words (or lemmas) in a sentence and an associated set of directed binary grammatical relations that hold among the words"* (Jurafsky and Martin, 2020: pag. 280). In order to achieve such result we rely on spaCy[11] (Honnibal and Montani, 2017), a specialised NLP library which leverages on neural networks which is increasingly used both in industrial and academic applications, e.g. Meindl and Mendonça (2021). The model represents texts through dependency parsing, which reconstructs the grammar relationship[12] between words and the overall hierarchical structure of sentences. This allows to perform sophisticated text queries which go beyond the simple co-occurrence of keywords. One of the very interesting features of the spaCy algorithm is the possibility to deploy such grammatical structures through graphical representation of dependency trees, as we do in figures 13, 14, 15, 16, 17, 18, where the arrows describe both the dependency type (grammatical nature of a word in a specific phrase, e.g. adjectival modifiers) and the relationship between words (how a noun is related to another one through an adjective or verb, for example). The usage of ex-ante defined grammatical structure enables to rule out false positives like the examples provided in section 4.1.

Two "ingredients" are therefore necessary to allow the algorithm to work: a dictionary of target keywords and a specified dependency structure. We extend the keyword lists used in Montobbio et al. (2022b), as reported in figure 5.

Together with the dictionary of target keywords, we require the relevant sentences of potential LS green

[11]spaCy.

[12]A more technical discussion on the various dependency types is offered in De Marneffe et al. (2014).

$$\underbrace{\begin{bmatrix}\texttt{`reduce'}\\ \texttt{`reduction'}\\ \texttt{`replace'}\\ \texttt{`replacement'}\\ \texttt{`eliminate'}\\ \texttt{`elimination'}\\ \texttt{`save'}\\ \texttt{`saving'}\\ \texttt{`lower'}\\ \texttt{`substitute'}\\ \texttt{`substitution'}\\ \texttt{`automate'}\\ \texttt{`automation'}\\ \texttt{`efficiency'}\\ \texttt{`less'}\\ \texttt{`minimize'}\\ \texttt{`minimal'}\\ \texttt{`minimization'}\\ \texttt{`without'}\\ \texttt{`not require'}\end{bmatrix}}_{\text{predicate}} \times \underbrace{\begin{bmatrix}\texttt{`worker'}\\ \texttt{`labor'}\\ \texttt{`human'}\\ \texttt{`employee'}\\ \texttt{`manpower'}\\ \texttt{`job'}\\ \texttt{`labour'}\\ \texttt{`manual'}\end{bmatrix}}_{\text{object}} \times \underbrace{\begin{bmatrix}\texttt{`cost'}\\ \texttt{`expenditure'}\\ \texttt{`expense'}\\ \texttt{`hour'}\\ \texttt{`intensity'}\\ \texttt{`task'}\\ \texttt{`time'}\\ \texttt{`skill'}\\ \texttt{`effort'}\\ \texttt{`intervention'}\\ \texttt{`amount'}\end{bmatrix}}_{\text{attribute}}$$

Figure 5: Dictionary lists for the dependency parsing model

patents (previously identified based on the triplets in figure 5) to exhibit either of the following dependency structures:

- Baseline pattern: predicate $\rightarrow$ attribute $\rightarrow$ object
- Pattern I: predicate $\leftarrow$ attribute $\rightarrow$ object
- Pattern II: predicate $\rightarrow$ object $\rightarrow$ attribute
- Pattern III: object $\rightarrow$ attribute $\rightarrow$ predicate
- Pattern IV: object $\rightarrow$ predicate $\rightarrow$ attribute

With the symbol $\rightarrow$ or $\leftarrow$ we indicate the relationship between keywords and their semantic order within the dependency tree. According to the baseline pattern (predicate $\rightarrow$ attribute $\rightarrow$ object), we ask the algorithm to look for a semantic structure that starts with a predicate which connects to an attribute, and further to an object. In order to read dependency trees, presented in Appendix A, we recall that the algorithm assigns two types of tagging. Firstly, for each single word in the patent text, the algorithm assigns a tag identifying the core part-of-speech category, that is, its grammar definition (e.g. a noun, an adjective, a verb, etc.). Such tag is called Universal Point of Speech (POS) tagging and is provided below each word represented in the figures. The second tag instead characterises the grammatical relationship between words (dependency) and is depicted along the connecting arrows; for instance, "attr" means that the word upon which the arrows land is an attribute with respect to the the word from which the arrow starts. In particular, the algorithm adopts the Universal Dependency (Dep) in terms of nomenclature. The POS tagging list is described in more detail in Appendix E where we also provide the appropriate source to the interested reader concerning the Dep nomenclature and description.

4.3 A snapshot of true LS green patents subset

Repeated random samples were hand-validated in order to gain an insight on the magnitude of algorithm accuracy, which shows a level superior to 85%. To study the underlying technological content of the subset of true LS green patents, we map the CPC associated to climate change-related patents, distinguishing between LS and non-LS ones. Table 1 and 2 show, respectively, the top 20 CPC at 4 digits[13] in terms of frequency for LS and non-LS green patents. The frequency is computed as the number of times a certain code is specified across all patents and takes into account the fact that the same 4-digit code may appear more than once in each patent. In grey we show common CPC between LS and non-LS green patents, while coloured CPC were the ones uniquely appearing in each list (green for the LS patents, orange for the non-LS ones): the majority of them are shared among the two lists, a sign of pervasiveness of green technology both between LS and not LS patents and also the lack of specific LS applications in some circumscribed domains.

In table 2 we notice specific non-LS CPC concerning medical/therapeutic areas (A61P, A61K), automotive with a focus on combustion engine (F02D, F01N, B60W) and chemistry (B01J, C01B). The LS green patents instead show more heterogeneous areas in terms of CPC, which include control systems (G05B), data processing for administrative purposes (G60Q), heating system (H05B, F24S), telephonic communication (H04M), in line with the digital content of this set. The very fact the LS patents are not restricted to a selected CPC but are rather widespread signals that LS heuristics are not a restricted phenomenon. However, this pervasiveness also highlights the fact that to be identified such hueristics require a more complex procedure rather than only ex-ante focusing on specific CPC.

Notably, the algorithm correctly pinpoints technological applications related to human treatments as false positive (non-LS); in addition, climate-change related innovation in automotive are concentrated in non-LS patents rather than in LS ones. The latter evidence might hint at the fact that more innovative efforts in the automotive sector are currently focusing on product, rather than process innovation, such as the electric engine, the production of batteries and their internal components.

[13] In Appendix D we present similar tables but using full digits codes.

Top 20 TRUE LS			
CPC 4 dig.	Rank	Freq	**Description**
H04L	1	3323	TRANSMISSION OF DIGITAL INFORMATION, e.g. TELEGRAPHIC COMMUNICATION
H04W	2	2966	WIRELESS COMMUNICATION NETWORKS
G05B	3	2132	CONTROL OR REGULATING SYSTEMS IN GENERAL; FUNCTIONAL ELEMENTS OF SUCH SYSTEMS; MONITORING OR TESTING ARRANGEMENTS FOR SUCH SYSTEMS OR ELEMENTS
B29C	4	1994	SHAPING OR JOINING OF PLASTICS; SHAPING OF MATERIAL IN A PLASTIC STATE, NOT OTHERWISE PROVIDED FOR; AFTER-TREATMENT OF THE SHAPED PRODUCTS, e.g. REPAIRING
G06Q	5	1744	DATA PROCESSING SYSTEMS OR METHODS, SPECIALLY ADAPTED FOR ADMINISTRATIVE, COMMERCIAL, FINANCIAL, MANAGERIAL, SUPERVISORY OR FORECASTING PURPOSES; SYSTEMS OR METHODS SPECIALLY ADAPTED FOR ADMINISTRATIVE, COMMERCIAL, FINANCIAL, MANAGERIAL, SUPERVISORY OR FORECASTING PURPOSES, NOT OTHERWISE PROVIDED FOR
Y02E	6	1526	REDUCTION OF GREENHOUSE GAS [GHG] EMISSIONS, RELATED TO ENERGY GENERATION, TRANSMISSION OR DISTRIBUTION
G06F	7	1333	ELECTRIC DIGITAL DATA PROCESSING
Y02P	8	1304	CLIMATE CHANGE MITIGATION TECHNOLOGIES IN THE PRODUCTION OR PROCESSING OF GOODS
Y10T	9	1154	TECHNICAL SUBJECTS COVERED BY FORMER US CLASSIFICATION
H01L	10	1097	SEMICONDUCTOR DEVICES; ELECTRIC SOLID STATE DEVICES NOT OTHERWISE PROVIDED FOR
F24S	11	1074	SOLAR HEAT COLLECTORS; SOLAR HEAT SYSTEMS
H01M	12	1053	PROCESSES OR MEANS, e.g. BATTERIES, FOR THE DIRECT CONVERSION OF CHEMICAL ENERGY INTO ELECTRICAL ENERGY
Y02T	13	1003	CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TO TRANSPORTATION
H02J	14	971	CIRCUIT ARRANGEMENTS OR SYSTEMS FOR SUPPLYING OR DISTRIBUTING ELECTRIC POWER; SYSTEMS FOR STORING ELECTRIC ENERGY
Y02B	15	782	CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TO BUILDINGS, e.g. HOUSING, HOUSE APPLIANCES OR RELATED END-USER APPLICATIONS
B60L	16	764	PROPULSION OF ELECTRICALLY-PROPELLED VEHICLES [...]; SUPPLYING ELECTRIC POWER FOR AUXILIARY EQUIPMENT OF ELECTRICALLY-PROPELLED VEHICLES[...]; ELECTRODYNAMIC BRAKE SYSTEMS FOR VEHICLES IN GENERAL[...]; MAGNETIC SUSPENSION OR LEVITATION FOR VEHICLES; MONITORING OPERATING VARIABLES OF ELECTRICALLY-PROPELLED VEHICLES; ELECTRIC SAFETY DEVICES FOR ELECTRICALLY-PROPELLED VEHICLES
C02F	17	745	TREATMENT OF WATER, WASTE WATER, SEWAGE, OR SLUDGE
B01D	18	652	SEPARATION
H04M	19	624	TELEPHONIC COMMUNICATION
H05B	20	619	ELECTRIC HEATING; ELECTRIC LIGHT SOURCES NOT OTHERWISE PROVIDED FOR; CIRCUIT ARRANGEMENTS FOR ELECTRIC LIGHT SOURCES, IN GENERAL

Table 1: TOP 20 CPC, true LS patents

Top 20 FALSE LS			
CPC 4 dig	Rank	Freq	**Description**
H01M	1	403572	PROCESSES OR MEANS, e.g. BATTERIES, FOR THE DIRECT CONVERSION OF CHEMICAL ENERGY INTO ELECTRICAL ENERGY
H01L	2	234719	SEMICONDUCTOR DEVICES; ELECTRIC SOLID STATE DEVICES NOT OTHERWISE PROVIDED FOR
Y02E	3	202559	REDUCTION OF GREENHOUSE GAS [GHG] EMISSIONS, RELATED TO ENERGY GENERATION, TRANSMISSION OR DISTRIBUTION
Y02T	4	182576	CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TO TRANSPORTATION
G06F	5	146749	ELECTRIC DIGITAL DATA PROCESSING
B60L	6	137688	PROPULSION OF ELECTRICALLY-PROPELLED VEHICLES; SUPPLYING ELECTRIC POWER FOR AUXILIARY EQUIPMENT OF ELECTRICALLY-PROPELLED VEHICLES ; ELECTRODYNAMIC BRAKE SYSTEMS FOR VEHICLES IN GENERAL; MAGNETIC SUSPENSION OR LEVITATION FOR VEHICLES; MONITORING OPERATING VARIABLES OF ELECTRICALLY-PROPELLED VEHICLES; ELECTRIC SAFETY DEVICES FOR ELECTRICALLY-PROPELLED VEHICLES
B01J	7	115215	CHEMICAL OR PHYSICAL PROCESSES, e.g. CATALYSIS OR COLLOID CHEMISTRY; THEIR RELEVANT APPARATUS
B01D	8	114560	SEPARATION
H04W	9	114510	WIRELESS COMMUNICATION NETWORKS
Y02P	10	112623	CLIMATE CHANGE MITIGATION TECHNOLOGIES IN THE PRODUCTION OR PROCESSING OF GOODS
H02J	11	105938	CIRCUIT ARRANGEMENTS OR SYSTEMS FOR SUPPLYING OR DISTRIBUTING ELECTRIC POWER; SYSTEMS FOR STORING ELECTRIC ENERGY
F02D	12	104515	CONTROLLING COMBUSTION ENGINES
A61P	13	94281	SPECIFIC THERAPEUTIC ACTIVITY OF CHEMICAL COMPOUNDS OR MEDICINAL PREPARATIONS
H04L	14	91749	TRANSMISSION OF DIGITAL INFORMATION, e.g. TELEGRAPHIC COMMUNICATION
F01N	15	86433	GAS-FLOW SILENCERS OR EXHAUST APPARATUS FOR MACHINES OR ENGINES IN GENERAL; GAS-FLOW SILENCERS OR EXHAUST APPARATUS FOR INTERNAL COMBUSTION ENGINES
B60W	16	84747	CONJOINT CONTROL OF VEHICLE SUB-UNITS OF DIFFERENT TYPE OR DIFFERENT FUNCTION; CONTROL SYSTEMS SPECIALLY ADAPTED FOR HYBRID VEHICLES; ROAD VEHICLE DRIVE CONTROL SYSTEMS FOR PURPOSES NOT RELATED TO THE CONTROL OF A PARTICULAR SUB-UNIT
Y10T	17	78931	TECHNICAL SUBJECTS COVERED BY FORMER US CLASSIFICATION
A61K	18	72087	PREPARATIONS FOR MEDICAL, DENTAL, OR TOILET PURPOSES
C01B	19	59927	NON-METALLIC ELEMENTS; COMPOUNDS THEREOF
Y02B	20	57705	CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TO BUILDINGS, e.g. HOUSING, HOUSE APPLIANCES OR RELATED END-USER APPLICATIONS

Table 2: TOP 20 CPC, non-LS patents

5 Results

In the following, we present our results. We start discussing the temporal evolution and technological composition of both LS and non-LS patents in subsection 5.1. Then, we move into the analysis of sectoral penetration and employment impacts on US labour markets in subsection 5.2.

5.1 Temporal evolution and technological composition

This section presents the temporal evolution and technological composition of the overall number of patents retrieved from the different steps of the algorithm, as shown in table 3. First, applying the procedure of identification of LS patents developed in Montobbio et al. (2022b), we end up with potential 10430 patents out of 475597 (about 2.29%). Second, applying the validation procedure via dependency parsing described in section 4.2, we end up with 3901 true LS patents (about 0.8% of all green patents).

In terms of temporal trends, shown in figure 6, both sets present a steady increase up to 2008-2010, with a fierce acceleration onward. In the third plot of figure 6, we report the relative share of LS patents over time, in order to compare the relative trend of the two categories: notably, until the '80, LS green patents had a steeper increase with respect to the overall green patents while, since then, the share fluctuates around 0.8% across the years.

In tables 4 we show the Y02-Y04S CPC tags frequency associated to the identified patents. The overall bulk of patents is classified into nine scopes of application, according to the USPTO definitions.[14] The relative distributions across categories appear coherent between LS and non-LS patents, with some notable exception. For instance, while Energy and CSSD ("Climate Storage Sequestration or Disposal") are respectively the most and less frequent CPC tag in both groups, Digital shows a lower level of frequency for LS green patents then the non-LS one. On the contrary, Smart grids CPC is relatively more prevalent in LS green patents with respect to the total of patents. Although all patents are inside the climate-change related domain, these technologies are characterised by different "stages of life-cycle", namely, some are infant technologies while some others are mature ones. Different stages in life-cycle might manifest in heterogeneous trends over time.

Type of patents	**Number of patent**
All green patents	475597
Potential LS green patents	10430
True LS green patents	3901
False LS green patents	6529

Table 3: N° of patents along identification procedures

In order to draw a patent composition analysis, beyond tags, we rely on prevalent CPC associated to each patent, which is the first CPC that appears for each granted patent. The bulk of green technologies are concentrated in what we labeled "complementary green" patents, meaning that the primary associated

[14]The official labels and descriptions can be found at Espacenet.

Tag	Comparison		LS		NO LS	
	Rank LS	**Rank non-LS**	**Freq**	**Rel. Freq. (%)**	**Freq**	**Rel. Freq. (%)**
Energy	1	1	1526	24.13	202559	28.73
Transportation	3	2	1003	15.86	182576	25.9
Products & processes	2	3	1304	20.62	112623	15.98
Building	4	4	782	12.37	57705	8.19
Digital	8	5	338	5.34	50529	7.17
Adaptation	5	6	539	8.52	46340	6.57
Waste	7	7	354	5.6	23612	3.35
Smart grids	6	8	463	7.32	22610	3.21
CSSD	9	9	15	0.24	6430	0.91

Table 4: Tag composition, LS vs. non-LS green patents

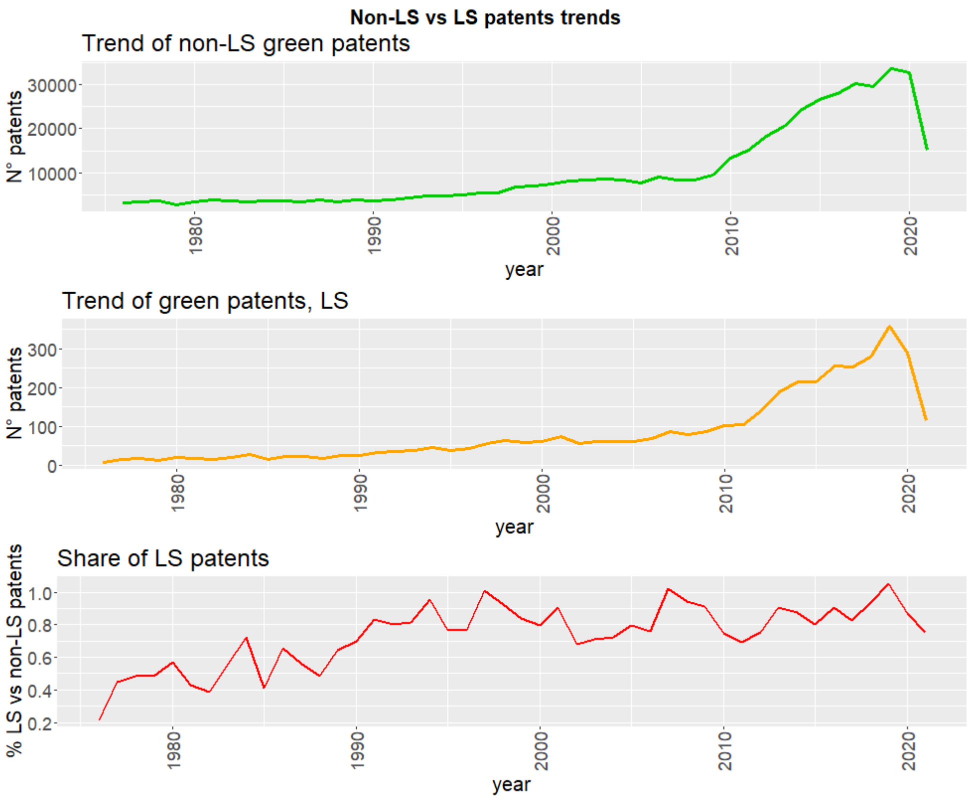


Figure 6: Patents' trend: all patents vs. LS

CPC tag is different from the Y02-Y04S classification. Therefore, restricting the analysis to those patents which present a green-tag as the first element (figure 7), in the non-LS set, energy, transportation, product & process, digital and building patents do represent the majority of them with a steep increase since 2008 onward. In particular, digital and building patents show notable growth rates, while product & process patents manifest a slowdown. The other technological classes instead have a more sluggish trend. Focusing on LS green patents (bottom part of figure 7), the low numbers of patents per year generate more volatile trends. In this subset we highlight the higher relative importance of product & process patents and the lower importance of transportation ones.

From the temporal and composition analyses of green patents it emerges that, green technologies are quite heterogeneous in themselves, with some technological domains almost disregarded, as waste. Energy and transportation are the workhorse but they generally come as secondary scope and use. Indeed, it appears that green technologies are more complementary rather than uniquely sourced, considering that the largest fraction of patents do not report a green CPC as primary code. With reference to LS green technology, the ubiquity across domains is evident, and even the temporal trend, although sluggish, is increasing.

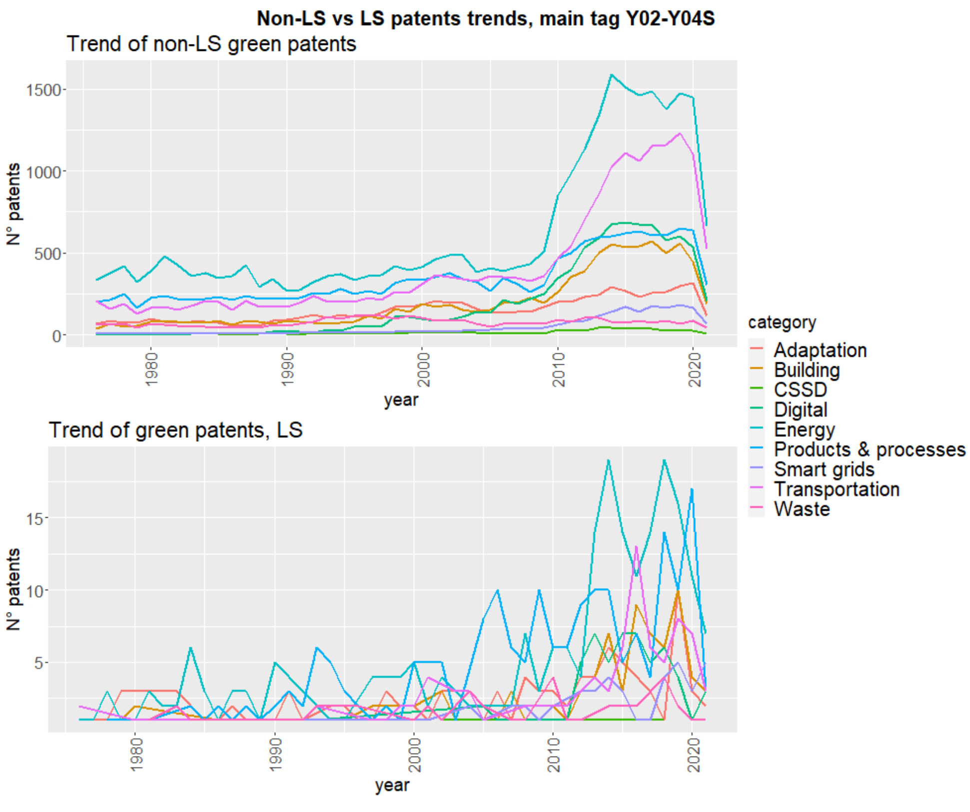


Figure 7: Patents' trend: non-LS patents vs. LS, only Y02-Y04S tags

Given the different life cycle of the underlying technologies, in figure 8 we present the temporal dynamics of each specific share of technology. A higher temporal volatility derives from LS green patents as shown in the bottom part of the figure where both LS and non-LS patents are considered together. If we focus on the comparison between LS and non-LS green patents, we distinguish some specific patterns. Starting from non-LS patents we observe stable levels of patent shares in transportation and energy, even if the latter shows a small decrease in the last decades. Adaptation and waste present a similar pattern, with a relative increase until mid-1990s, followed by a constant decrease. Digital, smart grids and building patents instead have acquired higher shares over time, in particular digital technologies. LS shares present remarkable differences with respect to the non-LS ones: for instance, adaptation and product & process look to be more relevant than respectively non-LS patents and, on the contrary, transportation LS patents do not have a very high importance. Waste, digital patents and smart grids present similar (yet more volatile) patterns than non-LS green patents. Smart grids technologies are relatively more important in the case of LS patents than in non-LS one. It is however important to highlight that the bulk of innovative efforts are concentrated toward the complementary green patents, as it is possible to notice in the right side of figure 8.

LS patents have shown to be a tiny fraction. However, what if new emerging green patents progressively are born embedding LS heuristics? How does dynamically change the importance of LS heuristics? Temporally weighted growth rates, where the weight is based on the lagged annual share of patents, allow to account for the underlying behaviour. This measure is useful to capture different life cycle stages of green technologies and it is constructed as follows, for both LS and non-LS patent sets, respectively defined by the superscript H={LS; non-LS} to indicate the underlying heuristics:

1. For each green technological category i, we define the annual growth rate of patents:

$$growth_{i,t}^{(H)} = \frac{n_{i,t}^{(H)} - n_{i,t-1}^{(H)}}{n_{i,t-1}^{(H)}} \tag{1}$$

2. For each single year, we compute the share of each specific technology, with respect to total green patents:

$$Share_{i,t}^{(H)} = \frac{n\_patents\_category_{i,t}^{(H)}}{\sum_{i=1}^{k} n\_patents\_category_{i,t}^{(H)}} \tag{2}$$

3. The weighted growth for each category is defined as the product between the yearly growth rate by single category and its lagged year share, namely:

$$Weighted\ Growth_{i,t}^{(H)} = growth_{i,t}^{(H)} * Share_{i,t-1}^{(H)} \tag{3}$$

4. Finally, we apply a five years rolling average to smooth the trends and we compute the cumulative growth.

The results offer a clear accounting of the technological pervasiveness of LS heuristics in the development of new green technologies across the majority of domains. Figure 9 represents the cumulative weighted growth, distinguished by the Y02-Y04S technological tag schemes and dividing between LS and non-LS patents. With the exception of CSSD technologies, digital and partly smart grids, the weighted cumula-

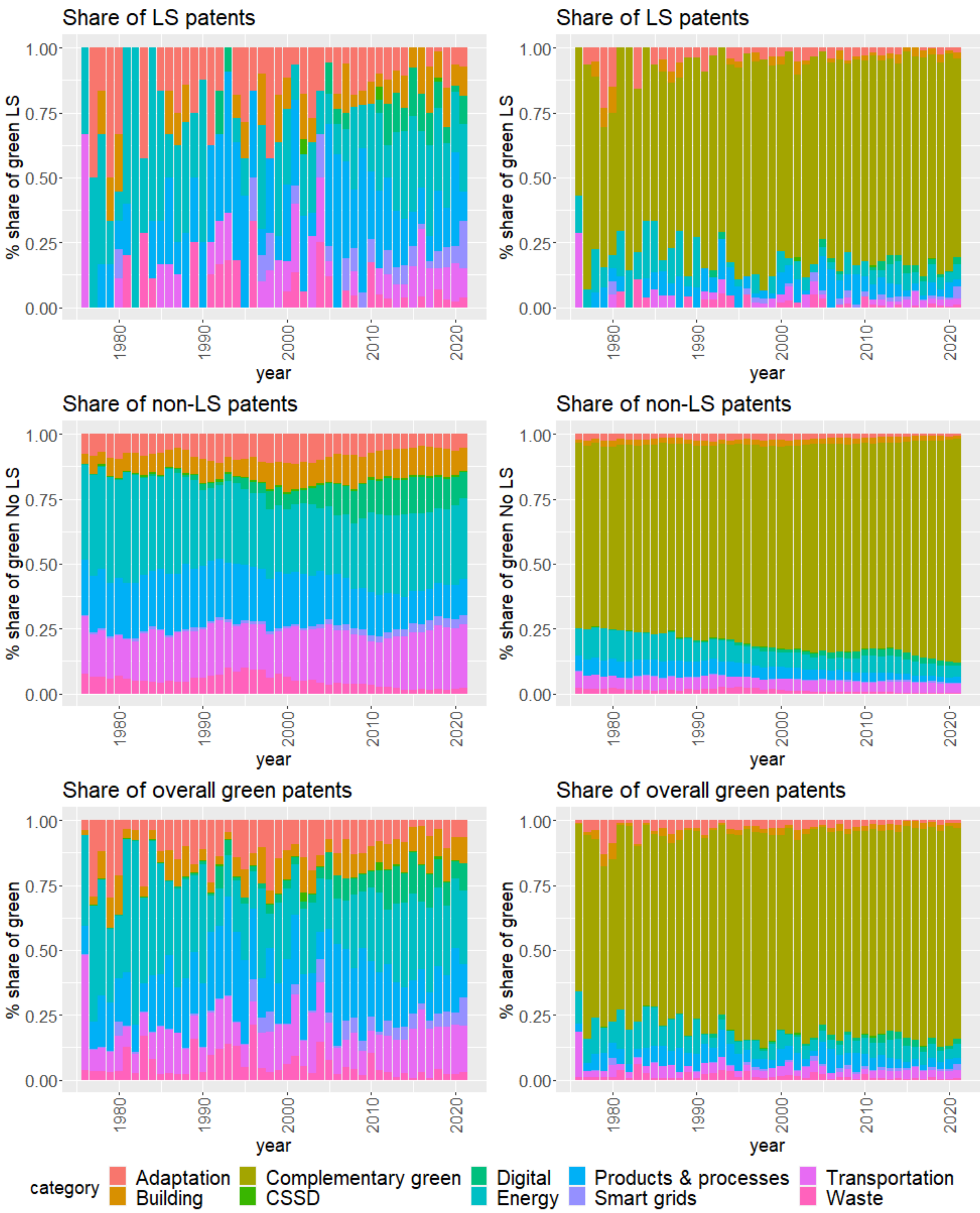


Figure 8: Share of patents, LS vs. non-LS

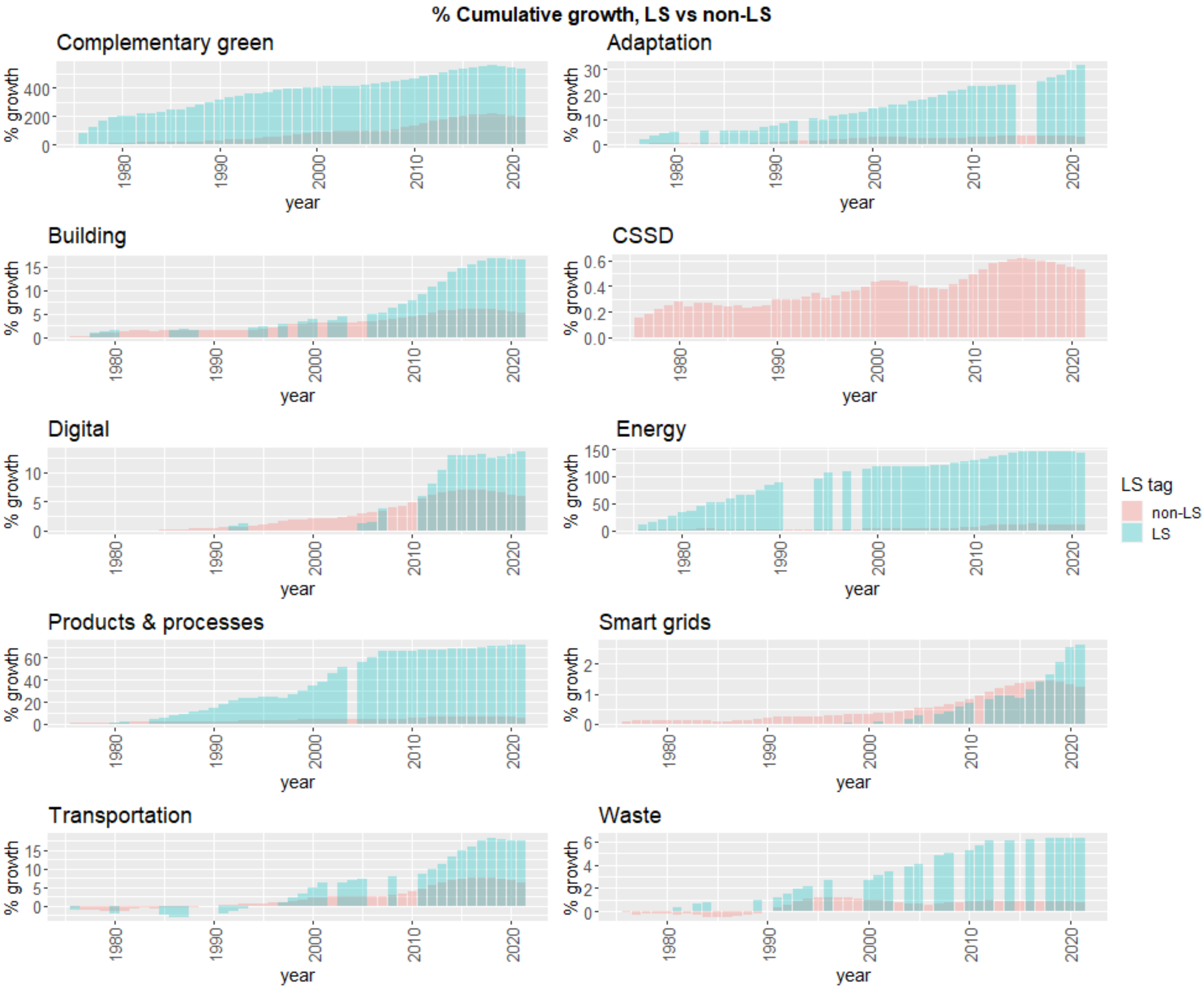


Figure 9: % Cumulative weighted growth, LS vs. non-LS by green component

tive growths of all other remaining technologies are much higher for the LS set of patents with respect to the non-LS one. The result highlights that green patents progressively arrive embedding labour-saving heuristics, therefore, although relatively low in number, LS patents show to be concentrated in newly developed technologies. Indeed, such heuristics, if at a first approximation might appear a secondary concern, once concentrating in newly emerging technologies, and during the last period, might reveal to be far more relevant.

This evidence raises the question about the job creation capacity of the green paradigm. While new jobs might arrive because of new product creation, technological upgrading and greenifying technologies might embrace a labour saving content potentially able to expel labour force. In the next subsection we are going to go deeper into the link between technological penetration of LS green technologies, at the geographical level, and the related employment growth trends.

5.2 Penetration of green LS technologies on employment growth

We now move to analyse the nexus between employment growth and LS heuristics' penetration, at the sectoral and state level in the US. With this scope, we first link patents to sectors, and then, controlling for state level sectoral composition of employment, we link sectoral penetration of technologies to employment dynamics.

The first step entails mapping the relevance of LS patents among sectors: to accomplish the task, we use the concordance table between IPCs and NACE sectors provided by the European Patent Office which can also be adopted for CPC.[15] The concordance table allows to map each patent i with its NACE codes (thereafter, sectoral codes) and weights associated, such that for each patent i the sectoral weights sum up to one.

After associating CPC to sectoral codes, we build a *sectoral penetration index* (normalised between zero and one) associated to each sector j, which represents the overall sectoral exposure to each of the two patent sets (LS vs. non-LS). For instance, a level close to one of such indicator in the non-LS patent set characterises a sector not exposed to LS technological penetration, therefore in which the majority of non-LS patents are concentrated.

To construct an indicator of sectoral penetration we firstly build an identifier composed by each patent ID and sectoral weights associated, and we then uniquely select the rows based on such identifier. Afterwards, for each j sectoral code, we build the average, according to the following procedure:

$$Avg.\ Sectoral\ penetration_j^{(H)} = \frac{\sum_{i=1}^{n_j} Sector\ weight_{i,j}^{(H)}}{n_j^{(H)}} \tag{4}$$

where n_j is the number of patents in sector j, and $Sector\ weight_{i,j}$ is the weight associated to patent i in

[15]Excel concordance tables and metadata can be found atNACE & IPC concordance table.

sector j.

Such procedure however equally weights more and less prevalent CPC, thus presenting potential biases. In order to make this measure meaningful for intersectoral comparison, we weight the average sectoral penetration for the patent share in each sector:

$$Sectoral\ patent\ share_j^{(H)} = \frac{n_j^{(H)}}{\sum_{j=1}^{k} n_j^{(H)}} \tag{5}$$

These sectoral shares assume higher values for sectors characterised by higher patent intensity of a specific CPC, and vice-versa. Finally, we build a Sectoral penetration index which allows for weighted comparisons between sectors:

$$Sectoral\ penetration\ index_j^{(H)} = Avg.\ Sectoral\ penetration_j^{(H)} * Sectoral\ patent\ share_j^{(H)} \tag{6}$$

The attributions of technological penetration to sectors are shown in figure 10 and in the third column of tables 5 and 6 for non-LS and LS patents respectively. If we consider the top twenty more exposed sectors, it is possible to notice a high level of overlapping between the two sets, in line with the results on CPC prevalence. Restricting the attention to the top five sectors, we highlight the relative less importance of automotive/transportation sector in LS technologies with respect to the non-LS green patents. In fact, in the 1$^{\text{st}}$ and 3$^{\text{rd}}$ positions in the set of non-LS green patents we find respectively sector 27.2 ("Manufacture of batteries and accumulators") and 29.1 ("Manufacture of motor vehicles"), while the same industrial sectors are in the 7$^{\text{th}}$ and 13$^{\text{th}}$ positions for the LS patents set. Such results are indeed not surprising and signal a lack of specific concentration of LS heuristics in some specific sectors, and therefore the ensuing pervasiveness all but limited to specific sectors/technology. Indeed, the pervasiveness of LS heuristics might be considered a potential warning of the embedded labour-saving effects. Recall that the majority of the identified patents are "complementary green", therefore the presence of sectors not strictly related to green products should not come as a surprise. In addition, this evidence suggests the relevance of interpreting the usage of the underlying patented technologies also in terms of green processes.

We now move from sectoral to employment penetration at the geographical level. In the US, the sectoral employment composition deeply differs across states. Therefore, we can compare state by state the change in employment shares in more versus less exposed sectors to LS patents, where the exposure is measured by the sectoral penetration index presented in tables 5 and 6, and then aggregated at the state level. Indeed, the analysis is meant to understand the extent to which states that do present an employment composition in sectors more exposed to LS patents record a different dynamics vis-à-vis states whose employment composition is less concentrated in sectors exposed to LS technologies.

We employ SUSB data from the United States Census Bureau and the concordance table (reported in the fourth column of tables 5 and 6), linking therefore each sector to the level of employment in 2019. In order

Table 5: TOP 20 sectoral codes, non-LS green patents

Sectoral code	Sectoral code description	Sectoral penetration index	EMPL sector (absolute values)	Rank Emp
27,2	Manufacture of batteries and accumulators	1	449911	3
29,1	Manufacture of motor vehicles	0,892928	453170	2
26,3	Manufacture of communication equipment	0,564028	77046	22
26,2	Manufacture of computers and peripheral equipment	0,527544	39505	30
28,29	Manufacture of other general-purpose machinery n.e.c.	0,354456	640863	1
28,11	Manufacture of engines and turbines, except aircraft, vehicle and cycle engines	0,30955	137586	16
27,9	Manufacture of other electrical equipment	0,234689	319205	7
27,12	Manufacture of electricity distribution and control apparatus	0,231457	68926	24
26,51	Manufacture of instruments and appliances for measuring, testing and navigation	0,168164	4685	5
25,3	Manufacture of steam generators, except central heating hot water boilers	0,154136	61724	25
28,3	Manufacture of agricultural and forestry machinery	0,079312	181638	13
28,25	Manufacture of non-domestic cooling and ventilation equipment	0,074527	301782	8
32,5	Manufacture of medical and dental instruments and supplies	0,056016	426927	4
28,23	Manufacture of office machinery and equipment (except computers and peripheral equipment)	0,055596	78223	21
27,33	Manufacture of wiring devices	0,054702	40303	29
27,4	Manufacture of electric lighting equipment	0,048128	103335	18
28,99	Manufacture of other special-purpose machinery n.e.c.	0,037413	230878	9
42,91	Construction of water projects	0,030119	72980	23
26,7	Manufacture of optical instruments and photographic equipment	0,024796	19414	34
26,4	Manufacture of consumer electronics	0,023935	15533	35

Table 6: TOP 20 sectoral codes, LS green patents

Sectoral code	Sectoral code description	Sectoral penetration index	EMPL sector (absolute values)	Rank Empl
26,2	Manufacture of computers and peripheral equipment	1	39505	28
26,3	Manufacture of communication equipment	0,875315519	77046	21
26,51	Manufacture of instruments and appliances for measuring, testing and navigation	0,556333898	373743	5
27,2	Manufacture of batteries and accumulators	0,433191707	449911	3
28,3	Manufacture of agricultural and forestry machinery	0,393741431	181638	13
28,29	Manufacture of other general-purpose machinery n.e.c.	0,367832238	640863	1
29,1	Manufacture of motor vehicles	0,330763648	453170	2
27,12	Manufacture of electricity distribution and control apparatus	0,234918604	68926	23
25,3	Manufacture of steam generators, except central heating hot water boilers	0,173308396	61724	24
28,23	Manufacture of office machinery and equipment (except computers and peripheral equipment)	0,147993413	78223	20
27,4	Manufacture of electric lighting equipment	0,14345348	103335	18
28,11	Manufacture of engines and turbines, except aircraft, vehicle and cycle engines	0,114136857	137586	16
28,25	Manufacture of non-domestic cooling and ventilation equipment	0,102149119	301782	8
27,9	Manufacture of other electrical equipment	0,098981025	319205	7
32,5	Manufacture of medical and dental instruments and supplies	0,09572913	426927	4
27,33	Manufacture of wiring devices	0,08913564	40303	27
28,99	Manufacture of other special-purpose machinery n.e.c.	0,072506004	230878	9
42,91	Construction of water projects	0,053415885	72980	22
20,2	Manufacture of pesticides and other agrochemical products	0,031378868	27457	30
28,22	Manufacture of lifting and handling equipment	0,028898217	353483	6

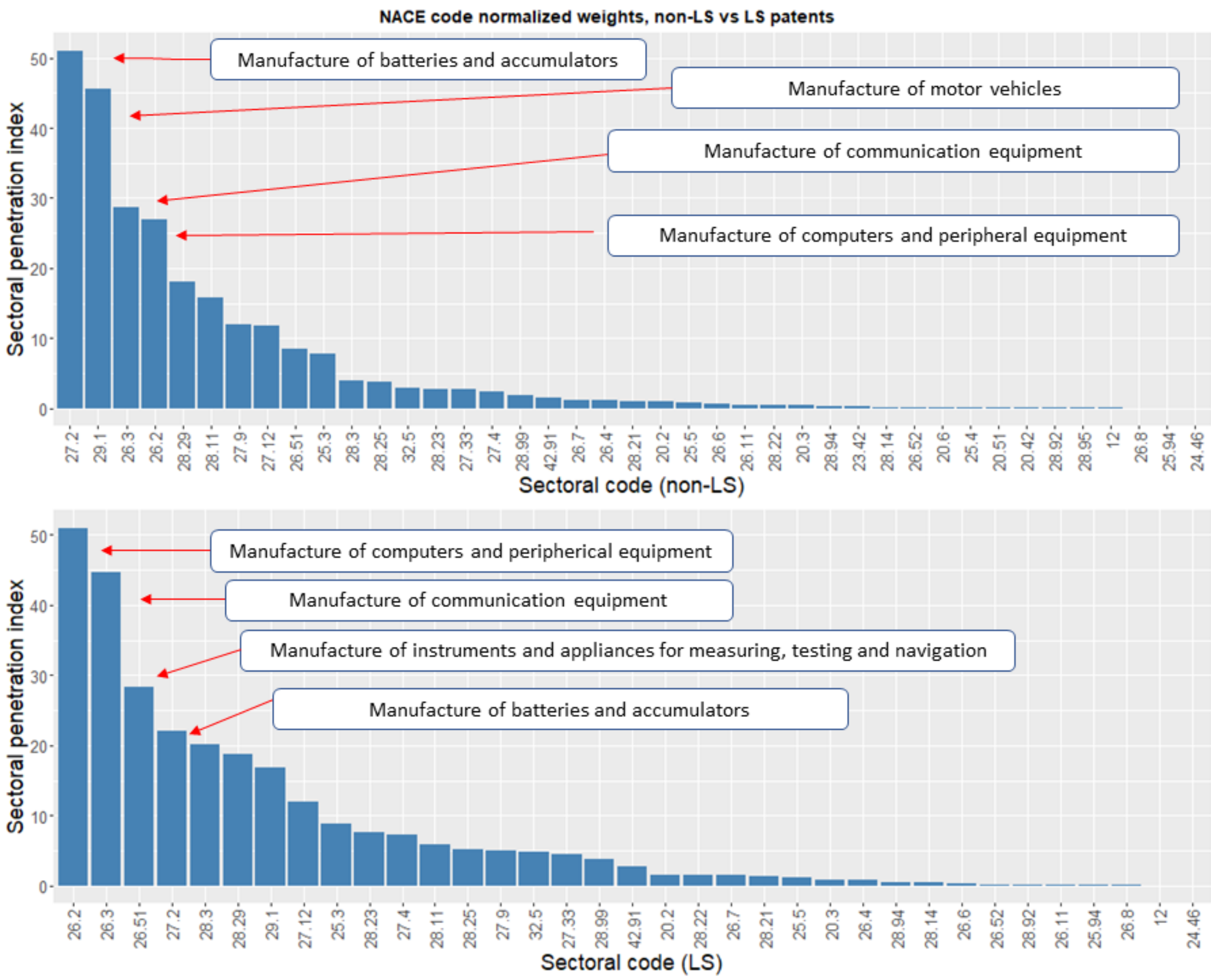


Figure 10: Most prevalent sectoral codes, normalised values

to map if and the extent to which the sectoral penetration of LS patents has an impact on employment share growth, and eventually a differentiated one with respect to non-LS patents, we set up a quantile regression analysis, conducted at the state-sectoral level.

Given the lack of sectoral employment data at the state level, we build state level employment weights to impute the share of each sector (available at federal level) for each state, that is:

$$state\,weight_{i,t}^{(H)} = \frac{Employment\,state_{i,t}^{(H)}}{Federal\,Employment_{t}^{(H)}} \tag{7}$$

with $i = 1, ..., 50$ that represents state dummies, and $t = \{1999, 2009, 2019\}$ years considered. For both sets we then compute the share of employment in state i, in sector j, at time t:

$$Total\,share_{i,j,t}^{(H)} = \frac{Federal\,Employment\,sector_{j,t}^{(H)}}{Federal\,Employment_{t}^{(H)}} * state\,weight_{i,t}^{(H)} \tag{8}$$

We finally calculate ten-year growth rates of the implied sectoral employment shares in order to have a relatively long time span to capture any structural change process. We perform the following quantile regression estimation for each of the two sectoral penetration indices of patents (as usual for both LS and non-LS sets), including both a linear and a quadratic term:

$$\begin{aligned} Empl\ Share\ growth_{\theta,i,j}^{(\tau)(H)} =& \beta_1^{(\tau)(H)} Sectoral\ penetration\ index_{i,j} + \beta_2^{(\tau)(H)} Sectoral\ penetration\ index_{i,j}^2 \\ &+ \alpha_i^{(\tau)(H)} + \epsilon_{i,j,\theta}^{(\tau)(H)} \end{aligned} \tag{9}$$

with $\tau = 0.5$, indicating the proportion of the population having scores below the quantile at τ; θ represents the interval periods considered (2019 vs. 2009, 2009 vs. 1999, 2019 vs. 1999), for which we perform distinct regressions; i=1,...,50, indicating one of the US states; j the sector. State level fixed effects were included to account for geographical heterogeneity and to counterbalance the fixed within sectoral composition across states (see equation 7). Finally, the inclusion of the quadratic term allows for a non linear relationship and a more flexible regression estimate.

In table 7 we present the results of the regression exercise and in figure 11 we plot the intensity of the coefficients along the distribution of the sectoral penetration index. Both OLS and quantile regression at the median estimates are reported, while each point represents the sectoral-state share of employment in each year of the estimation period.

Results are in line with our expectations. Along the distribution of the sectoral penetration index, between zero and one, the coefficients show an overall negative concave relationship between the relevance of LS patents across sectors and employment share growth, while the opposite holds for the non-LS patents set. Notably, the quadratic relationship signals the existence of a non linear-threshold behaviour that is in general around a penetration index of 0.5. Indeed, sectors/states with high sectoral penetration are not the majority of observed points but, whenever the index of penetration is high, negative effects for sectoral employment share growth are strong.

Our dependent variable informs about changes in the structural composition across sectors, being both services and manufacturing included in the overall employment. A robustness test is conducted in Appendix C, restricting the analysis to manufacturing shares only, confirming the result: whenever the within-manufacturing sectoral employment is more exposed to LS green patents, employment shares in that sector decline.

According to these results, first, our identification methodology of LS heuristics looks to be ex-post validated given that sectors which are exposed to a large number of LS green patents do present decreasing share growth, and therefore manifest LS effects on employment. Notably, the opposite result holds for non-LS patents. In addition, despite the low number of LS patents, results are significant and robust to the inclusion of a different dependent variable, that is employment share across manufacturing. Finally, the nexus between the unfolding of green technology upon employment strengthens in the last decade (2009-2019), while in the first (1999-2009) does not show any significant result in either set (LS vs. non-LS).

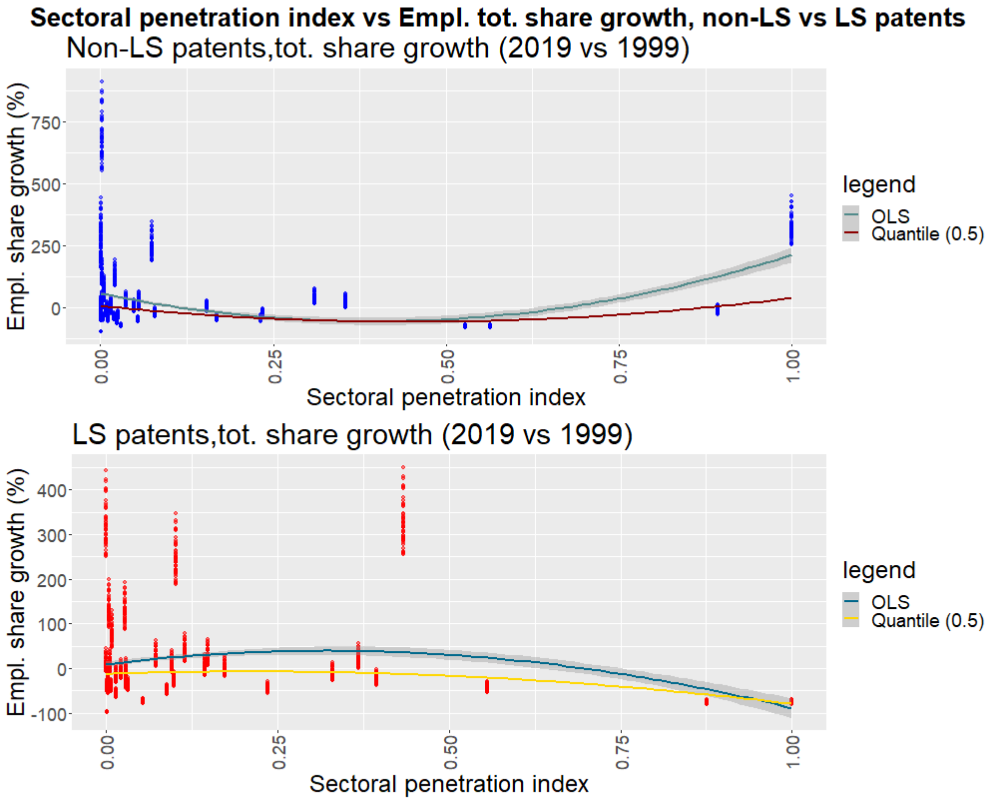
In order to account for geographical heterogeneity, in figure 12 we plot the estimated beta coefficients at the state level. From the maps, the depressed Rust Belt area and the inner Wyoming and Missisipi states notably stand out. These states record negative employment growth shares in sectors more exposed to LS green technologies, given the differences between the two maps (left-hand side vs. right-hand side). Winning states are instead located into the east and southern areas. Notably, states with higher sectoral exposure to non-LS patents record positive employment share growth, as shown in the left-hand side figure. Another striking result is that the distinction between LS and non-LS technologies becomes relevant in the second decade (2009-2019), while the dynamics in the decade 1999-2009 shows an almost overlapping pattern between the two sets of technologies. Considering that the sectoral penetration index of LS technologies is a time-invariant indicator, the intensification of the effects over time can only be attributed to a retardation, time to display-effect, of both LS and non-LS technologies over time.

In a nutshell, the most exposed sector to LS patents is manufacturing of computers and peripheral equipments, while, at the opposite, the most exposed sector to non-LS technologies is manufacturing of batteries and accumulators. The two most exposed sectors are a clear distinct example of the different effects that process vs. product innovation can exert on employment even in the green segment.

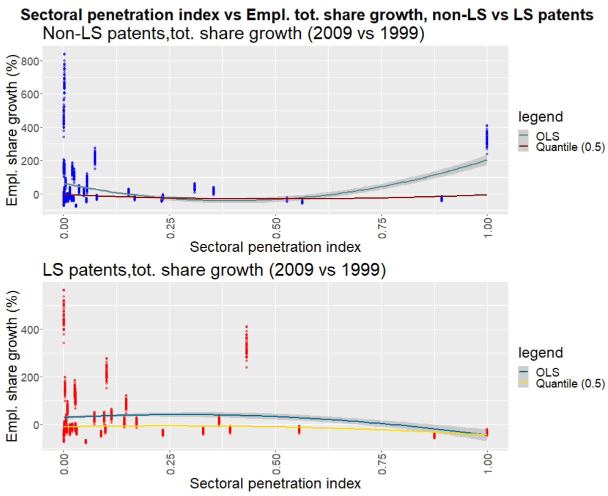
Table 7: Quantile regression results (0.5), LS vs. non-LS patents

	NACE Employment total share growth, LS vs. non-LS patents					
	2019vs1999	2009vs1999 LS	2019vs2009	2019vs1999	2009vs1999 non-LS	2019vs2009
	(1)	(2)	(3)	(4)	(5)	(6)
Sectoral penetration index	0.244288**	−0.065534	0.719042***	−3.076947***	−0.741661*	−0.210307***
	(0.120790)	(0.230282)	(0.066488)	(0.480215)	(0.413646)	(0.054767)
Sectoral penetration index$\hat{2}$	−1.051315***	−0.188366	−1.262192***	3.158888***	0.613342	0.275940***
	(0.140848)	(0.227818)	(0.073499)	(1.168288)	(1.193818)	(0.054104)
Observations	1,785	1,785	1,785	2,091	2,091	2,091

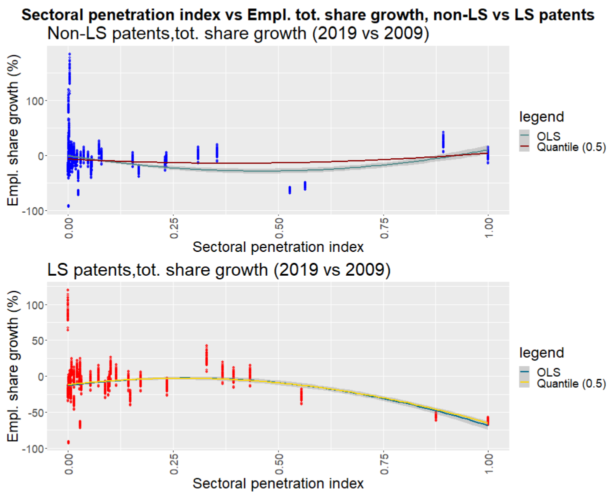
Note: *p<0.1; **p<0.05; ***p<0.01



(a) Total share growth, 2019 vs. 1999

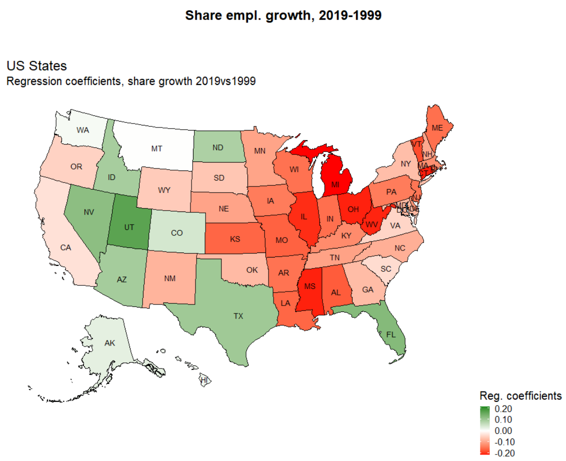


(b) Total share growth, 2009 vs. 1999

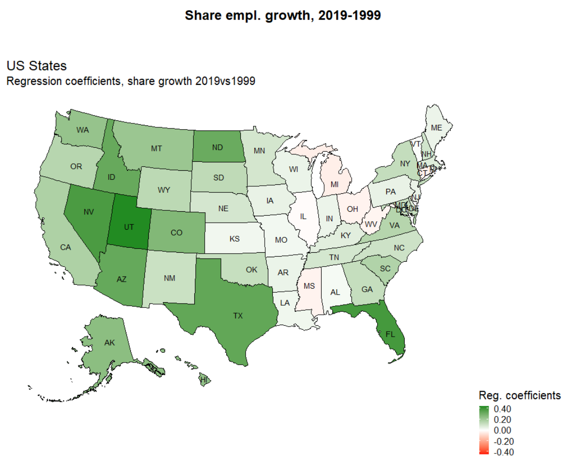


(c) Total share growth, 2019 vs. 2009

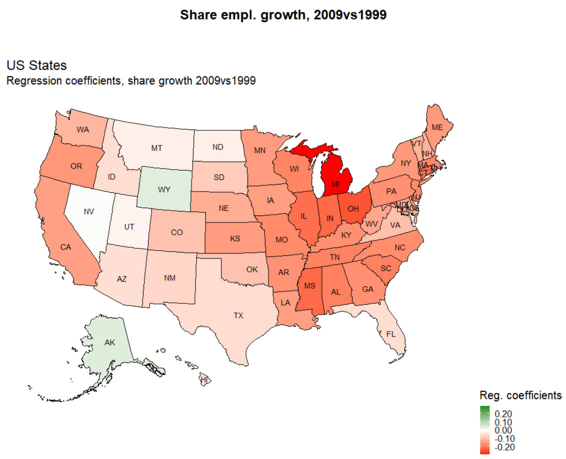
Figure 11: Regression plots



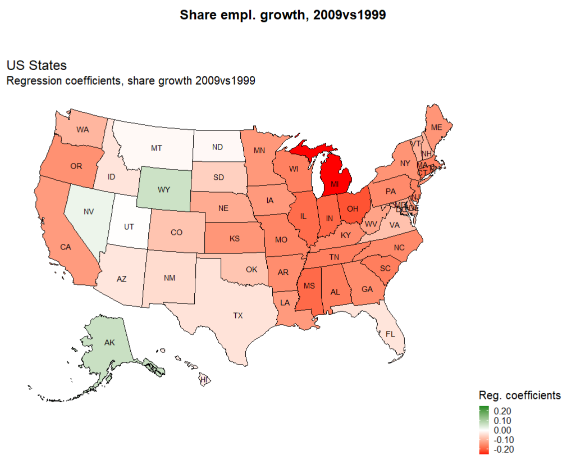
(a) LS patents, 2019 vs. 1999



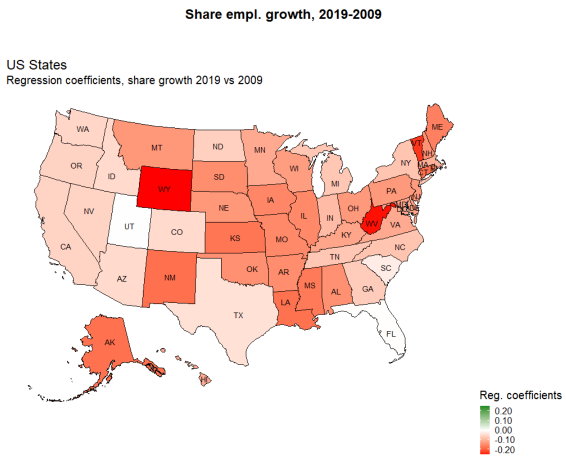
(b) Non-LS patents, 2019 vs. 1999



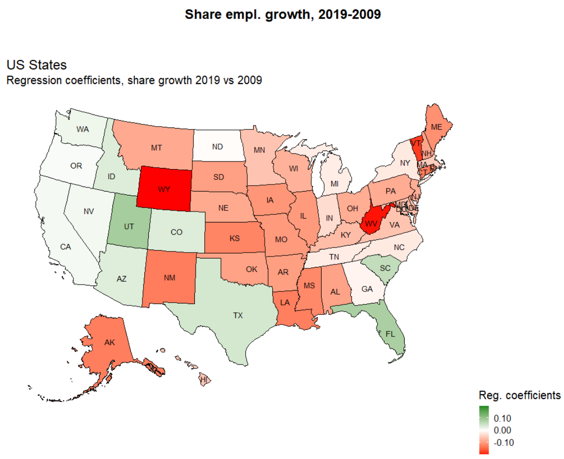
(c) LS patents, 2009 vs. 1999



(d) Non-LS patents, 2009 vs. 1999



(e) LS patents, 2019 vs. 2009



(f) Non-LS patents, 2019 vs. 2009

Figure 12: State level coefficients, empl. share growth

6 Conclusions

Climate change urges for policy actions: green transition and digitalisation/automation efforts are seen as pivotal and are currently under the lens of practitioners and scholars. However, while recognised as part of a common transition, defined as twin transition, green and digitalisation/automation impacts are often treated separately, especially with respect to labour markets. The literature currently tends to emphasise the job creation effects of the green paradigm (IRENA and ILO, 2021) and the job-destruction effects of automation, digitalisation, and more recently AI (Montobbio et al., 2022a). However, there is still no clear understanding of the couple dynamics between green technologies, which should support the green transition, and labour-saving heuristics embedded in innovative green efforts.

Given the extant literature, the first contribution of this paper is to detect the existence of LS heuristics in climate change mitigation/adaptation patents, therefore to link these two countervailing forces upon labour market restructuring. To empirically accomplish the task, we delve into the analysis of textual contents of patents relying on Natural Language Processing (NLP) techniques. In addition, we adopt a semantic analysis validation method, dependency parsing, allowing to produce quite restrictive but reliable results. The methodological advancement in identifying LS green patents represents our second contribution.

We then construct a direct measure of sectoral technological penetration linking patents, distinguished into LS and non-LS, and connected to sectors via the patent-sector concordance table. The sectoral exposure allows then to move to state level labour markets, accounting for sectoral employment distributions and the net effects deriving from LS penetration. The construction of a direct measure of technological exposure linked to labour markets is an innovative advancement with respect to the literature on green jobs, which until now has adopted indirect measures of degrees of greeness at the task-level unit of analysis, inheriting the approach from the Routine-Biased Technical Change literature, the latter lacking of the actual measurement of technology in use. In addition, the green jobs literature, so far, has not delved into understanding green as a process but rather as a product or new emerging sector. The construction of a penetration index connecting technology-sectors-employment, together with the focus on green as a process, represents our third contribution.

According to our results, first, LS and non-LS patents do manifest differences in terms of technological composition (Y02-Y04S main tag): for instance the transport (Y02T) and digital sectors (Y02D) exert a less relevant role in LS green patents than in non-LS ones, while product & process (Y02P) and smart grids are relatively more important in LS patents. More remarkably, from the patterns of cumulative weighted growth in the two different sets, it emerges a clear evidence that LS patents are becoming progressively more pervasive in recent technological applications, considering that the measure accounts for the maturity stage of technologies.

Finally, we explore the effect of penetration of the two different sets of technology on employment (SUSB data) at the sector/state level. We study changes in the share of of employment along the last ten and twenty years. Our evidence shows that employment shares in sectors characterised by a higher exposure to LS (non-LS) technologies present an overall negative (positive) growth dynamics. Such results are robust when using manufacturing shares instead of overall employment, and provide further validation of the identification steps of LS heuristics embedded into green technologies. Remarkable state level heterogeneity emerges in the second decade, hinting at a time-to-display effect of technologies upon employment, with the Rust Belt area dramatically loosing in contrast with Texas and California gaining employment shares, with reference to non-LS patents.

The flexibility of the index is a such that it can be distinctively adopted to measure both labour expelling but also labour creating effects of the green paradigm. Our concern here has been towards labour expelling patterns, however labour creation effects might be studied as well. This represents a natural continuation of our work. In addition, cosine similarity measures of textual contents might be applied linking patents and tasks embedded into occupations, via O*NET, along the line of Montobbio et al. (2021). The latter would represent a second avenue of research. The study of the effects upon wage and functional inequality would be a third realm of investigation. Further extensions may include the distinction between product and process innovation, together with a more fine-grained decomposition of the geographical distribution.

There are however a number of limitations: patent data are not the unique proxy of technological innovation, and they do not exhaust the multidimensional aspects of innovation realm and scope. Secondly, our study deals with technological penetration but it does not address the actual adoption of such technologies by firms: our results, therefore, must be interpreted in terms of potential LS impacts and not as realised ones. Finally, other NLP methods are spurring and may be considered valuable alternatives, especially supervised machine learning techniques (Mann and Püttmann, 2021; Do et al., 2022).

References

Acemoglu, Daron and Pascual Restrepo (2020). "Robots and jobs: Evidence from US labor markets". In: *Journal of Political Economy* 128(6), pp. 2188–2244.

Angelucci, Stefano, F. Javier Hurtado-Albir, and Alessia Volpe (2018). "Supporting global initiatives on climate change: The EPO's "Y02-Y04S" tagging scheme". In: *World Patent Information* 54, S85–S92.

Barbieri, Laura, Chiara Mussida, Mariacristina Piva, and Marco Vivarelli (2019). *Testing the employment impact of automation, robots and AI: A survey and some methodological issues.* Tech. rep. IZA Discussion Paper No. 12612.

Barbieri, Nicolò, Davide Consoli, Lorenzo Napolitano, François Perruchas, Emanuele Pugliese, and Angelica Sbardella (2022). "Regional technological capabilities and green opportunities in Europe". In: *The Journal of Technology Transfer*, pp. 1–30.

Barbieri, Nicolò, Alberto Marzucchi, and Ugo Rizzo (2020). "Knowledge sources and impacts on subsequent inventions: Do green technologies differ from non-green ones?" In: *Research Policy* 49(2), p. 103901.

Barbieri, Nicolò, Alberto Marzucchi, and Ugo Rizzo (2021). *Green technologies, complementarities, and policy*. Tech. rep. SWPS 2021-08.

Bowen, Alex and Karlygash Kuralbayeva (2015). "Looking for green jobs: the impact of green growth on employment". In: *Grantham Research Institute Working Policy Report. London: London School of Economics and Political Science*.

Brynjolfsson, Erik and Andrew McAfee (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company.

Corradini, Carlo (2019). "Location determinants of green technological entry: evidence from European regions". In: *Small Business Economics* 52(4), pp. 845–858.

Curtis, E. Mark and Ioana Marinescu (2022). *Green Energy Jobs in the US: What Are They, and Where Are They?* Tech. rep. National Bureau of Economic Research.

D'Alessandro, Simone, André Cieplinski, Tiziano Distefano, and Kristofer Dittmer (2020). "Feasible alternatives to green growth". In: *Nature Sustainability* 3(4), pp. 329–335.

De Marneffe, Marie-Catherine, Timothy Dozat, Natalia Silveira, Katri Haverinen, Filip Ginter, Joakim Nivre, and Christopher D Manning (2014). "Universal Stanford dependencies: A cross-linguistic typology". In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pp. 4585–4592.

Dechezleprêtre, Antoine, David Hémous, Morten Olsen, and Carlo Zanella (2019). "Automating labor: evidence from firm-level patent data". In: *Available at SSRN 3508783*.

Dierdorff, Erich C., Jennifer J. Norton, Donald W. Drewes, Christina M. Kroustalis, David Rivkin, and Phil Lewis (2009). *Greening of the world of work: Implications for O* NET®-SOC and new and emerging occupations*.

Do, Salomé, Étienne Ollion, and Rubing Shen (2022). "The Augmented Social Scientist: Using Sequential Transfer Learning to Annotate Millions of Texts with Human-Level Accuracy". In: *Sociological Methods & Research*.

Dosi, Giovanni (1988). "Sources, procedures, and microeconomic effects of innovation". In: *Journal of Economic Literature*, pp. 1120–1171.

Frey, Carl Benedikt and Michael A Osborne (2017). "The future of employment: How susceptible are jobs to computerisation?" In: *Technological forecasting and social change* 114, pp. 254–280.

Hain, Daniel S., Roman Jurowetzki, Tobias Buchmann, and Patrick Wolf (2022). "A text-embedding-based approach to measuring patent-to-patent technological similarity". In: *Technological Forecasting and Social Change* 177, p. 121559.

Hain, Daniel S., Roman Jurowetzki, Primoz Konda, and Lars Oehler (2020). "From catching up to industrial leadership: towards an integrated market-technology perspective. An application of semantic patent-to-patent similarity in the wind and EV sector". In: *Industrial and Corporate Change* 29(5), pp. 1233–1255.

Hickel, Jason and Giorgos Kallis (2020). "Is green growth possible?" In: *New political economy* 25(4), pp. 469–486.

Honnibal, Matthew and Ines Montani (2017). "spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing". In: *To appear* 7(1), pp. 411–420.

International Monetary Fund (2022). "War Sets Back the Global Recovery". In: *World Economic Outlook, April*.

International Labour Office (2018). *World employment and social outlook 2018: Greening with jobs*. International Labour Organisation (ILO).

IRENA and ILO (2021). *Renewable Energy and Jobs – Annual Review 2021*. Tech. rep. International Renewable Energy Agency, International Labour Organization.

Jurafsky, Daniel and James Martin (Feb. 2020). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Third edition draft.

Mann, Katja and Lukas Püttmann (2021). "Benign effects of automation: New evidence from patent texts". In: *The Review of Economics and Statistics*, pp. 1–45.

Mazzei, Julia, Tommaso Rughi, and Maria Enrica Virgillito (2022). *Knowing brown and inventing green? Incremental and radical innovative activities in the automotive sector*. Tech. rep. LEM Working Paper Series.

Meindl, Benjamin and Joana Mendonça (2021). "Mapping Industry 4.0 Technologies: From Cyber-Physical Systems to Artificial Intelligence". In: *arXiv preprint arXiv:2111.14168*.

Montobbio, Fabio, Jacopo Staccioli, Maria Virgillito, and Marco Vivarelli (2022a). *The empirics of technology, employment and occupations: lessons learned and challenges ahead*. Tech. rep. IZA Discussion Paper No. 15731.

Montobbio, Fabio, Jacopo Staccioli, Maria Enrica Virgillito, and Marco Vivarelli (2022b). "Robots and the origin of their labour-saving impact". In: *Technological Forecasting and Social Change* 174, p. 121122.

Montobbio, Fabio, Jacopo Staccioli, Maria Enrica Virgillito, Marco Vivarelli, and Giovanni Vecchi (2021). *Labour-saving automation and occupational exposure: a text-similarity measure*. Tech. rep. LEM Working Paper Series.

Montresor, Sandro and Francesco Quatraro (2020). "Green technologies and Smart Specialisation Strategies: a European patent-based analysis of the intertwining of technological relatedness and key enabling technologies". In: *Regional Studies* 54(10), pp. 1354–1365.

Noble, David (1986). *Forces of production: A social history of industrial automation.* Oxford University Press. Oxford.

Nuvolari, Alessandro et al. (2002). "The 'Machine Breakers' and the industrial revolution". In: *Journal of European Economic History* 31(2), pp. 393–426.

Pavitt, Keith (1985). "Patent statistics as indicators of innovative activities: possibilities and problems". In: *Scientometrics* 7(1-2), pp. 77–99.

Porter, Martin F. (1980). "An algorithm for suffix stripping". In: *Program* 14(3), pp. 130–137.

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

Rosenberg, Nathan (1976). *Perspectives on technology.* Cambridge University Press.

Santoalha, Artur, Davide Consoli, and Fulvio Castellacci (2021). "Digital skills, relatedness and green diversification: A study of European regions". In: *Research Policy* 50(9), p. 104340.

Saussay, Aurélien, Misato Sato, Francesco Vona, and Layla O'Kane (2022). *Who's fit for the low-carbon transition?: Emerging skills and wage gaps in job and data.* Tech. rep. Fondazione Eni Enrico Mattei (FEEM).

Staccioli, Jacopo and Maria Enrica Virgillito (2021). "The present, past, and future of labor-saving technologies". In: *Handbook of Labor, Human Resources and Population Economics.* Springer, pp. 1–16.

Stefan, Muench, Stoermer Eckhard, Jensen Kathrine, Asikainen Tommi, Salvi Maurizio, and Scapolo Fabiana (2021). *Towards a green digital future.* Tech. rep. EUR 31075 EN, Publications Office of the European Union, JRC129319.

Tanner, Anne Nygaard (2014). "Regional branching reconsidered: Emergence of the fuel cell industry in European regions". In: *Economic Geography* 90(4), pp. 403–427.

Tunzelmann, G. Nicholas von (1995). "Time-saving technical change: the cotton industry in the English Industrial Revolution". In: *Explorations in Economic History* 32(1), pp. 1–27.

Unmüßig, Barbara, Wolfgang Sachs, and Thomas Fatheuer (2012). "Critique of the green economy". In: *Publication series on ecology. Berlin: Heinrich Böll Foundation.*

Van Vuuren, Detlef P., Elke Stehfest, David EHJ Gernaat, Jonathan C Doelman, Maarten Van den Berg, Mathijs Harmsen, Harmen Sytze de Boer, Lex F Bouwman, Vassilis Daioglou, Oreane Y Edelenbosch, et al. (2017). "Energy, land-use and greenhouse gas emissions trajectories under a green growth paradigm". In: *Global Environmental Change* 42, pp. 237–250.

Veefkind, Victor, J Hurtado-Albir, Stefano Angelucci, Konstantinos Karachalios, and Nikolaus Thumm (2012). "A new EPO classification scheme for climate change mitigation technologies". In: *World Patent Information* 34(2), pp. 106–111.

Vona, Francesco (2021). *Labour Markets and the Green Transition: a practitioner's guide to the task-based approach.* Tech. rep. Joint Research Centre (Seville site).

Vona, Francesco, Giovanni Marin, Davide Consoli, and David Popp (2018). "Environmental regulation and green skills: an empirical exploration". In: *Journal of the Association of Environmental and Resource Economists* 5(4), pp. 713–753.

Zuboff, Shoshana (1988). *In the age of the smart machine: The future of work and power*. Basic Books, Inc.

A Pattern examples

Pattern I: predicate ← attribute → objects

The first example is represented by patent US9062327B2 and the section of the text identified is the following:

"[...] *there is also less total operational expense, even assuming that the operational expense for a single ear corn harvester is the same as that for a single combine; and there is* ***less*** *total* ***labor expense.***" Figure

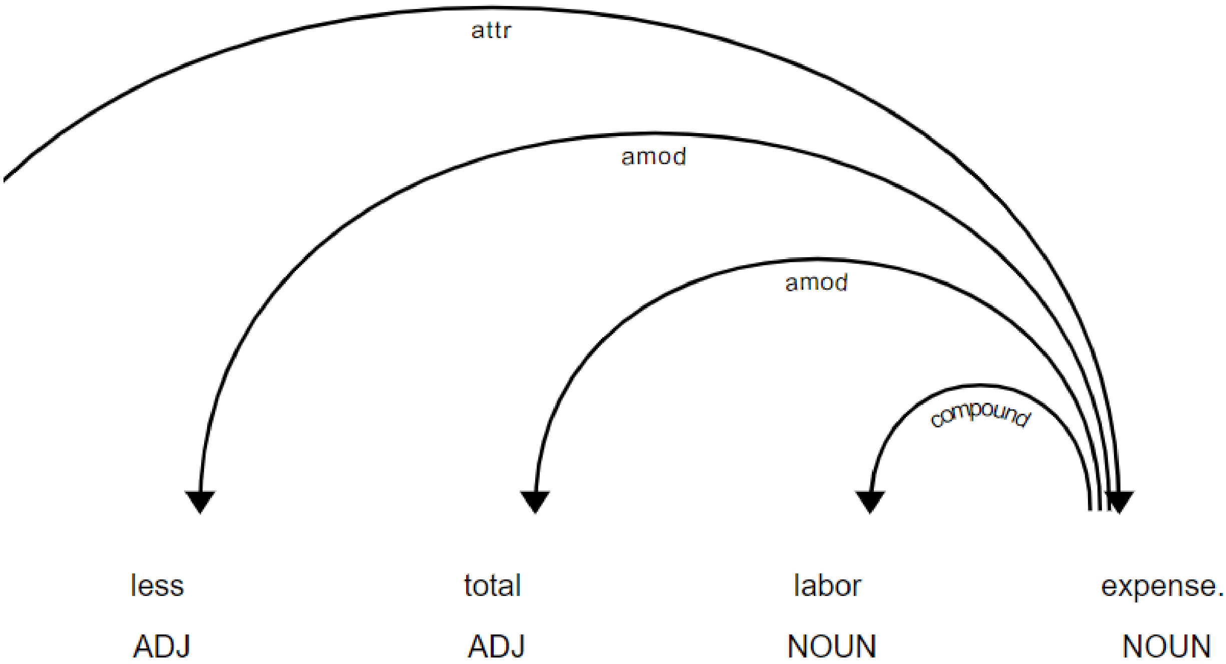


Figure 13: Example 1, pattern I

13 shows the portion of the dependency tree containing our target keywords. Here, we see that the word "expense" (NOUN) is connected to both "labor" (NOUN, "compound" of "expense") and "less" (ADJective, also identified as an adjective modifier, "amod"). The word "expense" belongs to the *attribute* list, "labor" to *object* and "less" to *predicate*.

The second example we provide is patent US10005267B1 and we focus on this section of text:

"[...] *this translates to* ***reduced*** *assembly* ***time*** *and* ***labor*** *for quicker and more cost effective manufacture.*"

Here the keyword is "time" which is connected to both the conjuction ("conj") "labor" and the adjective modifier "reduced". Again, "time" is in the *attribute* list, while the other two terms are respectively in the *object* and in the *predicate* lists.

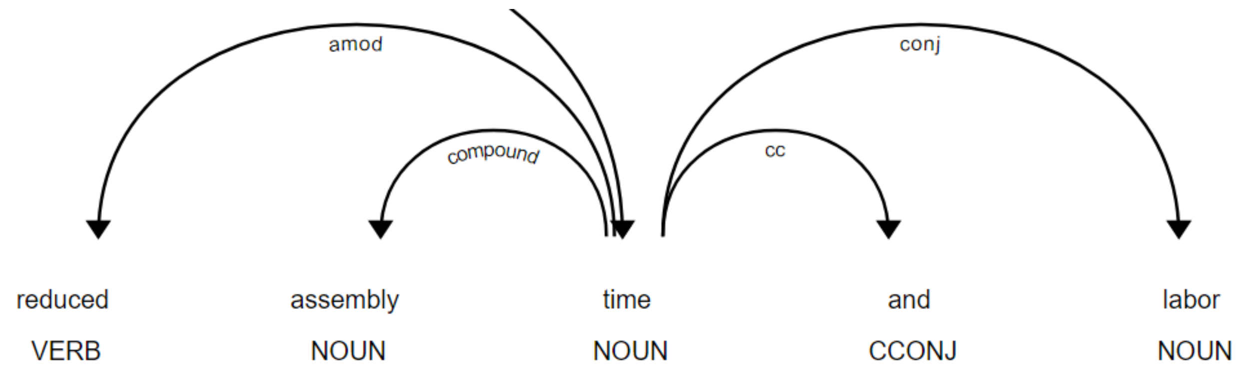


Figure 14: Example 2, pattern I

Pattern II: predicate → object → attribute

As first example of the pattern presented, we show the following patent US10003090B2:

"[...] *this* ***reduces labor*** *and* ***expenses*** *associated with assembly* [sic] *a cell stack assembly"*

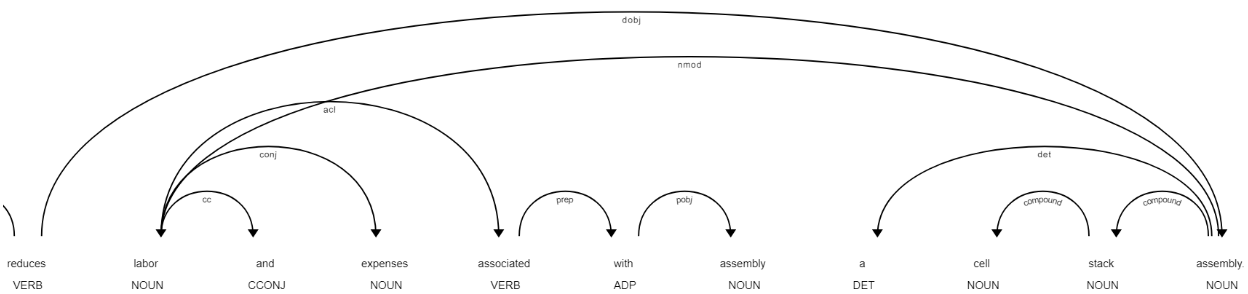


Figure 15: Example 1, pattern II

The graph in figure 15 appears more complex, but it is possible to see that it starts from the word "reduces", passing through "assembly" ("dobj": direct object), then "labor" ("nmod": nominal modifier) and finally concludes with "expenses" ("conj": conjunction). The structure is therefore *predicate* ("reduces") → [assembly] → *object* ("labor") → *attribute* ("expenses").

Another example, with a simpler semantic structure, is in patent US10010936B2:

"[..]accordingly, improved methods and articles of manufacture are needed to ***reduce labor*** *and* ***time*** *required for fabrication and to improve the quality of the part."*

The structure of figure 16 is indeed the following: from "reduce" (*predicate*) the link is directed to "labor"(*object*) which is connected to "time" (*attribute*).

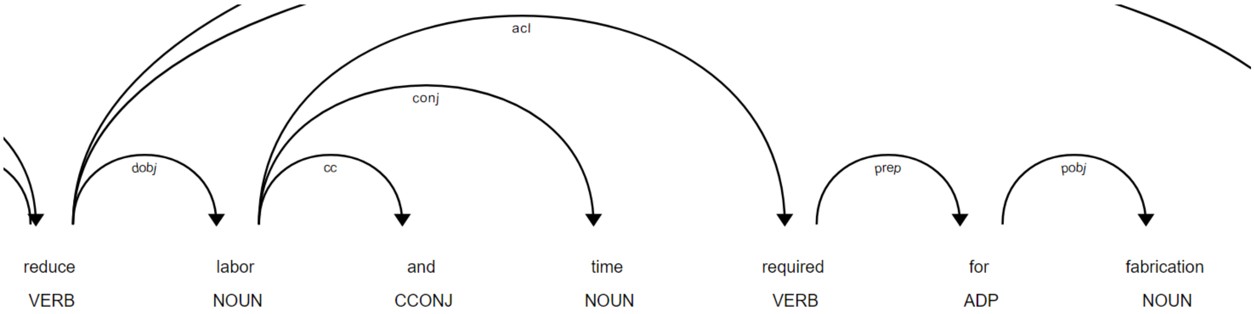


Figure 16: Example 2, pattern II

Pattern III: object → attribute → predicate

The text belongs to patent US8410636B2:

*"[...] installation of solar panels integrated with wireless power transfer may require **less skilled labor** since fewer electrical contacts need to be made."* Here, in figure 17 we have a minimalist structure starting from

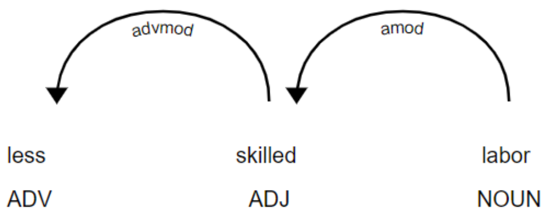


Figure 17: Example 1, pattern III

the word "labor" (*object*), connected through the adjective modifier "skilled" (*attribute*) which is connected with the adverbial modifier "less" (*predicate*).

Pattern IV: object → predicate → attribute

Patent US4723220A presents a more convoluted structure. As usual, we start from the text: *"the invention results in significant investment, installation **labor** and **time savings**."*

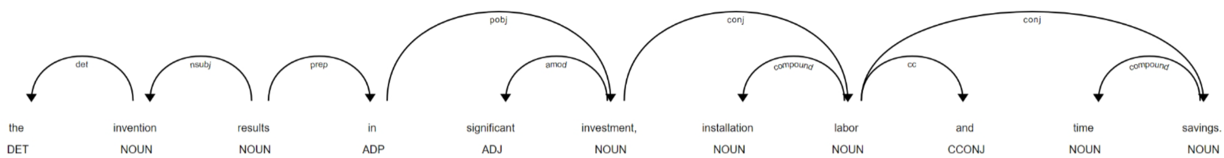


Figure 18: Example 1, pattern IV

The structure we are interested in starts with "labor" (*object*), connected through the conjunction "savings (*predicate*) to the compound "time" (*attribute*).

B Employment data and concordance tables

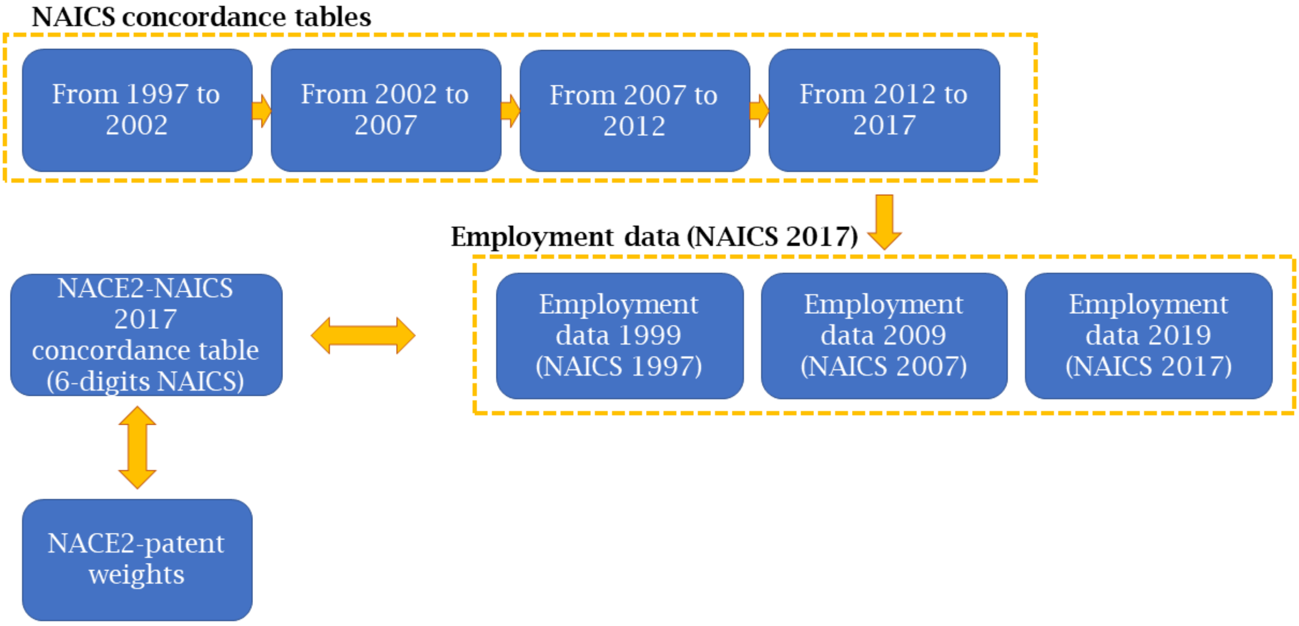


Figure 19: NAICS update and NACE concordance

The American classification for industrial activities (NAICS) is different from the European one (NACE): thus the concordance table used for this work [16] is the NACE2-NAICS concordance table at 6-digits level, referred at 2017. A second issue regards the release of revised NAICS classifications over time, as well described in US Census site (census.gov). We use employment data for three different years, 1999, 2009 and 2019, each belonging to a specific NAICS classification, in particular: year 1999 with NAICS 1997; year 2009 with NAICS 2007; year 2019 with NAICS 2017. Subsequent releases of NAICS classification (see upper part of 19) were used to uniform the data.

C Robustness check: manufacturing shares

We replicate the regression exercise of section 5.2 exclusively using employment in manufacturing. The exercise allows to capture within-manufacturing share changes. In addition, we can assess the robustness of

[16]The table used was downloaded during October 2022 at Eurostat-RAMON, while a new versions in now available.

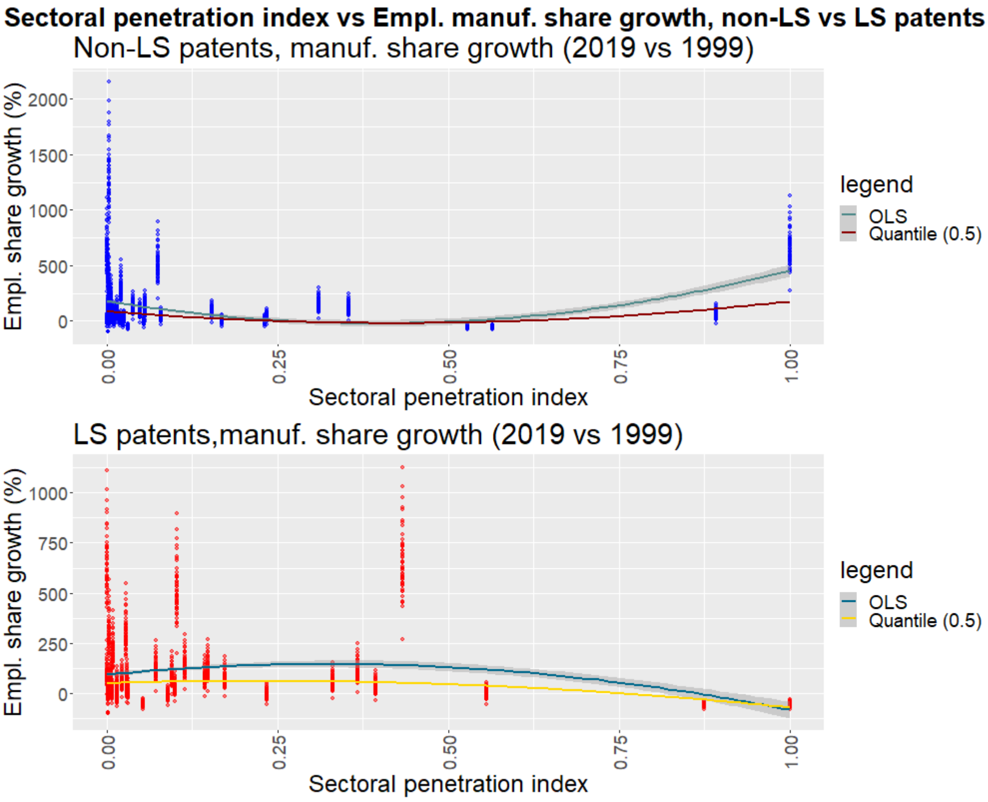
our identification and empirical strategy. We construct the following variable:

$$Manuf\,share_{i,j,t} = \frac{Employment\_sector_{j,t}}{Manufacture\,Federal\,Employment_t} * state\_weight_{i,t} \tag{10}$$

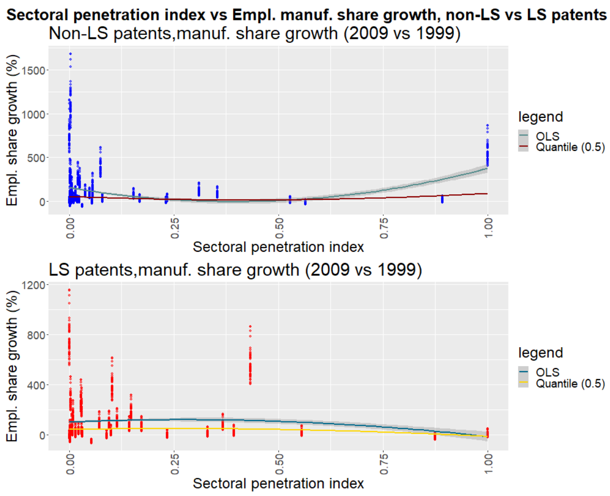
The regression coefficients are shown in table 8 and plotted in figure 20: results are all in line with the baseline specification.

Table 8: Quantile regression results (0.5), LS vs. non-LS patents

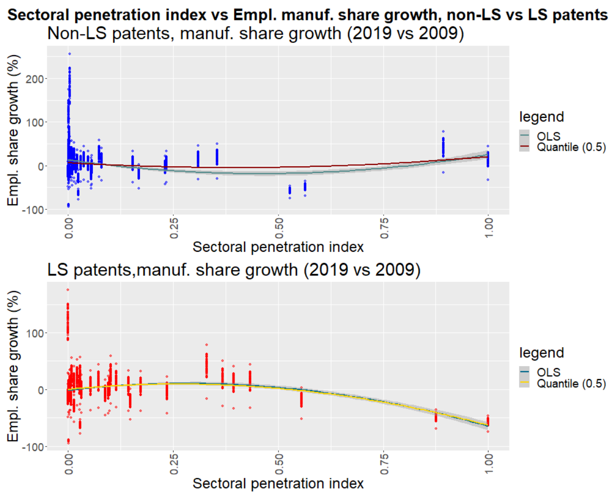
	NACE Employment manufacturing share growth, LS vs. non-LS patents					
	2019vs1999	2009vs1999 LS	2019vs2009	2019vs1999	2009vs1999 non-LS	2019vs2009
	(1)	(2)	(3)	(4)	(5)	(6)
Sectoral penetration index	1.061473^{***}	0.074066	0.824139^{***}	-4.983590^{***}	-1.122979^{*}	-0.261887^{***}
	(0.211299)	(0.351547)	(0.072674)	(0.750901)	(0.618508)	(0.065887)
Sectoral penetration index$\hat{2}$	-2.480830^{***}	-0.551530	-1.457756^{***}	5.118344^{***}	0.928259	0.337653^{***}
	(0.228018)	(0.347845)	(0.076132)	(1.834030)	(1.736114)	(0.065147)
Observations	1,785	1,785	1,785	2,091	2,091	2,091
Note:						*p<0.1; **p<0.05; ***p<0.01



(a) Manufacture share growth, 2019 vs. 1999



(b) Manufacture share growth, 2009 vs. 1999



(c) Manufacture share growth, 2019 vs. 2009

Figure 20: Regression plots, manufacturing

D Full digits CPC, LS vs. non-LS green patents

Top 20 TRUE LS		
CPC	Freq	**Description**
Y02P90/02	341	Enabling technologies with a potential contribution to greenhouse gas [GHG] emissions mitigation: Total factory control, e.g. smart factories, flexible manufacturing systems [FMS] or integrated manufacturing systems [IMS]
Y02E10/50	270	Energy generation through renewable energy sources: Photovoltaic [PV] energy
Y02P70/50	178	Climate change mitigation technologies in the production process for final industrial or consumer products: Manufacturing or production processes characterised by the final manufactured product
Y02E60/10	177	Enabling technologies; Technologies with a potential or indirect contribution to GHG emissions mitigation: Energy storage using batteries
Y02E10/47	175	Energy generation through renewable energy sources: Mountings or tracking
Y02D10/00	169	**Energy efficient computing, e.g. low power processors, power management or thermal management**
Y02P90/80	159	Enabling technologies with a potential contribution to greenhouse gas [GHG] emissions mitigation: Management or planning
Y02B10/10	141	CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TO BUILDINGS, e.g. HOUSING, HOUSE APPLIANCES OR RELATED END-USER APPLICATIONS:Integration of renewable energy sources in buildings: Photovoltaic
Y02T50/40	138	Aeronautics or air transport: Aeronautics or air transport; Weight reduction
Y02T10/70	126	Road transport of goods or passengers: Energy storage systems for electromobility, e.g. batteries
Y02D30/70	125	**Reducing energy consumption in communication networks: in wireless communication networks**
G06Q10/06	110	DATA PROCESSING SYSTEMS OR METHODS, SPECIALLY ADAPTED FOR ADMINISTRATIVE, COMMERCIAL, FINANCIAL, MANAGERIAL, SUPERVISORY OR FORECASTING PURPOSES; SYSTEMS OR METHODS SPECIALLY ADAPTED FOR ADMINISTRATIVE, COMMERCIAL, FINANCIAL, MANAGERIAL, SUPERVISORY OR FORECASTING PURPOSES, NOT OTHERWISE PROVIDED FOR: Administration; Management; Resources, workflows, human or project management
Y02E10/72	104	REDUCTION OF GREENHOUSE GAS [GHG] EMISSIONS, RELATED TO ENERGY GENERATION, TRANSMISSION OR DISTRIBUTION: Energy generation through renewable energy sources Wind turbines with rotation axis in wind direction
Y02T10/7072	103	CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TO TRANSPORTATION: Road transport of goods or passengers, Electromobility specific charging systems or methods for batteries, ultracapacitors, supercapacitors or double-layer capacitors
Y02B10/20	102	CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TO BUILDINGS, e.g. HOUSING, HOUSE APPLIANCES OR RELATED END-USER APPLICATIONS: Integration of renewable energy sources in buildings Solar thermal
Y02B20/40	100	CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TO BUILDINGS, e.g. HOUSING, HOUSE APPLIANCES OR RELATED END-USER APPLICATIONS: Energy efficient lighting technologies, e.g. halogen lamps or gas discharge lamps; Control techniques providing energy savings, e.g. smart controller or presence detection
Y02A90/10	94	TECHNOLOGIES FOR ADAPTATION TO CLIMATE CHANGE: Technologies having an indirect contribution to adaptation to climate change; Information and communication technologies [ICT] supporting adaptation to climate change, e.g. for weather forecasting or climate simulation
H04W84/12	92	WIRELESS COMMUNICATION NETWORKS: Network topologies WLAN [Wireless Local Area Networks]
H04L9/3247	89	TRANSMISSION OF DIGITAL INFORMATION, e.g. TELEGRAPHIC COMMUNICATION: arrangements for secret or secure communications; Network security protocols involving digital signatures
H04W88/08	89	WIRELESS COMMUNICATION NETWORKS: Devices specially adapted for wireless communication networks, e.g. terminals, base stations or access point devices; Access point devices

Top 20 FALSE LS		
CPC	**Freq**	**Description**
Y02E60/10	51758	Enabling technologies; Technologies with a potential or indirect contribution to GHG emissions mitigation: Energy storage using batteries
Y02P70/50	33856	Climate change mitigation technologies in the production process for final industrial or consumer products: Manufacturing or production processes characterised by the final manufactured product
Y02T10/12	31863	Road transport of goods or passengers: Improving ICE efficiencies
Y02T10/70	24847	Road transport of goods or passengers: Energy storage systems for electromobility, e.g. batteries
Y02E60/50	24069	Enabling technologies; Technologies with a potential or indirect contribution to GHG emissions mitigation: Fuel cells
Y02D10/00	22434	Energy efficient computing, e.g. low power processors, power management or thermal management
Y02D30/70	21575	Reducing energy consumption in communication networks: in wireless communication networks
Y02T10/40	17578	Road transport of goods or passengers: Engine management systems
Y02A50/30	15603	TECHNOLOGIES FOR ADAPTATION TO CLIMATE CHANGE: in human health protection, e.g. against extreme weather; Against vector-borne diseases, e.g. mosquito-borne, fly-borne, tick-borne or waterborne diseases whose impact is exacerbated by climate change
Y02T50/60	14617	CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TO TRANSPORTATION: Aeronautics or air transport; Efficient propulsion technologies, e.g. for aircraft
Y02T10/7072	13348	CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TO TRANSPORTATION: Road transport of goods or passengers, Electromobility specific charging systems or methods for batteries, ultracapacitors, supercapacitors or double-layer capacitors
H01M10/0525	11996	PROCESSES OR MEANS, e.g. BATTERIES, FOR THE DIRECT CONVERSION OF CHEMICAL ENERGY INTO ELECTRICAL ENERGY: Secondary cells; Manufacture thereof; Rocking-chair batteries, i.e. batteries with lithium insertion or intercalation in both electrodes; Lithium-ion batteries
Y02B70/10	11904	CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TO BUILDINGS, e.g. HOUSING, HOUSE APPLIANCES OR RELATED END-USER APPLICATIONS: Technologies for an efficient end-user side electric power management and consumption; Technologies improving the efficiency by using switched-mode power supplies [SMPS], i.e. efficient power electronics conversion e.g. power factor correction or reduction of losses in power supplies or efficient standby modes
Y02T10/62	11875	CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TO TRANSPORTATION: Road transport of goods or passengers; Hybrid vehicles
Y02E10/50	11087	Energy generation through renewable energy sources: Photovoltaic [PV] energy
Y02T10/72	10037	CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TO TRANSPORTATION: Road transport of goods or passengers; Electromobility specific charging systems or methods for batteries, ultracapacitors, supercapacitors or double-layer capacitors
H01M10/052	9256	PROCESSES OR MEANS, e.g. BATTERIES, FOR THE DIRECT CONVERSION OF CHEMICAL ENERGY INTO ELECTRICAL ENERGY: Secondary cells; Manufacture thereof; Li-accumulators
Y02E30/30	9124	REDUCTION OF GREENHOUSE GAS [GHG] EMISSIONS, RELATED TO ENERGY GENERATION, TRANSMISSION OR DISTRIBUTION: Energy generation of nuclear origin; Nuclear fission reactors
Y02E10/72	8982	REDUCTION OF GREENHOUSE GAS [GHG] EMISSIONS, RELATED TO ENERGY GENERATION, TRANSMISSION OR DISTRIBUTION; Energy generation through renewable energy sources Wind turbines with rotation axis in wind direction
Y02T10/64	8794	CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TO TRANSPORTATION; Road transport of goods or passengers; Electric machine technologies in electromobility

E Tagging legend

POS tagging of spaCy is based on the Universal POS tags. Here we show the list:

- ADJ: adjective
- ADP: adposition
- ADV: adverb
- AUX: auxiliary
- CCONJ: coordinating conjunction
- DET: determiner
- INTJ: interjection
- NOUN: noun
- NUM: mnumeral
- PART: particle
- PRON: pronoun
- PROPN: proper noun
- PUNCT: puncuation
- SCONJ: subordinating conjuction
- SYM: Symbol
- VERB: verb
- X: Other

For what concerns the Universal Dependency, we refer to the table at Universal Dependency for the complete list and full description. Here we report only some of the acronyms of the examples shown in the paper:

- *amod*: adjective modifier → "An adjectival modifier of a noun (or pronoun) is any adjectival phrase that serves to modify the noun (or pronoun). The relation applies whether the meaning of the noun is modified in a compositional way (e.g., large house) or an idiomatic way (hot dogs). An amod dependent may have its own modifiers (e.g., very large house) but the dependent should not be a clause. If it is a clause, then acl should be used"

- *acl:attr*: attributive adnominal clause → The acl:attr subtype of the acl relation is used for adnominal clause with attributive morphology [17]

- *compound*: compound → "The compound relation is one of three relations for multiword expressions (MWEs) (the other two being fixed and flat). It is used:

  - "for any kind of X^0 compounding: noun compounds (e.g., phone book), but also verb and adjective compounds that are more common in other languages (such as Persian or Japanese light verb constructions) [...]"
  - "for particle verbs (with the subtype compound:prt)"
  - "for serial verbs (with the subtype compound:svc)".

  The compound relation (nor any subtype thereof) is not used to link an inherently reflexive verb with the reflexive morpheme, despite the similarity of this construction to particle verbs. The current UD guideline is to use an appropriate subtype of the expl relation. Each language that uses compound should develop its own specific criteria based on morphosyntax (rather than lexicalisation or semantic idiomaticity), though elsewhere the terms "compound" and "multiword expression" may be used more broadly [...]"

- *conj*: conjunct → "A conjunct is the relation between two elements connected by a coordinating conjunction, such as and, or, etc. We treat conjunctions asymmetrically: The head of the relation is the first conjunct and all the other conjuncts depend on it via the conj relation."

- *dobj*: direct object → "The direct object of a VP is the noun phrase which is the (accusative) object of the verb." [18]

- *nmod*: nominal modifier → "The nmod relation is used for nominal dependents of another noun or noun phrase and functionally corresponds to an attribute, or genitive complement [...]"

- *advmod*: adverbial modifier → "An adverbial modifier of a word is a (non-clausal) adverb or adverbial phrase that serves to modify a predicate or a modifier word.

  In some contexts and languages, a limited set of adverbs can also modify nominals (e.g., only on Monday). The advmod relation or its subtype has to be used in such cases, too (see also advmod:emph).

  Note that in some grammatical traditions, the term adverbial modifier covers constituents that function like adverbs regardless whether they are realised by adverbs, adpositional phrases, or nouns in particular morphological cases. We differentiate adverbials realised as adverbs (advmod) and adverbials realised by noun phrases or adpositional phrases (obl). However, we do not differentiate between modifiers

[17] `https://universaldependencies.org/ckt/dep/acl-attr.html`
[18] `https://universaldependencies.org/docs/en/dep/dobj.html`

of predicates (adverbials in a narrow sense) and modifiers of other modifier words like adjectives or adverbs (sometime called qualifiers). These functions are all subsumed under advmod.”