Will green innovation augment or automate labor?

Aurélien Saussay*[1], Misato Sato[1], and Francesco Vona[2]

[1] *London School of Economics*
[2] *University of Milan*

March 26, 2026

Preliminary draft. Please do not cite.

Abstract

We measure the labor orientation of green innovation-whether patented low-carbon technologies are oriented towards automating existing tasks or towards the creation of new work. Using 1.3 million U.S. patents filed between 1980–2022 and CPC climate tags (Y02 and Y04S), we construct a patent-level labor-orientation index from semantic similarity between patent text and occupational task descriptions and newly emerging work titles, following Autor et al. (2024). We find heterogeneity across green technologies: green ICT and smart grids are relatively labor-augmenting, while adaptation is more labor-augmenting. Within renewable and electric vehicle patents, the labor-augmenting share declines over time. We then relate patent orientation to diffusion in a patent-by-technology panel that matches patents to country-level deployment in physical units: higher lagged deployment predicts a lower labor-augmenting share conditional on rich sets of country and technology-specific fixed effects and trends. These estimates are consistent with learning-by-doing whereby scaling and standardization tilt green inventive activity towards more labor-saving content.

*Corresponding author: a.saussay@lse.ac.uk.

1 Introduction

To be cited: Kalyani et al., 2025

Decarbonizing the economy is not only an energy transition—it is also a labor-market transition. Reaching net-zero requires a rapid scale-up of low-carbon technologies and the jobs and skills needed to deploy, operate, and maintain them. Yet the employment consequences of the green transition are often assessed assuming that the labor intensity of "green technologies" are fixed. This is difficult to justify in a context where the underlying technologies evolve rapidly.

In a broader induced innovation framework (Acemoglu, 1998; Hicks, 1963), the labor content of a technology depends on design choices that are made during invention and refinement, and those choices can shift as markets expand and production and deployment become more standardized, allowing an easier substitution of workers with machines (Acemoglu & Restrepo, 2018; Autor et al., 2003). This paper studies the labor orientation of green innovation and asks whether it systematically shifts with technology deployment, towardss task automation or towardss the creation of new work. By doing so, we aim to shed light on an endogenous margin with important implications for the distributional and spatial incidence of decarbonisation.

The labor economics literature emphasizes that the employment consequences of technical change depend on whether it automates tasks or creates new tasks and new occupations (Autor et al., 2024). Task-based frameworks distinguish technologies that substitute directly for existing tasks from technologies that complement workers by expanding the set of tasks performed, reorganizing production, or raising the value of human inputs (Acemoglu & Restrepo, 2018, 2019; Autor et al., 2003). Within that framework, the same broad technology can generate both displacement and new work, and the relative strength of these margins can evolve as technologies diffuse (Acemoglu & Restrepo, 2019). Recent work has developed tools to measure the degree of automation versus augmentation orientation of innovation exploiting the textual similarity between patents and tasks (Autor et al., 2024; Kogan et al., 2023), moving beyond indirect inference where the task content of occupation is fixed (Atalay et al., 2020). To our knowledge, these tools have not been applied to the technologies at the core of decarbonization.

In environmental economics, evidence on induced-innovation is well-established and shows that climate policy shift inventive activity towards cleaner technologies (Popp et al., 2020; Probst et al., 2021). That literature has primarily studied how much clean innovation occurs, where it occurs, and how it diffuses, using patent counts, citations, and self-reported measures of technology adoption as outcomes and using either demand-pull or regulatory push as drivers. These contributions have established that policy and market incentives shape the level and composition of climate-related inventive activity. However, they are largely silent on whether induced green innovation complements or saves labour. A patent-based measure that allows distinguishing automation-oriented from augmentation-oriented content provides important insights on the potential effects of a green demand-pull on labor demand. A green market expansion can shift not only the direction and

quantity of green innovation, but also the type of work that the resulting technologies are designed to economize on or to complement.

A parallel literature studies green jobs and labor market adjustment to environmental policy. This evidence documents heterogeneity in the skills required by green occupations and the geography of green work, and it measures worker transitions associated with the expansion of low-carbon industries (Consoli et al., 2016; Curtis & Marinescu, 2023; Curtis et al., 2023; Elliott et al., 2024; Popp et al., 2022; Saussay et al., 2025; Vona et al., 2018, 2019). Related work emphasizes that the employment consequences of environmental policy or a green demand push operate through sectoral reallocation and changes in the occupational structure (Castellanos & Heutel, 2019; Frattini et al., 2025; Hafstead & Williams III, 2018; Popp et al., 2022), also highlighting some similarities with the impact of trade shocks (Weber, 2020). These empirical studies measure labor demand in production, installation, and operations - outcomes shaped jointly by adoption, regulation and local labour supply. They do not directly observe the intended planned labor content of inventive activity that precedes and accompanies adoption, and they cannot separate ex-ante technology choices from changes in regulation, local labor supply, and other structural factors. We instead focus on an upstream innovation margin: the task content that patented technologies appear designed to automate versus complement, and how that content evolves with diffusion.

We bring these strands of literature together by measuring whether green invention is oriented towards labor-saving automation or towards labor augmentation, and by relating that orientation to technology deployment. The first step is measurement. Following Autor et al., 2024, we construct a patent-level labor-orientation index from semantic proximity between patent text and two labor-related corpora. First are the occupational task descriptions from the Dictionary of Occupational Titles (DOT) from the U.S. Department of Labor, Employment and Training Administration, which capture language associated with existing task inputs of workers. The second is the titles of "new work" occupations, which instead capture that associated with the creation of new tasks. The index is a continuous share that is high when a patent text is more similar to the new-work corpus, indicating that the technology is more labour-augmenting. Conversely, the index is low when more similar to descriptions of established tasks, indicating the technology is more labour-saving. It is a relative orientation measure based on patent text, and does not claim to measure realized employment effects, wages, or occupational composition.

Applying this measure to 1.3 million U.S. green patents (1980–2022), we document substantial heterogeneity in labor orientation across green technology categories, with differences of around five percentage points between the most labor-augmenting and most labor-saving categories. Focusing on deployable technologies, the labor-augmenting share trends downward within renewables and electric-vehicle technologies over time. In a matched patent–technology panel, higher lagged deployment predicts a lower labor-augmenting share within technology classes.

The second step is to apply this measure to a broad universe of green patents and to connect it

to technology maturity. We study U.S. patents[1] with priority years 1980–2022[2] and identify green patents using Cooperative Patent Classification (CPC) climate-change mitigation and adaptation tags (Y02 and Y04S). This green definition is based on CPC assignments following convention in the literature, and does not use keyword screening. We report descriptive statistics using citation weights to give more importance to heavily cited patents. Green patenting spans across multiple technology domains, including software and control systems, heavy engineering, chemical processes, and building technologies. In our data, the difference between the highest- and lowest-orientation green categories is on the order of five percentage points when aggregated over 1980–2022 with citation weights, a dispersion comparable to differences across the whole range of broad non-green technology classes under the same measure. This heterogeneity implies that statements about "green innovation" and labor orientation are sensitive to aggregation across green categories.

The analysis of the effect of deployment focuses on a subset of "deployable" green technologies for which diffusion can be measured in physical units. We construct consistent time series for renewable energy technologies and for electric-vehicle-related technologies, allowing us to match patent classes to deployment paths at the technology-year level. Our deployment series begin in 2000, which constrains the earliest patent cohorts that can be used once lags are imposed. Our empirical design therefore addresses a substantive question with a more precise scope than the descriptive cross-category evidence: how the labor orientation of innovation evolves within green technologies that are scaling through deployment as they move from early diffusion towards maturity?

Deployment is a natural proxy for market size and diffusion stage in deployable technologies, and models of induced technical change imply that market size can affect the direction of innovation. Our working hypothesis is that technologies become relatively more labor-saving as they scale because the associated tasks become more standardized and easier to automate (Acemoglu & Restrepo, 2018; Goldin & Katz, 2009; Vona & Consoli, 2015). As technologies move from early adoption to maturity, firms often reallocate inventive effort from frontier exploration and performance improvements (which requires tacit knowledge and creativity) towards manufacturability, reliability, process engineering, and cost reduction (which are easier to automate). In settings where labor is an adjustable and replaceable input into these more standardized tasks, innovation can plausibly have a more labor-saving character, even within narrowly defined technology classes. A descriptive implication in our data is that the labor-augmenting share of renewables and electric-vehicle technologies declines over time, but this can simply reflect a spurious correlation. Our empirical strategy tests whether this pattern lines up with diffusion rather than being a simple reflection of technology-specific time trends. In practice, we regress the labor-orientation index of deployable green patents on lagged global deployment within technology classes, conditioning on year fixed effects and progressively finer technology controls and technology-specific trends.

We find a robust negative association between deployment and the labor-augmenting share of

[1] In this paper, we focus on utility patents, which protect functional inventions such as products, processes, machines, or compositions of matter, in contrast to other patent types i.e. design and plant patents.

[2] Priority year indicates the year in which the patent was first filed, rather than grant date

green patents, conditional on the controls implemented in the analysis. The currently implemented design does not resolve the potential endogeneity of deployment. The appropriate inference at this stage is therefore about the direction of the relationship between diffusion and the labor orientation of inventive activity, not about downstream employment levels. In the automation versus new-tasks framework, the estimates are consistent with a maturity-induced tilt of green innovation towards task-substituting patents as technologies scale up, while leaving open the mechanisms and general-equilibrium effects that would be required to map patent text into employment changes.

The paper contributes to three literatures. First, we contribute to the literature on induced clean innovation and directed technical change by adding a direction-of-innovation outcome that is defined with respect to labor rather than with respect to emissions intensity. Evidence around demand-pull in environmental economics has established that policies and markets can redirect inventive activity towards low-carbon domains. We show that diffusion is also associated with a reorientation of innovation content within those domains, moving along an automation-versus-augmentation margin. This complements the climate innovation literature that measures clean inventive activity using patent-based classifications and emphasizes diffusion dynamics, because it shows that the composition of innovation within a fixed green technology definition is not invariant to deployment (Popp et al., 2020; Probst et al., 2021).

Second, we contribute to the literature on automation and new tasks by applying a task-based, patent-text measure of labor orientation to green technologies and by linking that measure to a technology state variable with a clear economic interpretation. Our results indicate that, within green technologies, the direction of innovation covaries with diffusion as measured by deployment scale. Work on automation emphasizes that displacement and new task creation are both endogenous responses to incentives and can be observed in innovation outputs (Acemoglu & Restrepo, 2019; Autor et al., 2024). Recent patent-based evidence shows that firms shift innovative effort towards automation in response to incentives, and that such shifts can be measured directly in patent data (Dechezleprêtre et al., 2023; Hémous et al., 2024). Our setting differs in that the conditioning variable is deployment scale in low-carbon technologies rather than a labor-cost shock or an automation-specific policy. This connects the automation and climate innovation literatures through a common patent-based measurement framework. Related "twin transition" evidence uses patent text to study the interaction between digitalization and decarbonization. Our analysis is complementary in that it measures labor orientation explicitly and ties it to deployment, which provides an empirical anchor for maturity rather than relying solely on time trends or broad digital classifications (Bianchini et al., 2023; Rughi et al., 2023).

Third, we contribute to the green jobs literature by providing an upstream measure of the labor orientation of the technologies that are being invented and refined during the transition. Evidence on green jobs documents differences in skills, wages, and mobility across green work and measures adjustment costs and reallocation in response to environmental policy (Curtis et al., 2023; Marin & Vona, 2019; Saussay et al., 2025; Vona et al., 2018). Those outcomes reflect adoption, regulation,

and local labor market conditions, and they often occur downstream of innovation. By focusing on patents, we isolate an innovation-content margin that can precede and accompany adoption and that can evolve as deployment scales. We document that the labor orientation of green invention is heterogeneous across technologies and systematically related to deployment within deployable green technologies, which provides a channel through which the task content of green production and installation work could evolve over time.

Section 2 defines the labor-orientation measure and describes the patent data and green technology definitions. Section 3 documents heterogeneity across green technologies and time. Section 4 lays out the deployment-linked designs. Section 5 reports results, and Section 6 concludes.

2 Measuring the labor orientation of green innovation

Patent text provides a tractable proxy for the task content of innovation. We measure whether green inventions are oriented towards labor-saving automation or towards augmenting labor using a task-based, embedding-driven approach adapted from Autor et al. (2024). The adaptation is twofold. First, we apply the method to patents tagged as climate-related using CPC classifications. Second, we implement the semantic matching using a subword embedding model, fastText, which is well suited to patent text because it yields vectors for rare technical terms and morphologically complex compounds that are frequently out of vocabulary under fixed-dictionary embeddings (Bojanowski et al., 2017).

2.1 Patent data and green identification

We obtain the raw patent universe from the Google Patents dataset (Google, 2017) and restrict the sample to U.S. patent publications with publication years 1980–2022. For each publication, we use the title, abstract, CPC codes, assignee information, and forward citations. We construct a single text field for embedding by concatenating the patent title, the full title of the patent's main CPC-6 digit class, and the abstract. Using the CPC-6 class title adds a standardized technical descriptor that is present even when abstracts are terse and that provides a bridge between specialized patent language and more general descriptors.

Green patents are identified using CPC "climate" tags. Concretely, a patent is classified as green if it has at least one CPC code in Y02 (climate change mitigation and adaptation technologies) or Y04S (smart grids). This classification is deterministic given CPC codes and does not rely on keyword screening or supervised learning, consistent with the use of CPC climate tags in the environmental innovation literature (Haščič & Migotto, 2015). For patents with multiple green-tagged CPC codes, we assign a primary green CPC-4 and CPC-6 class as the modal green CPC-4 and CPC-6 among the patent's CPC list. We use CPC-4 green tags to group patents into nine green technology families: climate adaptation (Y02A), green buildings (Y02B), carbon capture and storage (Y02C), green ICT (Y02D), renewables and green energy (Y02E), industrial process

decarbonization (Y02P), green transportation (Y02T), waste management (Y02W), and smart grids (Y04S).

2.2 Embedding-based labor-orientation measure

Our conceptual starting point is the task approach to technological change (Acemoglu & Autor, 2011; Autor et al., 2003). Patents that describe technologies substituting for existing workplace tasks should be textually close to occupational task descriptions, while patents that embody new kinds of work should be close to titles of newly emerging occupations. Autor et al. (2024) operationalize this idea by measuring semantic proximity between patent text and (i) task statements from the U.S. Dictionary of Occupational Titles and (ii) occupation titles from the U.S. Census Alphabetical Index of Occupations that proxy for "new work". We follow their reference-corpus construction and scoring logic and reimplement the measure using our patent corpus and embedding pipeline.

Let p index patent publications. Let x_p denote the patent document text (title, main CPC-6 title, abstract). We represent each patent as a vector embedding $e_p \in \mathbb{R}^d$ constructed as a TF-IDF weighted average of token embeddings. Specifically, after standard tokenization and stopword removal, we compute TF-IDF weights ω_{pw} for each token w appearing in patent p and obtain a word vector $v_w \in \mathbb{R}^d$ from a pretrained fastText model. The patent embedding is

$$e_p = \frac{\sum_{w \in \mathcal{V}_p} \omega_{pw} v_w}{\left\| \sum_{w \in \mathcal{V}_p} \omega_{pw} v_w \right\|}, \tag{1}$$

where $\mathcal{V}_p$ is the set of tokens appearing in p after preprocessing. The key implementation choice is the use of fastText embeddings with subword information (Bojanowski et al., 2017). Patent language is dense in technical vocabulary, abbreviations, hyphenated compounds, and inflected forms. Under fixed-vocabulary embeddings, tokens absent from the model dictionary must be dropped or mapped to an undifferentiated unknown token, which can be consequential in patents where rare terms carry substantial semantic content. In contrast, a subword model constructs embeddings from character n-grams, providing meaningful vectors for rare or previously unseen tokens and improving the link between specialized patent language and the more common terminology used in occupational descriptions.

We then compute semantic similarity between each patent embedding and two sets of reference embeddings. Let a_j denote embeddings of $j \in [1, J]$ task-description reference texts (automation-oriented content), and let g_k denote embeddings of $k \in [1, K]$ "new work" occupation titles (augmenting content), as in Autor et al. (2024). We compute semantic similarity using cosine similarity between the embedding vectors. To remove the common component that reflects generic overlap between patent language and the reference corpus, we demean similarity scores within each patent by subtracting the median similarity across reference texts:

$$\tilde{s}_{pj}^{\text{aut}} = e_p' a_j - \text{median}_\ell(e_p' a_\ell), \quad \tilde{s}_{pj}^{\text{aug}} = e_p' g_j - \text{median}_\ell(e_p' g_\ell) \tag{2}$$

Following Autor et al. (2024), we convert similarity profiles into exposure counts by thresholding. For each of the two reference sets (automation and augmentation), and within publication-year decades, we compute a scalar threshold as the 85^{th} percentile of the empirical distribution of $\tilde{s}$ across all patents and all reference texts. We define the patent's automation exposure as the number of task reference texts for which similarity exceeds the automation threshold. The augmentation exposure is defined analogously:

$$\text{Aut}_p = \sum_{j=1}^{J} \mathbf{1}\left[\tilde{s}_{pj}^{\text{aut}} > \tau_{\text{aut}}(d_p)\right], \quad \text{Aug}_p = \sum_{k=1}^{K} \mathbf{1}\left[\tilde{s}_{pj}^{\text{aug}} > \tau_{\text{aug}}(d_p)\right] \tag{3}$$

where d_p denotes the publication-year decade containing patent p.

Our core object of interest is a patent-level labor-orientation index defined as the relative weight of augmentation exposure:

$$y_p \equiv \frac{\text{Aug}_p}{\text{Aug}_p + \text{Aut}_p} \in [0, 1] \tag{4}$$

defined for patents with $\text{Aug}_p + \text{Aut}_p > 0$. Higher values of y_p indicate that the patent's semantic proximity is tilted towards "new work" titles relative to task descriptions; lower values indicate a tilt towards task-replacing content. The index is a relative orientation measure derived from the semantic analysis of patent text. As such it should not be interpreted as a direct estimate of realized employment effects, task adoption, or productivity impacts. Further algorithmic and implementation details are provided in the Appendix.

3 Stylized facts on heterogeneity across green technologies

Green patenting spans wide ranging technological domains from software and control systems, heavy engineering, chemical processes, to building technologies. A blanket statement about whether green innovation as a whole tends to automate or augment labor could therefore be misleading. The stylized facts in this section characterize heterogeneity in the labor orientation of green innovation at a level that is closer to technology content than existing evidence based on occupations, establishments, or policy-induced changes in aggregate green patenting. The empirical labor literature emphasizes that the direction of technical change matters for labor market adjustment because automation and augmentation operate through different margins, including task displacement and the creation of new tasks (Acemoglu & Restrepo, 2019; Autor et al., 2003). The climate innovation literature, by contrast, has mainly measured where and when clean innovation occurs and how it diffuses (Popp et al., 2020; Probst et al., 2021). The descriptive evidence below links these two approaches by mapping green technologies onto the patent-text labor-orientation scale recently developed to measure automation vs augmentation-oriented invention (Autor et al., 2024; Hémous et al., 2024).

3.1 Labor orientation across green technology classes

Figure 1 locates green technology categories within the broader distribution of labor orientation across patenting fields. The outcome is the patent-level labor-orientation index, aggregated to technology cells and expressed as the share of labor-augmenting content in a citation-weighted mixture of augmenting and automating signals. The figure uses a common scale for green and non-green technologies, which allows a direct comparison between the distribution of labor orientation across standard technology classes and the point estimates for green categories.

Two facts emerge. First, green technologies are dispersed across the overall distribution rather than clustering at one end. In Figure 1, the distance between the lowest and highest green-category markers is about five percentage points, which is comparable to the spread across broad non-green technology classes. This dispersion is large relative to the level differences typically discussed in work that treats "clean" technologies as a single sectoral aggregate. Second, the ranking across green categories aligns with intuitive differences in technology content that matter in the automation and new-tasks framework. Green ICT and smart grids lie towards the upper end of the distribution (more labour augmenting), transportation, and several engineering-oriented categories sit closer to the middle of the overall distribution, while adaptation sits at the lower end (more labour saving).

The comparison to non-green technologies is informative for two reasons. First, it separates heterogeneity that is specific to "green" content from heterogeneity that is characteristic of patenting fields more generally. Instruments and information technologies have relatively high labor-augmenting shares in the non-green distribution, and the green ICT and smart-grid categories occupy a similar region of the scale. Second, it clarifies that 'green" status is not a sufficient statistic for the labor orientation of the underlying innovation. This distinction is absent from most empirical work on green innovation, which uses CPC-based definitions to study induced clean innovation, diffusion, and the location of inventive activity (Popp et al., 2020; Probst et al., 2021). It is also absent from the green jobs literature, which documents heterogeneity in skill requirements and worker transitions across green occupations and industries (Bowen et al., 2018; Consoli et al., 2016; Curtis & Marinescu, 2023; Curtis et al., 2023; Saussay et al., 2025). That evidence speaks to realized labor demand in production and installation activities. Figure 1 instead measures the content of inventive activity and shows that heterogeneity arises already at the level of patented technologies.

A measurement-based interpretation follows from how the labor-orientation index is constructed. The index increases when a patent's text is closer to the language of new occupation titles than to the language of occupational task descriptions (Autor et al., 2024). Technologies that rely on software, monitoring, and control systems may be more likely to introduce new tasks and complementary inputs, while technologies that prioritize mechanization, process control in physical production, or labor-saving equipment may be more likely to resemble task descriptions associated with displacement (Acemoglu & Restrepo, 2019; Hémous et al., 2024). Figure 1 is consistent with such a mapping. It further establishes a constraint for any explanation that treats green innovation as

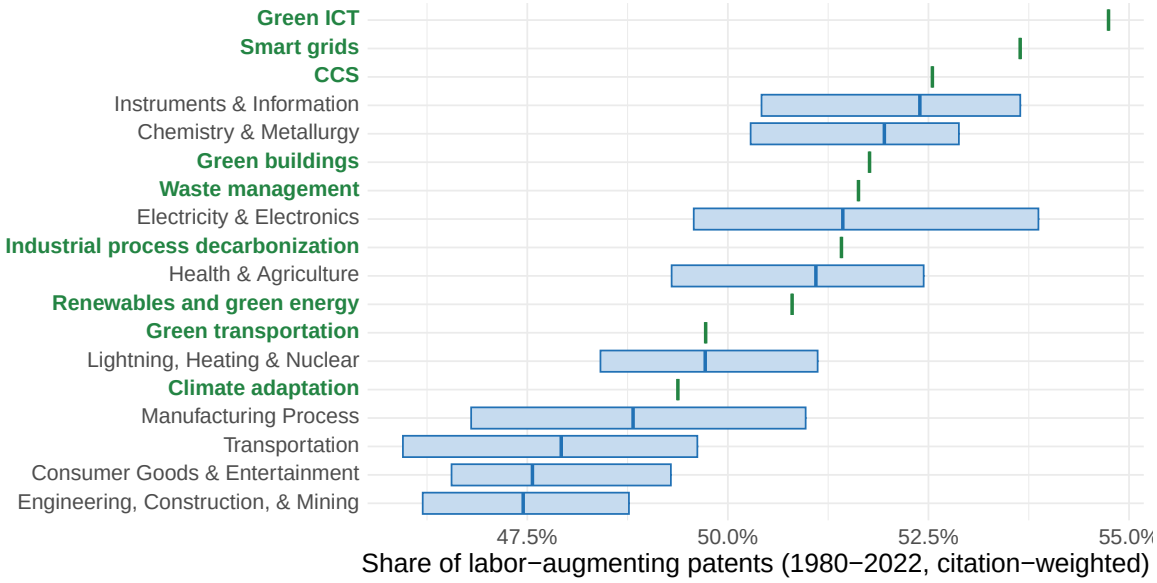


Figure 1: Labor orientation across green and non-green technology classes

Notes: The horizontal axis reports the citation-weighted share of labor-augmenting content, constructed from patent-level augmenting and automating exposure scores following Autor et al. (2024). Blue boxplots summarize the distribution across 4-digits CPC subfields within each broad non-green technology class; each boxplot is formed from technology-by-subfield cells over 1980–2022, with forward-citation weights as in the main analysis. Green vertical markers report the corresponding citation-weighted share for green categories defined by single 4-digits CPC climate tags aggregated to nine green technology groups (e.g., renewables, green transportation, smart grids, green ICT).

homogeneous: within-green differences are of the same order as differences across major technology fields.

3.2 Evolution over time within deployable green technologies

Cross-sectional heterogeneity does not imply stability. Directed technical change models allow the direction of innovation to evolve as technologies diffuse and production processes change, even within a narrowly defined technology (Acemoglu & Restrepo, 2019). In the climate domain, diffusion and scale often follow long adoption cycles with sharp changes in deployment during the 2000s and 2010s (Probst et al., 2021). Figure 2 documents time variation within two deployable green technology groups for which deployment data can later be matched to patents: renewables and electric-vehicle-related technologies.

Two patterns stand out. First, the labor-augmenting share trends downward in both groups over the 1980–2020 priority-year window. The decline is modest in absolute magnitude, but must be considered within the context of the indicator, which only spans five percentage points across all technologies (see previous sub-section). Second, year-to-year variation is sizable, particularly in the early part of the sample, which indicates that annual averages are noisy even when citation-weighted. Figure 2 therefore supports treating labor orientation as a dynamic variable rather than

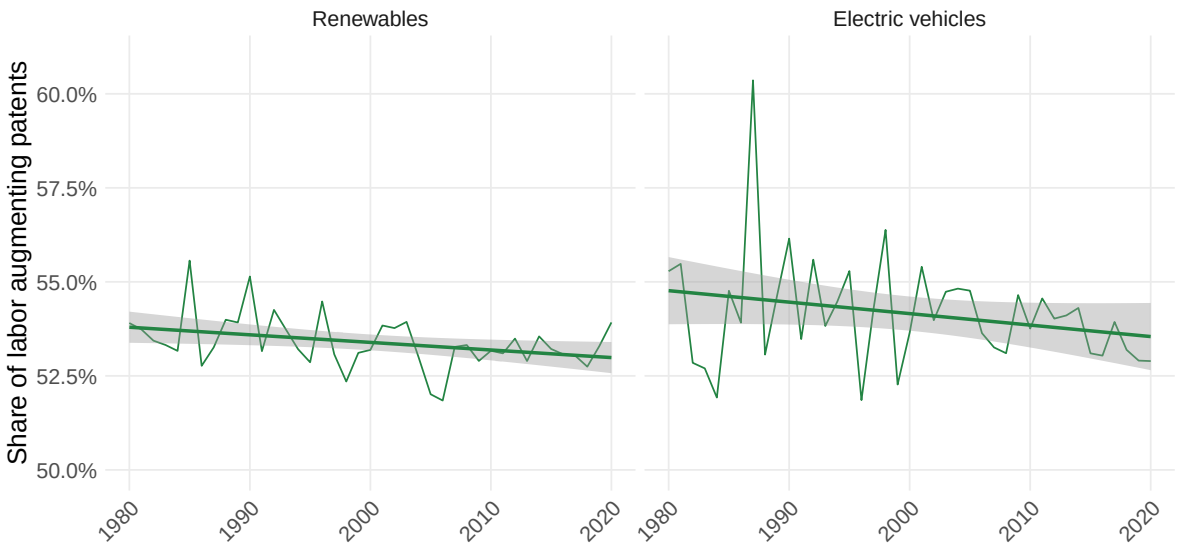


Figure 2: Labor orientation over time in deployable green technologies

Notes: The figure plots, by priority year, the citation-weighted mean of the patent-level labor-orientation index for (i) renewables and green energy and (ii) green transportation patents mapped to electric-vehicle-related subcategories (hybrid, plug-in electric, and hydrogen vehicles) as defined by CPC climate tags. The solid lines show annual citation-weighted averages; the fitted lines and shaded bands correspond to linear trends and 95% confidence intervals estimated by year within each panel.

a fixed attribute of a green category.

These time patterns relate to three active strands of the literature. In the labor and automation literature, the interaction between diffusion and task content is central, because the same broad technology can shift from frontier experimentation towards incremental improvements, standardization, and process engineering, which can change the balance between new tasks and displacement (Acemoglu & Restrepo, 2019; Hémous et al., 2024). In the climate innovation literature, the focus has been on the timing, diffusion, and determinants of low-carbon inventive activity, typically measured by patent counts and citations (Popp et al., 2020; Probst et al., 2021). Figure 2 adds information about direction within a given green technology group, holding fixed the definition of the technology. In the green jobs literature, the dominant outcomes are occupational transitions, job postings, and the skill content of green work (Bowen et al., 2018; Consoli et al., 2016; Curtis et al., 2023; Saussay et al., 2025). Those outcomes combine technology adoption, regulation, and local labor market conditions. By focusing on patent text, Figure 2 isolates the orientation of inventive activity that can precede, accompany, or follow shifts in production and installation work.

A final implication of Figure 2 is that within-technology time trends can mechanically confound comparisons across technologies or policy eras. In later sections we therefore treat time effects and technology-specific trends as baseline controls when linking labor orientation to deployment. The evidence in Figure 2 motivates that choice by documenting that time variation in labor orientation is empirically relevant within deployable green technologies.

3.3 Alignment with an external automation classification

The interpretation of the labor-orientation index rests on a mapping between patent text and task content. Here we test the external validity of our mapping by comparing our continuous index to an independent, discrete automation classification from (Hémous et al., 2024). Because the Hémous et al. indicator is constructed using a different methodology and for a different purpose, it provides a useful benchmark for assessing whether our index meaningfully captures automation-oriented content.

Appendix Figure A.2 reports the citation-weighted time series of the share of green patents classified as automation patents using the Hémous et al. (2024) definition, implemented on our patent corpus. While this variation does not mechanically imply changes in the continuous labor-orientation index, it provides an external reference series based on different inputs. Appendix Table B.1 then relates the two measures at the patent level: patents classified as automation patents by Hémous et al. (2024) have systematically lower values of the labor-orientation index in regressions that condition on increasingly demanding sets of year and technology fixed effects. This association is a necessary property of the index if it is to be interpreted as an orientation measure, and it supports using the index to compare technologies and track within-technology changes over time. The Hémous et al. (2024) indicator is a discrete classification built for a different purpose, which makes it a good candidate to assess the external validity of our continuous patent-level index.

4 Empirical strategy

Deployment is an observable proxy for the scale of a green technology's market and its stage of diffusion. Our empirical analysis links the labor-orientation index for green patents, y_p, to deployment measures for a set of technologies for which deployment is directly observed. The core challenge is that deployment is an equilibrium outcome that may covary with unobserved determinants of patent text and, in some cases, may respond to innovation. Our strategy therefore implements a technology-country design that relates the labor orientation of patents to changes in country-specific deployment within narrow green technology classes over time.

4.1 Deployment data and the patent–technology match

We assemble a deployment dataset for a set of *deployable* green technologies that have consistent time series in administrative or statistical sources. The deployable set covers seven renewables and three electric-vehicle related categories. In the patent data, a patent is assigned to a deployable technology based on the presence of specific CPC climate-mitigation subclasses within the CPC Y02 and related tags. The deployable technology indicators are defined at the CPC 6-digit level using CPC string patterns, and include: biofuels (Y02E50/1), marine energy (Y02E10/3), hydro energy (Y02E10/2), wind energy (Y02E10/7), fuel from waste (Y02E50/3), solar photovoltaic

(Y02E10/5), solar thermal energy (Y02E10/4), hybrid vehicles (Y02T10/62), plug-in electric vehicles (Y02T90/14), and hydrogen vehicles (Y02T90/40).

Spatial and temporal coverage in the deployment dataset varies by technology, from an average of 15 countries over 10 years for *e.g.* marine energy to more than 130 countries over 25 years for solar photovoltaic or wind energy. Our analysis is restricted to patents that can be matched to a deployment series for at least one deployable technology. Appendix Table B.2 documents the crosswalk between the deployable technologies, the CPC definitions used to select patents, and the corresponding matched patent counts.

Deployment series are constructed at two levels. For the technology-country design, deployment is measured at the country-level for each deployable technology and year. Appendix Figure B.1 plots global deployment by technology from 2000 to 2022 and shows the strong heterogeneity in diffusion rates across technologies.

Two features of the patent–technology match matter for interpretation and inference. First, a patent can meet the CPC criteria for multiple deployable technologies. In the technology–country design, the unit of observation is therefore a patent–technology pair, and observation weights split each patent equally across its mapped technologies. Second, the CPC-based definitions select patents based on classification rather than keywords. The approach is transparent and replicable, but it inherits any reclassification or definitional drift in CPC assignment.

4.2 Technology-country design at the patent level

Our empirical design relates the labor orientation of deployable green patents to changes in global deployment for the corresponding technology. Let p index patents, $c \in \mathcal{C}$ index patent assignee's countries, $k \in \mathcal{K}$ index deployable technologies, and t index patent priority year. Let $y_{p,c}$ denote the labor-orientation index defined in Section 2, constructed from patent text as in Autor et al. (2024), where patent p's assignee is based in country c. Let $D_{k,c,t}$ denote deployment for technology k in country c and year t. The estimating sample consists of patents that map to at least one (k, c) pair and have non-missing $y_{p,c}$. A patent that maps to multiple k contributes one observation per mapped technology. Denote that count by m_p, and define the splitting weight $s_{p,k} = 1/m_p$ for each mapped pair.

The main regressor is the log deployment level, $\log(1 + D_{k,c,t-L})$, where $L = 5$ years in our main specification. The lag is chosen to reduce mechanical simultaneity between patenting in year t and deployment in the same year, and it implies that the earliest priority year entering the matched sample is constrained by the length of the deployment series. With deployment observed from 2000 onward, a five-year lag implies that the matched patent sample begins in priority year 2005.

Since $y_{p,c}$ is a fractional outcome bounded between 0 and 1, we use a fractional logit specification (Papke & Wooldridge, 1996) so that the fitted conditional mean is constrained to the unit interval.

For a patent–technology observation (p, k, c, t), our baseline specification is:

$$\mathbb{E}\left[y_{p,c} \mid k, c, t, \mathbf{Z}_p\right] = \Lambda\left(\beta \log(1 + D_{k,c,t-5}) + \mathbf{Z}_p\right), \tag{5}$$

where $\Lambda(u) = (1 + \exp(-u))^{-1}$ and $\mathbf{Z}_p$ collects fixed effects and trends. The specification is estimated on the stacked patent–technology sample, with weights

$$w_{p,k} = s_{p,k} \cdot \text{cites}_p, \tag{6}$$

where cites_p is the citation-weight measure used throughout the descriptive analysis. This weight is intended to capture differences in the salience of patented inventions, while $s_{p,k}$ ensures that a patent with multiple deployable tags does not mechanically receive higher total weight.

All specifications include assignee-country-by-priority-year fixed effects, $\eta_{c,t}$, which absorb any time-varying determinants of patent composition that are common over time to assignees located in the same country. Relative to simple priority-year fixed effects, these controls flexibly net out country-specific policy environments, macroeconomic conditions, and shifts in the mix of assignees within country over time. Our baseline model additionally includes FEs for the patent's CPC 4-digit green class, $g(p)$, and some specifications add FEs for the patent's primary CPC 2-digit class. More demanding specifications allow for flexible evolution within broad technology groupings by including CPC 4-digit green-by-year and primary CPC 2-digit-by-year FEs. Our tightest specification further conditions on CPC 6-digit green FEs.

Inference allows for dependence along both the technology and country dimensions. Standard errors are clustered multi-way by assignee country and by the strictest CPC classification used in each specification. For example, in columns that use CPC 4-digit green fixed effects, we cluster by CPC 4-digit green class and assignee country. This clustering scheme accommodates serial correlation within countries and persistent unobservables within technology classes, which derives from the deployment regressor varying at the country-by-technology-by-year level.

The identifying variation in (5) comes from differences across deployable technologies in lagged deployment measured in the assignee's country. With country-by-year fixed effects, identification comes from relative deployment across technologies within the same assignee country–year, rather than from cross-country differences in overall deployment levels. A causal interpretation would require that, conditional on the fixed effects structure and the five-year lag, technology-specific changes in domestic deployment are orthogonal to unobserved determinants of patent-level labor orientation for assignees in that country. We interpret this specification as an induced-innovation style conditional association in which domestic deployment proxies for the market environment faced by assignees located in that country, rather than as a design that isolates exogenous deployment shocks.

5 Results

This section asks whether the scaling of deployable green technologies is accompanied by a systematic change in the labor content of green invention. Table 1 relates the labor-orientation index y to measures of technology deployment in the patent assignee's country. In our technology-country design, higher domestic deployment is associated with a lower labor-augmenting share of green patents, consistent with a shift in patent content toward more task-substituting innovation as deployment expands. While this relationship holds under increasingly strict sets of fixed effects, the evidence remains correlational at this stage given the absence of an explicitly exogenous deployment shock.

5.1 Deployment and labor orientation in the technology-country design

Table 1 reports patent-level fractional logit regressions of y_p on lagged log deployment for deployable green technologies. We match each patent to deployment of the corresponding technology in the assignee's country, lagged by five years relative to the patent priority year. Across all four specifications, the coefficient on $\log(1 + \text{deployment})$ is negative and statistically significant.

In the baseline specification with CPC4-green fixed effects and assignee-country-by-year fixed effects (column 1), we estimate a coefficient of -0.0072, statistically significant at the 1 percent level. The most granular specification, which includes CPC6-green fixed effects and retains CPC2-by-year and assignee-country-by-year fixed effects (column 4), yields a slightly less precisely estimated yet statistically significant coefficient of -0.0062. This corresponds to a shift in the task content of invention within deployable green technologies towards labor-automating, relative to the overall dispersion in the index across broad technologies (between 0.16 and 0.18 percentage point per doubling of deployment). This magnitude is economically meaningful considering that i) the deployment of most of the technologies considered in our analysis is expected to increase by more than an order of magnitude over their diffusion cycle and ii) our labor orientation index only covers around 5 percentage points over the whole spectrum of broad technologies. These results suggests that diffusion is accompanied by a shift in the task content of invention within deployable green technologies. The negative relationship is robust to progressively stricter controls for time-varying country conditions and technology trends.

Columns (2)–(4) progressively tighten identification by comparing relative deployment across technologies within the same country and year. The identifying variation in columns (2) through (4) comes from variation in domestic deployment across country $\times$ technology $\times$ year cells, after absorbing common shocks at the assignee-country-by-year level and progressively finer technology controls. The country-by-year fixed effects absorb any factors common across technologies, including aggregate demand conditions and policy environments. Column (3) goes further by absorbing technology-by-year FEs (at both CPC4-green and CPC2 levels), so identification comes from relative deployment intensity of a given technology within a country-year, net of global technology-year forces. The persistence of a negative coefficient under this saturation suggests that the baseline

Table 1: Deployment and the labor orientation of green patents: technology-country design

	(1)	(2)	(3)	(4)
log(1 + deployment)	-0.007***	-0.006***	-0.006**	-0.006**
	(0.003)	(0.002)	(0.003)	(0.002)
Assignee country FE x Year FE	Yes	Yes	Yes	Yes
Green CPC 4-digit FE	Yes	Yes	No	No
Primary CPC 2-digit FE	No	Yes	No	No
CPC 2-digit x Year FE	No	No	Yes	Yes
Green CPC 4-digit x Year FE	No	No	Yes	No
Green CPC 6-digit FE	No	No	No	Yes
Observations	178,224	178,224	178,195	178,198
AIC	95,645	95,615	96,585	96,412

Notes: Each observation is a U.S. patent mapped to at least one deployable green technology (renewables and EV-related CPC subclasses) and matched to technology-specific deployment in the patent assignee's country. The dependent variable is the patent-level labor-orientation index $y_p \equiv A_p/(A_p + S_p)$, where A_p and S_p are the labor-augmenting and labor-saving exposure scores described in Section 2 and based on Autor et al. (2024). The regressor is $\log(1 + \text{deployment})$, where deployment is measured in physical units at the country × technology × year level and lagged by five years relative to the patent priority year. Estimation is fractional logit (logit link); coefficients are log-odds. Observations are citation-weighted and split evenly across technologies when a patent maps to multiple deployable technologies. Standard errors are clustered by CPC4-green and assignee country in column (1), by CPC4-green, CPC2, and assignee country in columns (2)–(3), and by CPC6-green, CPC2, and assignee country in column (4).

relationship is not driven only by common time trends in particular technologies or by cross-country time-varying factors that affect all deployable technologies.

The coefficient remains negative even with CPC6-green fixed effects, suggesting the result is not driven by compositional reweighting across subfields. The stability of the negative coefficient in the specification with CPC6-green fixed effects (column 4) is informative because CPC6 partitions deployable technologies into relatively narrow subclasses. If the baseline relationship were driven primarily by reweighting across heterogeneous subfields within broader green categories, the coefficient would be expected to attenuate once CPC6-green fixed effects absorb that cross-subfield composition. The point estimate instead remains similar, consistent with an association between scale and the direction of patent text within the deployable green technology set.

The negative association between deployment and y_p is consistent with an induced direction-of-innovation interpretation in which scale shifts the private return from frontier exploration towards incremental refinements that economize on inputs, including labor. Induced-innovation studies in environmental economics typically focus on the level of clean innovation, using energy prices or policy variation to generate demand-pull for clean technologies (Acemoglu et al., 2012; Aghion et al., 2016; Calel & Dechezleprêtre, 2016; Johnstone et al., 2010; Popp, 2002). In those settings, expansion in the market for a clean technology raises innovative effort in that domain. Table 1 adds

a different margin: conditional on innovating in a deployable green technology, higher deployment in the assignee's country is associated with a shift in the content of that innovation towards labor-saving, as reflected in a lower y_p.

This pattern fits a technology-maturity channel that is common to empirical descriptions of diffusion and industry evolution. As technologies move from early adoption to large-scale deployment, innovation often reallocates from product variety and performance towards manufacturability, reliability, and cost reduction (Comin & Hobijn, 2010; Klepper, 1996). Many cost-reducing improvements in deployable technologies plausibly take a labor-saving form because labor remains an adjustable and often local component of production, installation, and operations. Our evidence here remains limited to patent labor orientation, and we do not attempt to identify which tasks are targeted. Our results nonetheless indicate that the orientation of green patenting is not invariant to diffusion: higher domestic deployment correlates with a more automation-oriented innovation in the patents that are produced.

These results also connect directly to recent work measuring automation-oriented innovation in patent data. Hémous et al. (2024) and Dechezleprêtre et al. (2023) show that firms respond to incentives by shifting patenting towards automation technologies. Table 1 provides an analogous association within green technologies using a task-based measure of labor orientation. The direction of the relationship is also consistent with evidence from climate-patent text that highlights labor-saving language in the joint green and digital transition (Bianchini et al., 2023; Rughi et al., 2023). Because y_p is a relative index, the result should be read as a change in the balance between automation-oriented and augmentation-oriented innovation within green patenting, rather than a claim that the modal green patent is automation or that automation dominates the green innovation agenda.

The labor-market implication is therefore indirect. The automation versus augmentation framework emphasizes that technologies can both displace tasks and generate new tasks, with the net employment effect depending on the relative strength of those margins (Acemoglu & Restrepo, 2019; Autor et al., 2024). A decline in y_p with deployment suggests that, within deployable green technologies, the innovation frontier shifts towards designs and processes that are more closely aligned with the language of task substitution than with the language of new-work creation. This is relevant to ongoing debates about "green jobs" because those debates often treat the labor intensity of the transition as fixed or policy-determined. The evidence we provide in this paper indicates that labor intensity can be endogenous to the stage of deployment through the direction of innovative effort, even holding the technology domain fixed.

5.2 Connection to the stylized facts

Here we show that the deployment regression evidence supports and helps interpret the descriptive patterns in Figures 1–2. Figure 1 shows that green categories span a wide range of labor-augmenting shares. This heterogeneity implies that statements about "green innovation" and automation are

sensitive to aggregation. The deployment analysis therefore focuses on a subset of green technologies where diffusion can be measured in physical units and where the underlying technologies are plausibly comparable across time.

Figure 2 shows that, within renewables and electric-vehicle technologies, the citation-weighted labor-augmenting share declines over time. Table 1 indicates that this decline is not mechanically a time effect in the subset of deployable technologies: conditional on year fixed effects and technology trends, higher deployment is associated with a lower labor-augmenting share. This matters for interpretation because it links the time pattern to a state variable with a clear economic meaning. Deployment is a proxy for market size and diffusion stage, so the regressions are consistent with a maturity-induced reorientation of innovation content.

5.3 Threats to interpretation and scope of claims

We discuss threats to interpretation to clarify the scope of inference: the results describe a robust conditional association, not a causal effect of deployment. A first concern is reverse causality. Deployment is partly a function of innovation, including the same patents whose text enters the dependent variable. In Table 1, the deployment measure is lagged by five years, which reduces mechanical simultaneity, but does not guarantee that deployment is predetermined with respect to unobserved determinants of patent orientation. If firms anticipate future deployment and redirect R&D effort in advance, the lag does not eliminate endogeneity. The negative coefficient should therefore be read as an association that is consistent with induced direction of innovation, not as a design-based causal estimate.

Omitted variable bias is a second concern. Policy bundles, commodity-price cycles, and global supply-chain shocks can move both deployment and innovation incentives. For example, a coordinated policy push can simultaneously accelerate adoption and increase the salience of cost reduction and manufacturability, which could change patent text orientation. The fixed effects and trends in Table 1 absorb many time-invariant technology differences and smooth drift, but they do not control for unobserved technology-year shocks that jointly affect deployment and innovation content.

Measurement imposes an additional boundary. The outcome y_p is a text-based index constructed from semantic proximity to task descriptions and new-work occupation titles. It is informative about the orientation of the innovation described in the patent, and Appendix Table B.1 provides external validation against an automation patent classification. It is not a direct measure of realized labor displacement or job creation. The appropriate inference is narrower: the content of green patenting shifts towards innovation that is closer to task substitution as deployment rises.

6 Conclusion

Green patenting spans technologies with very different task content. We measure the labor orientation of green inventive activity using U.S. patents with priority years 1980 to 2022 and CPC climate

tags (Y02 and Y04S). For each patent, we combine the automation and augmentation exposure scores produced by our implementation of the patent-text approach of Autor et al. (2024) and summarize orientation by the labor-augmenting share $y_p \equiv A_p/(A_p + S_p) \in [0, 1]$. We find that green categories are dispersed across the overall patenting distribution: green ICT and smart grids have higher y_p, while climate adaptation and green transportation sit nearer the middle or lower end. The spread across green categories is about five percentage points in citation-weighted averages, similar to the dispersion across broad non-green technology classes. "Green" status is therefore not a sufficient characteristic to determine whether inventive activity is closer to task substitution or to the creation of new work. Within deployable technologies, we also observe that the average y_p for renewables and electric-vehicle-related patents declines over time.

We then relate these descriptive patterns to diffusion. Table 1 matches patents in deployable technologies to global deployment measured in physical units and estimates fractional-logit models of y_p on lagged log deployment. Higher deployment five years earlier predicts a lower labor-augmenting share. Across specifications that include year fixed effects, green technology fixed effects, technology-specific trends, and, in the most saturated case, CPC 6-digit fixed effects, the coefficient on log deployment remains negative (around -0.006) and statistically significant.

One interpretation is a maturity channel. As a technology moves from frontier experimentation to large-scale deployment, private returns may shift towards process engineering, manufacturability, reliability, and cost reduction. Many of these improvements can be implemented through labor-saving design changes, mechanization, and tighter process control. We do not attempt to identify which tasks are affected, and y_p is a relative measure rather than a count of automation patents, but the deployment results indicate that diffusion is associated with a reorientation of innovation towards task substitution within the same green technology classes.

Our current evidence does not establish a causal effect of deployment on innovation content. Deployment is an equilibrium outcome and can respond to innovation, so lagging deployment reduces mechanical simultaneity but does not rule out reverse causality, anticipation, or omitted technology-year shocks. Policy packages, input prices, and supply-chain constraints can jointly shift adoption and the incentives that shape patent text, and the specifications do not include direct measures of these shocks. The outcome is also a text-based index, not an estimate of realized job creation or displacement. Thus a decline in y_p cannot be immediately interpreted as resulting in fewer green jobs.

The results still discipline claims about the labor intensity of the green transition. Forecasts of labor demand that treat task content as fixed can miss an endogenous margin: as deployment rises, inventive effort can tilt towards labor saving even within a narrowly defined technology. Two extensions are immediate. First, linking patent-based orientation to downstream outcomes in job postings, occupational transitions, and firm employment would quantify how much of the adjustment operates through task displacement versus new task creation. Second, isolating plausibly exogenous deployment variation, for example from staggered policy adoption or resource shocks, would separate

induced changes in innovation direction from correlated trends. Finally, extending the deployment-linked analysis beyond renewables and electric-vehicle technologies, where measurement is currently feasible, would clarify whether the maturity pattern generalizes to other parts of climate innovation.

References

Acemoglu, D. (1998). Why do new technologies complement skills? directed technical change and wage inequality. *The quarterly journal of economics*, *113*(4), 1055–1089.

Acemoglu, D., Aghion, P., Bursztyn, L., & Hemous, D. (2012). The environment and directed technical change. *American Economic Review*, *102*(1), 131–166. https://doi.org/10.1257/aer.102.1.131

Acemoglu, D., & Autor, D. (2011). Chapter 12 - skills, tasks and technologies: Implications for employment and earnings. In D. Card & O. Ashenfelter (Eds.), *Handbook of labor economics* (pp. 1043–1171, Vol. 4). Elsevier. https://doi.org/https://doi.org/10.1016/S0169-7218(11)02410-5

Acemoglu, D., & Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, *108*(6), 1488–1542.

Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of economic perspectives*, *33*(2), 3–30.

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

Atalay, E., Phongthiengtham, P., Sotelo, S., & Tannenbaum, D. (2020). The evolution of work in the united states. *American Economic Journal: Applied Economics*, *12*(2), 1–34.

Autor, D., Chin, C., Salomons, A., & Seegmiller, B. (2024). New Frontiers: The Origins and Content of New Work, 1940–2018*. *The Quarterly Journal of Economics*, *139*(3), 1399–1465. https://doi.org/10.1093/qje/qjae008

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, *118*(4), 1279–1333. https://doi.org/10.1162/003355303322552801

Bianchini, S., Damioli, G., & Ghisetti, C. (2023). The environmental effects of the "twin"green and digital transition in european regions. *Environmental and Resource Economics*, *84*(4), 877–918. https://doi.org/10.1007/s10640-022-00741-7

Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, *5*, 135–146. https://doi.org/10.1162/tacl_a_00051

Bowen, A., Kuralbayeva, K., & Tipoe, E. L. (2018). Characterising green employment: The impacts of 'greening' on workforce composition. *Energy Economics*, *72*, 263–275. https://doi.org/https://doi.org/10.1016/j.eneco.2018.03.015

Calel, R., & Dechezleprêtre, A. (2016). Environmental policy and directed technological change: Evidence from the european carbon market. *Review of Economics and Statistics*, *98*(1), 173–191. https://doi.org/10.1162/REST_a_00470

Castellanos, K. A., & Heutel, G. (2019). *Unemployment, labor mobility, and climate policy* (tech. rep.). National Bureau of Economic Research.

Comin, D., & Hobijn, B. (2010). An exploration of technology diffusion. *American Economic Review*, *100*(5), 2031–2059. https://doi.org/10.1257/aer.100.5.2031

Consoli, D., Marin, G., Marzucchi, A., & Vona, F. (2016). Do green jobs differ from non-green jobs in terms of skills and human capital? *Research Policy*, *45*(5), 1046–1060.

Curtis, E. M., & Marinescu, I. (2023). Green energy jobs in the united states: What are they, and where are they? *Environmental and Energy Policy and the Economy*, *4*, 202–237. https://doi.org/10.1086/722677

Curtis, E. M., O'Kane, L., & Park, R. J. (2023). *Workers and the green-energy transition: Evidence from 300 million job transitions* (Working Paper No. 31539). National Bureau of Economic Research.

Dechezleprêtre, A., Hémous, D., Olsen, M., & Zanella, C. (2023). Induced automation innovation: Evidence from firm-level patent data. *University of Zurich, Department of Economics, Working Paper*.

Elliott, R. J., Kuai, W., Maddison, D., & Ozgen, C. (2024). Eco-innovation and (green) employment: A task-based approach to measuring the composition of work in firms. *Journal of Environmental Economics and Management*, *127*, 103015.

Frattini, F. F., Vona, F., Bontadini, F., & Colantone, I. (2025). The local job multipliers of green industrialization.

Goldin, C. D., & Katz, L. F. (2009). *The race between education and technology*. harvard university press.

Google. (2017). Google patents public datasets [Data accessed through the Google Patents public data infrastructure].

Hafstead, M. A., & Williams III, R. C. (2018). Unemployment and environmental regulation in general equilibrium. *Journal of Public Economics*, *160*, 50–65.

Haščič, I., & Migotto, M. (2015). *Measuring environmental innovation using patent data* (OECD Environment Working Papers No. 89). OECD.

Hémous, D., Olsen, M., Dechezleprêtre, A., & Zanella, C. (2024). Automating labor: Evidence from firm-level patent data. *Journal of Political Economy*. Unpublished.

Hicks, J. (1963). *The theory of wages*. Springer.

Johnstone, N., Haščič, I., & Popp, D. (2010). Renewable energy policies and technological innovation: Evidence based on patent counts. *Environmental and Resource Economics*, *45*(1), 133–155. https://doi.org/10.1007/s10640-009-9309-1

Kalyani, A., Bloom, N., Carvalho, M., Hassan, T., Lerner, J., & Tahoun, A. (2025). The diffusion of new technologies. *The Quarterly Journal of Economics*, *140*(2), 1299–1365.

Klepper, S. (1996). Entry, exit, growth, and innovation over the product life cycle. *American Economic Review*, *86*(3), 562–583.

Kogan, L., Papanikolaou, D., Schmidt, L. D., & Seegmiller, B. (2023). *Technology and labor displacement: Evidence from linking patents with worker-level data* (Working Paper No. 31846). National Bureau of Economic Research. https://doi.org/10.3386/w31846

Marin, G., & Vona, F. (2019). Climate policies and skill-biased employment dynamics : evidence from EU countries. *Journal of Environmental Economics and Management*, *98*, 102253.

Papke, L. E., & Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics*, *11*(6), 619–632.

Popp, D. (2002). Induced innovation and energy prices. *American Economic Review*, *92*(1), 160–180. https://doi.org/10.1257/000282802760015658

Popp, D., Pless, J., Hascic, I., & Johnstone, N. (2020). Innovation and entrepreneurship in the energy sector. In *The role of innovation and entrepreneurship in economic growth*. University of Chicago Press.

Popp, D., Vona, F., Marin, G., & Chen, Z. (2022). The employment impact of a green fiscal push: Evidence from the american recovery and reinvestment act. *Brookings Papers on Economic Activity*.

Probst, B., Touboul, S., Glachant, M., & Dechezleprêtre, A. (2021). Global trends in the invention and diffusion of climate change mitigation technologies. *Nature Energy*, *6*, 1077–1086.

Rughi, T., Staccioli, J., & Virgillito, M. E. (2023). *Climate change and labour-saving technologies: The twin transition via patent texts* (LEM LEM Working Paper Series No. 2023/11). LEM Laboratory of Economics and Management, Institute of Economics, Scuola Superiore Sant'Anna.

Saussay, A., Sato, M., Vona, F., & O'Kane, L. (2025). Emerging skills and wage gaps in job ad data [No. 381].

Vona, F., & Consoli, D. (2015). Innovation and skill dynamics: A life-cycle approach. *Industrial and Corporate Change*, *24*(6), 1393–1415.

Vona, F., Marin, G., & Consoli, D. (2019). Measures, drivers and effects of green employment: evidence from US local labor markets, 2006–2014. *Journal of Economic Geography*, *19*(5), 1021–1048. https://doi.org/10.1093/jeg/lby038

Vona, F., Marin, G., Consoli, D., & Popp, D. (2018). Environmental Regulation and Green Skills: An Empirical Exploration. *Journal of the Association of Environmental and Resource Economists*, *5*(4), 713–753. https://doi.org/10.1086/698859

Weber, J. G. (2020). How should we think about environmental policy and jobs? an analogy with trade policy and an illustration from us coal mining. *Review of Environmental Economics and Policy*, *14*(1), 44–66.

A Labor orientation measurement details

This appendix provides additional detail on the construction of Aut_p, Aug_p, and y_p.

A.1 Text preprocessing and document construction

The patent document x_p concatenates three fields: (i) the patent title, (ii) the full title of the patent's main CPC-6 class, and (iii) the abstract. Missing fields are replaced with empty strings so that patents are not dropped mechanically due to incomplete metadata.

Tokenization uses a word-based tokenizer with punctuation, symbols, and numbers removed. Tokens are lowercased and common English stopwords are removed. Hyphenated terms are not mechanically split at hyphens. No stemming or lemmatization is applied. This choice is motivated by the prevalence of compound technical terms in patents and by the use of subword embeddings, which already encode morphological structure.

A.2 Embedding model and motivation for subword information

Word vectors v_w are obtained from a pretrained fastText model with subword information (Bojanowski et al., 2017). fastText represents a token as a sum of vectors associated with character n-grams, yielding a vector for tokens that are rare or unseen in the training vocabulary. This feature is especially relevant for patents, where technical terms, concatenations, and domain-specific acronyms are frequent and often absent from fixed-vocabulary embeddings. In the absence of subword information, an implementation based on fixed dictionaries would require dropping out-of-vocabulary tokens or collapsing them into an undifferentiated unknown token, potentially weakening semantic matching between patent text and occupational reference texts.

A.3 TF-IDF weighting and patent embeddings

Let N denote the number of patent documents and let $\text{df}(w)$ denote the number of documents containing token w. The TF-IDF weights ω_{pw} are computed using proportional term frequency and inverse document frequency with a natural-log scaling:

$$\omega_{pw} \propto \left(\frac{\text{tf}(w,p)}{|p|}\right) \log\left(\frac{N}{\text{df}(w)}\right) \tag{7}$$

where $\text{tf}(w,p)$ is the token count and $|p|$ is the number of tokens in patent p after preprocessing. Patent embeddings are constructed as in (1) and normalized to unit length. Normalization ensures that similarities are cosine similarities and prevents mechanically longer documents from having higher similarity.

A.4 Reference corpora and time variation

The automation reference corpus consists of occupational task statements (Dictionary of Occupational Titles). The augmentation reference corpus consists of occupation titles obtained from the U.S. Census Alphabetical Index of Occupations, which are intended to proxy for "new work," following Autor et al. (2024). Reference texts are embedded into unit-length vectors using the same embedding model. The augmentation reference set varies across broad time windows to reflect changes in the universe of new-work titles; patent-year similarity is computed against the appropriate vintage based on the patent publication year.

A.5 Similarity normalization, thresholds, and exposure construction

For each publication year, we compute similarity matrices between patent embeddings and reference embeddings for each reference set. Similarities are median-demeaned within patent as in (2). This step removes patent-specific baseline similarity that can arise from generic terms shared across documents and reference texts.

Thresholds $\tau_{\text{aut}}(\cdot)$ and $\tau_{\text{aug}}(\cdot)$ are computed separately for each reference set and within publication-year blocks, using the empirical 0.85 quantile of the pooled distribution of median-demeaned similarities across all patents in the block and all reference texts. Exposures Aut_p and Aug_p are counts of reference texts whose similarity exceeds the relevant threshold, as in (3). The labor-orientation index y_p in (4) is undefined when $\text{Aut}_p = \text{Aug}_p = 0$; in that case, the patent has no high-similarity matches to either reference set under the chosen thresholds, and the implementation treats y_p as missing since a ratio is required.

A.6 Scope and interpretation

The exposure counts and ratio are constructed for the full U.S. patent corpus used in the embedding pipeline and can then be merged to the green subsample via publication identifiers. The measure is designed to capture textual orientation towards task-replacing versus new-work-related content. It does not identify the adoption of the patented technology, the identity of affected occupations, or general equilibrium employment effects. Measurement error is possible, given that patents may describe enabling technologies without explicit task language, and that occupational reference texts are an imperfect stand-in for the set of tasks displaced or created by innovation. When y_p is used as an outcome, such error implies that estimates should be interpreted as relating observables to a noisy proxy for innovation orientation rather than to realized labor market impacts.

A.7 Additional descriptive exhibits

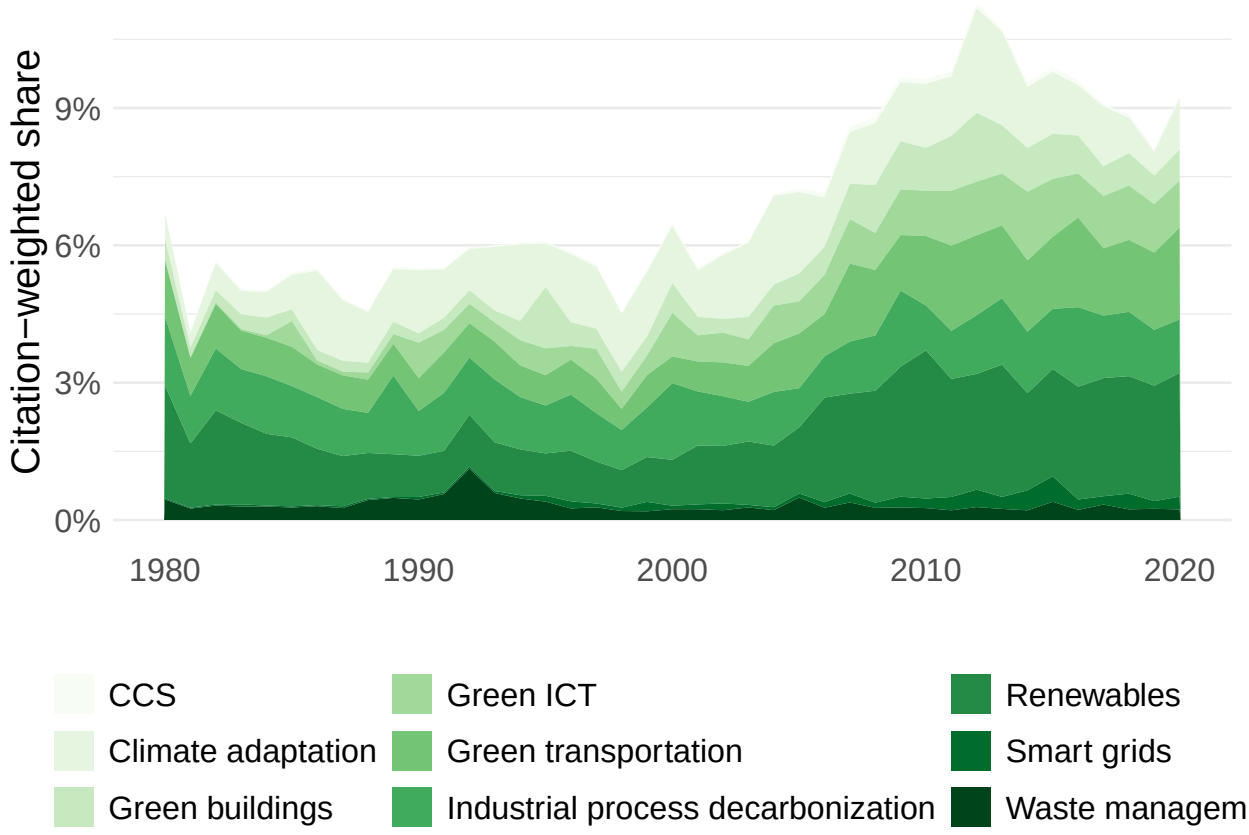


Figure A.1: Composition of green patenting over time

Notes: The figure plots, by priority year, the citation-weighted share of patents in each green CPC4 category relative to total patent citations in the same year. Green categories are defined by CPC climate tags at the four-digit level (Y02A, Y02B, Y02C, Y02D, Y02E, Y02P, Y02T, Y02W, and Y04S), mapped to the labels shown in the legend. Each shaded region is a category-specific share; the sum across regions in a year equals the total citation-weighted share of green patents in that year.

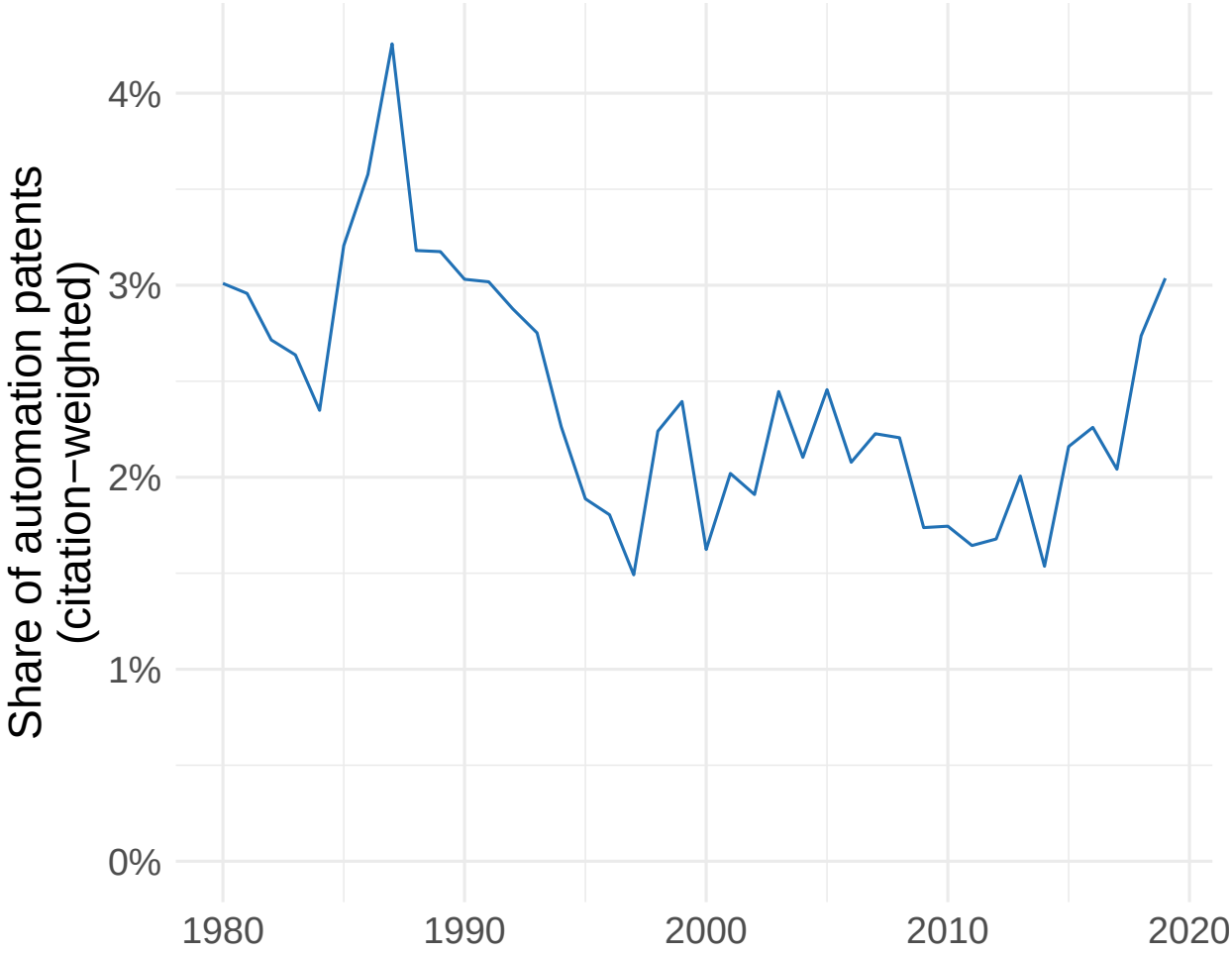


Figure A.2: Automation patents over time under Hémous et al. (2024)

Notes: The figure plots, by year, the citation-weighted share of patents classified as automation patents under the definition in Hémous et al. (2024). The underlying series is computed by implementing Hémous et al.'s (Hémous et al., 2024) keyword and CPC codes-based classification on our own patent dataset.

B Supplementary evidence

Table B.1: Relationship between the labor-orientation index and an external automation classification

	(1)	(2)	(3)
Automation patent (Hemous *et al.*, 2024)	-0.018***	-0.033**	-0.035**
	(0.003)	(0.013)	(0.016)
Year FE	✓	✓	✓
CPC 4-digit FE	✓	✓	✓
CPC 6-digit FE		✓	✓
CPC 4-digit FE × Year			✓
Observations	148,395	148,395	148,386

* p <0.1, ** p <0.05, *** p <0.01

Notes: The dependent variable is the patent-level labor-orientation index, defined as the share of augmenting exposure in the sum of augmenting and automating exposure scores (Autor et al., 2024). The key regressor is an indicator for whether the patent is classified as an automation patent under Hémous et al. (2024). Columns add increasingly demanding sets of technology and time fixed effects and, where indicated, technology-specific time controls. Standard errors and clustering follow the table output. A negative coefficient implies that automation-classified patents have lower labor-augmenting shares under the continuous index, as expected.

Table B.2: Deployable technology definitions and matched patent counts

Technology	CPC	CPC name	Patent count
Biogas, Liquid and Solid biofuels	Y02E50/1	Biofuels	34,680
Marine energy	Y02E10/3	Energy from the sea	6,332
Mixed Hydro Plants and Renewable hydropower	Y02E10/2	Hydro energy	6,315
Offshore and Onshore wind energy	Y02E10/7	Wind energy	47,765
Renewable municipal waste	Y02E50/3	Fuel from waste	24,653
Solar photovoltaic	Y02E10/5	Solar PV	112,860
Solar thermal energy	Y02E10/4	Solar thermal energy	18,900
PHEV	Y02T10/62	Hybrid vehicles	29,684
BEV	Y02T90/14	Plug-in electric vehicles	39,699
FCEV	Y02T90/40	Hydrogen vehicles	10,861

Notes: The table documents the crosswalk used to match patents to deployable green technologies with observed deployment series. Technologies are defined using CPC climate-tag subclasses within Y02E (energy generation) and Y02T (transport). A patent is assigned to a deployable technology if its CPC list contains the corresponding CPC pattern (for example, solar photovoltaic patents are those with at least one CPC code beginning with Y02E10/5). Patent counts are computed from the U.S. patent sample used in the paper and reflect the number of patents that meet each CPC definition. Because a single patent can contain CPC codes from multiple deployable technologies, counts can overlap across rows and do not necessarily sum to the number of unique patents in the matched sample.

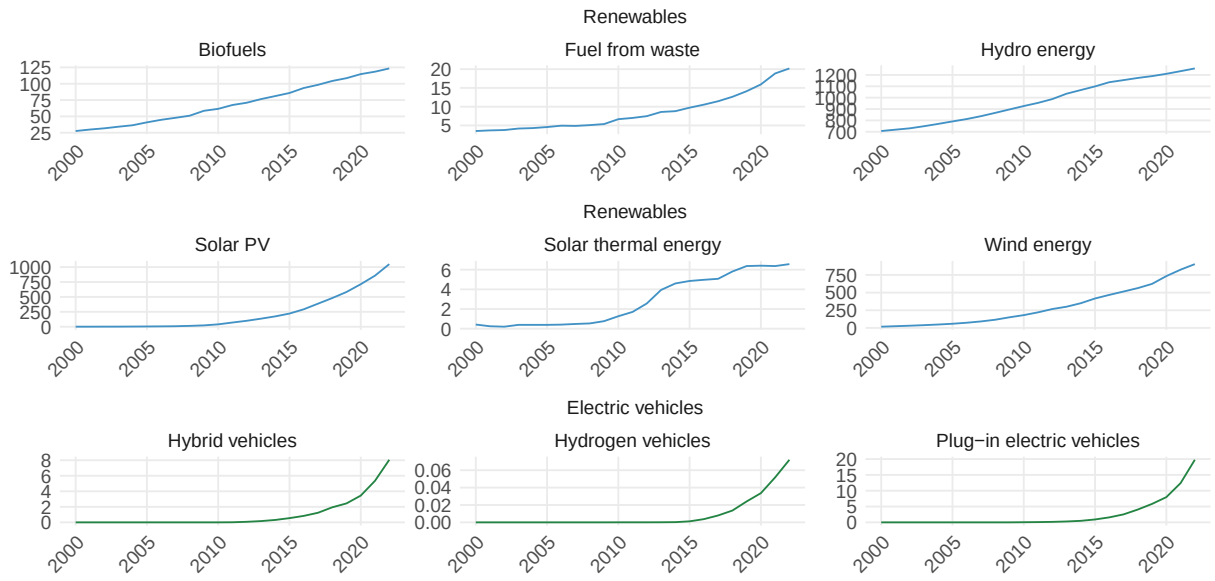


Figure B.1: Global deployment of deployable green technologies, 2000–2022

Notes: The figure plots annual global deployment for each technology observed in our dataset. Renewable energy deployment is measured as installed capacity in megawatts from International Renewable Energy Agency statistics; electric-vehicle deployment is measured as the global stock of vehicles, in millions, from International Energy Agency data, excluding two- and three-wheelers. Each panel corresponds to a single technology, and the vertical scales differ across panels because technologies are measured in different physical units and have different diffusion trajectories.