# The Employment Impact of Emerging Digital Technologies*

Ekaterina Prytkova† Fabien Petit‡ Deyu Li§
Sugat Chaturvedi¶ Tommaso Ciarli‖

February 2024

**Abstract**

This paper measures the exposure of industries and occupations to 40 digital technologies that emerged over the past decade and estimates their impact on European employment. Using a novel approach that leverages sentence transformers, we calculate exposure scores based on the semantic similarity between patents and ISCO-08/NACE Rev.2 classifications to construct an open–access database, 'TechXposure'. By combining our data with a shift–share approach, we instrument the regional exposure to emerging digital technologies to estimate their employment impact across European regions. We find an overall positive effect of emerging digital technologies on employment, with a one-standard-deviation increase in regional exposure leading to a 1.069 percentage point increase in the employment-to-population ratio. However, upon examining the individual effects of these technologies, we find that smart agriculture, the internet of things, industrial and mobile robots, digital advertising, mobile payment, electronic messaging, cloud storage, social network technologies, and machine learning negatively impact regional employment.

**Keywords:** Occupation Exposure; Industry Exposure; Text as Data; Natural Language Processing; Sentence Transformers; Emerging Digital Technologies; Automation; Employment
**JEL Codes:** C81, O31, O33, O34, J24, O52, R23

*We are grateful to Florencia Jaccoud, Önder Nomaler and Maria Savona as well as to audiences at the ESCoE Conference 2023, GTM Conference, Eu-SPRI 2023, VSI 2023, AI and the Economy Conference, JRD 2023 Autumn London Meeting, EMAEE Workshop, Paris PILLARS Workshop, PILLARS Final Conference, Untangled Final Conference, UNU-MERIT Internal Conference 2023, Digital Catapult Workshop, CESifo Big Data Workshop, and seminar participants at JRC–European Commission, IFS-UCL-IoE Human Capital and Labour Markets Seminar, UK Department for Education, and PILLARS Expert Stakeholder Group for useful comments and suggestions. This project received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 101004703.

†Côte d'Azur University and University of Sussex, SPRU. Email: e.prytkova@sussex.ac.uk.
‡University College London, CEPEO. Email: f.petit@ucl.ac.uk. Website: fabienpetit.com.
§Utrecht University. Email: d.li1@uu.nl
¶Ahmedabad University. Email: sugat.chaturvedi@ahduni.edu.in.
‖United Nations University, UNU-MERIT and University of Sussex, SPRU. Email: ciarli@merit.unu.edu

# 1 Introduction

The past decade has witnessed rapid technological advancements in digital technologies, including electric vehicles, self-driving cars, drones, wearable devices, artificial intelligence, augmented/virtual reality, internet of things, 3D printing, and blockchain. Understanding how occupations and industries are exposed to these emerging digital technologies is crucial to policymakers responsible for shaping education, training, and labor policies in anticipation of technological transitions. Furthermore, while there is a substantial body of evidence on the labor market impact of more established technologies, such as Information and Communication Technologies (ICT) and industrial robots, little is still known about how these emerging digital technologies affect employment.[1]

This paper measures the exposure of industries and occupations to 40 digital technologies that emerged over the past decade and estimates their impact on regional employment in Europe. Our set of technologies includes a wide spectrum of digital technologies encompassing the most significant innovations between 2012 and 2021. We cover both tangible and intangible technologies, including 3D printing and additive manufacturing, machine learning and neural networks, recent advances in industrial automation (including robotics), the IoT, autonomous vehicles and drones, intelligent logistics, mobile and e-payment systems, cloud technologies, and e-learning, among others.

We contribute to the existing literature in two distinct ways. First, we introduce an innovative methodology, using state-of-the-art Natural Language Processing (NLP) tools such as sentence transformers, for computing the exposure of industries and occupations to emerging digital technologies. This approach is based on the semantic similarity between patents and descriptions in standard classification systems, namely NACE (Statistical Classification of Economic Activities in the European Community) and ISCO (International Standard Classification of Occupations). The outcome of this methodology is the 'TechXposure' database, a pioneering resource that we have made publicly available. This database stands out as the first of its kind, offering an unprecedented level of granularity in measuring the exposure of industries (up to the 3-digit level) and occupations (up to the 4-digit level) to a comprehensive set of 40 emerging digital technologies.

Second, we estimate the impact of emerging digital technologies on regional employment across Europe. In a similar setup to that of Acemoglu and Restrepo (2020), we use an instrumental variable (IV) shift-share approach based on industry exposure scores, derived from our

[1] See, for instance, Autor et al. (2003), Autor et al. (2006), Goos and Manning (2007), Goos et al. (2009, 2014), Michaels et al. (2014), Akerman et al. (2015) for the labor market consequences of the technological change related to ICT; and Graetz and Michaels (2018), Acemoglu and Restrepo (2020), Vries et al. (2020), Dauth et al. (2021), Aghion et al. (2021), for the labor market effects of industrial automation and industrial robots.

database, and the baseline employment shares of these industries in the region. This provides valuable insights into the labor market consequences of regional exposure to these technologies.

We start by grouping patents into technologies based on semantic similarity in their titles. We consider the sample of patents identified as novel in Chaturvedi et al. (2023). This sample includes the most significant digital innovations that appeared between 2012 and 2021. We convert the text of patent titles into vector representations, or *embeddings*,[2] using the pre-trained sentence transformer model *all-mpnet-base-v2* (Reimers and Gurevych 2019).[3] We apply k-means clustering on these embeddings, resulting in the identification of 40 distinct emerging digital technologies, each defined as a group of patents.

Our obtained set of technologies comprises 40 individual emerging digital technologies, which can be categorized into nine emerging technology families. These families are: 3D Printing; Embedded Systems, including industrial automation such as industrial robots and IoT; Smart Mobility, featuring intelligent logistics such as mobile robots and autonomous vehicles; Food Services; E-Commerce; Payment Systems, encompassing recent advancements in e-payment and mobile payment; Digital Services, which cover technologies ranging from cloud computing to e-learning and social networks; Computer Vision technologies, such as augmented reality/virtual reality (AR/VR) and machine learning; and HealthTech, including health monitoring devices and e-healthcare.

We compute the exposure of occupations and industries to these technologies based on the semantic connection between patents and the descriptions of 3-digit NACE Rev.2 and 4-digit ISCO-08 classifications. For each industry–patent and occupation–patent pair, we compute the cosine similarity score which reflects the degree of similarity between both documents. To enhance the matching quality, we introduce a filtering procedure that enables us to retain only the most relevant pairs. Once filtered, we aggregate the cosine similarity scores from individual patents to technologies under which they were clustered by taking the citation-weighted sum.

This methodology results in the creation of the 'TechXposure' database, which provides a measure of exposure to each of the 40 emerging digital technologies for every industry and

[2]Text embedding is a Natural Language Processing (NLP) technique used to transform text (words, sentences, documents) into a numerical representation, i.e., high-dimensional numerical vectors, commonly referred to as embeddings. See Gentzkow et al. (2019) for a comprehensive review of NLP applications in the economic literature.

[3]A sentence transformer is a specific architecture of a deep neural network. The features of this architecture enable the model to capture the contextual significance of words in a text and leverage the ensemble effect to produce embeddings. The sentence transformer model *all-mpnet-base-v2* is fine-tuned on over a billion sentence or paragraph pairs from academic papers, Wikipedia, and Stack Exchange, among others, and has shown state-of-the-art results on sentence similarity tasks (Reimers and Gurevych 2019).

occupation. Our exposure metric reflects the degree to which an industry or occupation is *relevant* to a specific technology. For industries, relevance is determined by the integration of a technology into the production process and/or if the technology enhances the output of an industry. For occupations, relevance pertains to the importance of a technology in performing tasks and functions inherent to an occupation.[4] These exposure scores are available for all levels within the ISCO-08 and NACE Rev.2 classifications.

We identify several insights regarding the exposure of industries and occupations to emerging digital technologies. For occupations, we find clerical support workers, plant/machine operators, and assemblers are the most exposed to emerging digital technologies, followed by high-paying occupations, including managers, professionals, technicians, and associate professionals. Additionally, we observe that manual occupations are more exposed to *tangible* technology families, such as 3D Printing, Embedded Systems, and Smart Mobility, while cognitive occupations are more exposed to *intangible* technology families, such as Computer Vision, E-Commerce, Payment Systems, HealthTech, and Digital Services. We find a similar divide for industries, with agriculture, manufacturing industries, and services operating physical infrastructures, such as transportation and storage, being more exposed to tangible technologies as compared to other services which are more exposed to intangible technologies.

In building upon our new data, we estimate the causal effect of emerging digital technologies on regional employment. We employ an IV shift-share approach, using our industry exposure scores and the baseline employment shares of these industries, to instrument the exposure to emerging digital technologies at the regional level. Our methodology draws inspiration from Acemoglu and Restrepo (2020), who estimated the impact of robots on US local employment. However, our study diverges in its focus on European regions at the NUTS-2 level and includes a wider range of technologies, specifically 40 emerging digital technologies.

Our estimation starts with the overall impact of emerging digital technologies on the regional employment-to-population ratio from 2012 to 2019. Then, we conduct a more detailed analysis, assessing the effects of each of the nine technology families and the 40 individual technologies. In this latter process, we account for the potential regional co-integration of emerging digital technologies to disentangle their individual employment impact. The identification relies on two primary assumptions. First, we assume that regional industry shares are exogenous, conditional on observable factors. Second, we assume that regions with greater exposure to emerging digital technologies are not disproportionately affected by other labor

[4]Our exposure scores necessitate two clarifications regarding their interpretation. Firstly, they indicate the (contextual) relevance of each technology to a given industry or occupation, rather than their actual adoption. Secondly, they are neutral regarding the nature of the relationship between a technology and an industry/occupation. This means they do not specify whether the technology and industry/occupation are complementary or substitutive in producing output.

market shocks or trends.

We interpret our estimates using the canonical task-based framework of Acemoglu and Restrepo (2018, 2019). As emerging digital technologies develop and are adopted, they enable capital to substitute labor in a wider range of tasks. This may have three impacts on labor. First, new technologies may change the task content of production, reducing the role of labor and hence labor demand, leading to lower employment; this is the *displacement effect*. Second, new technologies may enhance worker productivity by allowing a more flexible task allocation, thereby increasing labor demand and employment; this is the *productivity effect*. Third, new technologies may create new tasks, consequently increasing labor demand and employment; this is the *reinstatement effect*.

Our work reveals several new findings. First, the overall impact of emerging digital technologies on regional employment is positive. We find that a one-standard-deviation increase in regional exposure leads to a 1.069 percentage point (pp.) change, corresponding to 2.1%, in the employment-to-population ratio from 2012 to 2019.

Second, the technology families of Smart Mobility and HealthTech have a positive impact on employment, whereas E-Commerce exhibits a negative impact. A one-standard-deviation increase in regional exposure to Smart Mobility results in a 0.62 pp. increase in the employment-to-population ratio. For HealthTech, this increase is 0.93 pp. Conversely, an equivalent increase in regional exposure to E-Commerce corresponds to a 1.54 pp. decline in the employment-to-population ratio, significant at the 10% level.

Third, we observe significant heterogeneity in the regional impact of individual technologies. We find that the displacement effect dominates for 3D Printing (-1.28 pp.), Smart Agriculture (-1.38 pp.), IoT (-1.80 pp.), Industrial Automation (-3.11 pp.), Intelligent Logistics (-1.33 pp.), Digital Advertising (-3.11 pp.), Mobile Payment (-1.78 pp.), Electronic Messaging (-5.08 pp.), Cloud Storage (-1.99 pp.), Recommender Systems (-4.79 pp.), Social Networking (-1.65 pp.), Digital Media Content (-3.26 pp.), Machine Learning (-1.58 pp.), and E-Healthcare (-5.78 pp.). Conversely, we find positive and significant employment effects for Additive Manufacturing (+2.88 pp., at the 10% level), Energy Management (+1.37 pp.), Remote Monitoring (+1.32 pp.), Smart Home (+0.83 pp.), Parking Management (+5.42 pp.), Vehicle Telematics (+3.68 pp.), E-Trading (+0.93 pp., at the 10% level), E-Payment (+1.62 pp., at the 10% level), Gaming (+1.46 pp.), E-Learning (+1.43 pp.), Workflow Management (+1.60 pp.), Information Processing (+3.21 pp.), and Medical Imaging (+1.43 pp.). Lastly, we find no employment effect for 3D Printer Hardware, Autonomous Vehicles, Passenger Transportation, Food Ordering, Online Shopping, E-Coupons, Digital Authentification, Location-Based Services, Voice Communication, Cloud Computing, AR/VR, Health Monitoring, and Medical Information.

Our work contributes to several strands of the literature, particularly in identifying the occupations and industries exposed to technological change. While previous studies have measured the exposure of employment to specific technologies such as robots,[5] broadly defined technologies such as Artificial Intelligence (AI),[6] or a mixed array of automation technologies,[7] our work stands out with its unique contributions.

First, our approach innovatively uses patents in this context. Although the use of patents to measure technical change is increasingly common, we are the first to define technologies as groups of patents clustered based on semantic distance. This novel method enables us to identify all digital technologies, not limited to general AI or robots, and provides a more precise and meaningful categorization of these technologies.

Second, we introduce a scalable and advanced methodology using state-of-the-art NLP techniques with sentence transformers. This methodology is universally applicable, bypassing the need for identifying specific keywords, or *tokens*, as it leverages text similarity, thereby requiring only a relevant set of patents.

Third, our work is pioneering in developing and making publicly available the 'TechXposure' database. This new database, which includes information about how different industries and occupations are exposed to emerging digital technologies, is set to help open up many new paths in analyzing the impact of these technologies in several domains of the economy. Additionally, it will be updated periodically to ensure ongoing relevance and utility.

Fourth, our work uniquely addresses a gap in the literature regarding exposure metrics. While most existing metrics focus on US classifications, our work is the first to provide detailed

[5] Prior work has estimated the impact of industrial robots on occupations, industries, or regions using data from a unique source, namely, the International Federation of Robots (IFR). Graetz and Michaels (2018) reports no significant effect on total employment at the country level in Europe. Dauth et al. (2021) observe that job losses in manufacturing due to industrial robots are balanced by new service sector jobs in Germany. In line with these findings, Vries et al. (2020) show a rise in non-routine analytic jobs and a decline in routine manual jobs in Europe. On the contrary, Acemoglu and Restrepo (2020) find a negative effect on employment in US commuting zones.

[6] Webb (2019) develops an AI exposure score based on verb-noun token pairs and finds that occupations with high AI exposure experience employment declines. Using AI keywords in online job vacancies to proxy AI adoption, Acemoglu et al. (2022) find reduced hiring in firms that adopt AI, whereas Alekseeva et al. (2021) reports an overall increased demand for AI skills. Felten et al. (2018, 2021) use survey data to assess the task exposure of occupations to 10 AI applications, creating AI exposure metrics at the occupation, industry, and regional levels. Frey and Osborne (2017) estimate the probability of 'computerization' of occupations using expert predictions to label tasks as automatable, finding that 47% of total US employment is at high risk of computerization due to AI.

[7] Kogan et al. (2021) identify breakthrough innovations with patents from 1850, applying Kelly et al. (2021)'s methodology, to estimate occupational exposure to these innovations via a TFIDF token-based approach. They find a negative correlation between such exposure and future employment in those occupations. Dechezleprêtre et al. (2021) develop a measure of automation innovation in machinery by analyzing the frequency of specific keywords in patent texts from 1997. Mann and Püttmann (2023) distinguishes US patents filed from 1976 to 2014 into automation and non-automation categories, finding a positive impact on employment in local labor markets, primarily driven by growth in the service sector.

exposure scores for European classifications, specifically NACE Rev. 2 and ISCO-08, at a highly granular level. This contribution extends the applicability of exposure metrics beyond the US context, offering valuable insights for European regions. Additionally, our exposure scores are based on worldwide patents, thereby considering global advances in technologies that extend beyond the US and Europe.

Fifth, while our results align with existing literature regarding the negative impact of some automation technologies, such as industrial robots and AI, on employment, our work suggests that the excessive focus on these specific technologies could potentially overshadow the positive impacts of other emerging digital technologies on employment.[8] This is particularly pertinent when considering the crucial role of co-integration among these technologies in determining their effects on employment.

Most closely related to our work is the paper by Autor et al. (2022). Using changes in the US Census Bureau's classification of occupations that occur every decade, both in terms of occupation titles and tasks, they establish the impact of augmentation and automation innovations on the emergence of new work and occupational labor demand in the US over the period 1940 to 2018. Our paper complements their work in two important aspects. First, our analysis focuses on the latest emerging digital technologies which are likely to be key drivers of the ongoing emergence of new work and those of the future. Second, while they find an overall positive impact of technological innovations on employment, focusing on occupations, our empirical analysis, which builds upon our measure of industry exposure, finds a similar impact of emerging digital technologies on employment.

Our work also differs from theirs in a methodological aspect. First, we construct an exposure measure that is agnostic of the augmentation or automation aspect of the innovation and captures only the relevance of that innovation to the occupation or the industry. Using our empirical analysis, we then identify whether that technology increases or decreases employment, hence determining if it is augmenting or diminishing. A second distinction comes from clustering patents into technologies, allowing us to isolate the effect of specific emerging digital technologies. Although we find a similar overall positive effect on employment, we also unveil substantive heterogeneity at the technology level, suggesting that not all emerging technologies have the same impact in terms of augmentation and automation of labor.

The paper is organized as follows. Section 2 outlines our methodology for deriving our set of emerging digital technologies from patent data. Section 3 introduces our state-of-the-art NLP-based method for calculating industry and occupation exposure scores to these tech-

[8]Consistent with our results, Mann and Püttmann (2023), who use a broader definition of automation technology as compared to industrial robots or software, also find a positive effect of the latter on employment in local US labor markets.

nologies. Section 4 provides descriptive statistics regarding the exposure of industries and occupations to emerging digital technologies. Section 5 estimates the causal impact of these technologies on regional employment, using an IV shift-share approach. Section 6 concludes.

## 2 Emerging Digital Technologies

In this section, we derive and describe our 40 emerging digital technologies, where each technology is a group of patents from the Derwent Database. First, we describe the different parts of a patent's text and the properties of our patent sample. Second, we explain our methodology to cluster patents based on semantic similarity and obtain our set of 40 emerging technologies. Finally, we describe the technologies.

We use a set $\mathcal{P}$ of 190,714 Derwent patents, filed between 2012 and 2021. This patent set constructed by Chaturvedi et al. (2023) comprises the most novel patents related to digital innovations, together with the patents that follow their semantic trajectory, that is, the most similar patents filed in subsequent years. Each patent is a document that describes the invention, and how it differs from existing inventions. The information provided for each patent includes a title, an abstract, and additional metadata such as the list of applicants and inventors (i.e., companies or individuals), filing year and authority, citations, and codified technical areas according to various classifications (such as the International Patent Classification or IPC), among others. In turn, the abstract is divided into labeled topical blocks such as novelty, use, independent claims, description of drawings, etc.

We use the title to extract semantic information from the patent. Compared to the abstract, the title has two significant advantages for our analysis. First, the title is always present, compact, and follows certain structure. Patent titles in the Derwent Database are comprised of two parts. The first part provides a concise description of the technology in a phrase or short sentence; we denote this part $p_1 \in p$. The second part describes *how the technology functions*; we denote this part $p_2 \in p$. The division between the two parts is marked by the first **comma–verb combination**.[9]

Second, the language used to represent the technology and its function in the title conveys technical information through comprehensive descriptions rather than highly technical jargon (included in the abstract). These two properties of the patent title map well onto characteristics of industrial and occupation descriptions making the title the most suitable candidate to use for semantic matching between each pair.

[9] Using Part-of-Speech (POS) tagging, we identify that this pattern appears in 87.3% of our patent sample, represented by the following combinations: ', has', ', includes', ', involves', and ', comprises'. For the remaining patents, we divide the patent title at the word space closest to the middle of the document.

We provide three examples of patent titles present in our sample:

1. Method for targeting television advertisement based on profile linked to online device**, involves** *selecting television advertisement to be directed to set-top box based on profile information pertaining to user or online activity.* (Patent ID 2013B87254, 2013)
2. Vehicle intelligent logistics control device**, has** *GPS locating module for obtaining position information of transport vehicle through main control chip, RFID reader for reading RFID tag information, and 4G module connected with server.* (Patent ID 201713859U, 2017)
3. System for recognizing training speech**, has** *process or which is configured to increment counter associated with word sequences, and train language model of automatic transcription system using word sequences and counter.* (Patent ID 202048118D, 2020)

Using the entire titles from our patent sample $\mathcal{P}$, we obtain their embeddings $Emb_p$, each being a 768-dimensional vector, using the sentence transformer model all-mpnet-base-v2. This model is specifically trained for text similarity tasks (Reimers and Gurevych 2019). Then, we cluster the embeddings using the k-means algorithm and obtain 40 clusters, each of which we denote as our set of emerging digital technologies $k \in \mathcal{K}$. We find 40 clusters to be optimal after manually exploring partitions ranging from 15 to 45 clusters to identify a meaningful partition that maximizes the differences between technologies and minimizes the differences within them.[10]

Table 1 lists our set of emerging digital technologies grouped by technology families. We provide a short description for each technology in Tables A.1 to A.3 in Appendix A.1. It is important to note that the grouping of these 40 technologies into 9 families is based on the correlation between the technologies' co-occurrence in occupations (more in the next section). Thus, a family comprises technologies whose occupation structure of semantic links is highly correlated; see Appendix A.5 for a more detailed discussion. Figure A.1, in the appendix, presents the distribution of patents across emerging digital technologies.

[10]The Davies and Bouldin (1979) Index, which assesses the quality of a partition by considering both the compactness of clusters and the separation between them, suggests that 33 clusters are optimal. However, a visual examination shows that a partition into 33 clusters based on semantic similarity combines very distinct technologies. For instance, it groups intelligent healthcare management systems with general data processing, and also combines robotic automation and augmented reality within monitoring systems. These technologies are often listed among the most relevant future technologies and may have very different impacts on labor. For instance, Acemoglu and Restrepo (2019) argues that augmented reality is an "area in which the use of AI can significantly change the production process in a way that may be favorable to labor".

Table 1: List of Emerging Digital Technologies

| | Family | | Emerging Technology |
|---|---|---|---|
| F1 | 3D Printing | 01 | 3D Printer Hardware |
| | | 02 | 3D Printing |
| | | 03 | Additive Manufacturing |
| F2 | Embedded Systems | 04 | Smart Agriculture & Water Management |
| | | 05 | Internet of Things (IoT) |
| | | 06 | Predictive Energy Management and Distribution |
| | | 07 | Industrial Automation & Robot Control |
| | | 08 | Remote Monitoring & Control Systems |
| | | 09 | Smart Home & Intelligent Household Control |
| F3 | Smart Mobility | 10 | Intelligent Logistics |
| | | 11 | Autonomous Vehicles & UAVs |
| | | 12 | Parking and Vehicle Space Management |
| | | 13 | Vehicle Telematics & Electric Vehicle Management |
| | | 14 | Passenger Transportation |
| F4 | Food Services | 15 | Food Ordering & Vending Systems |
| F5 | E-Commerce | 16 | Digital Advertising |
| | | 17 | Electronic Trading and Auctions |
| | | 18 | Online Shopping Platforms |
| | | 19 | E-Coupons & Promotion Management |
| F6 | Payment Systems | 20 | Electronic Payments & Financial Transactions |
| | | 21 | Mobile Payments |
| | | 22 | Gaming & Wagering Systems |
| F7 | Digital Services | 23 | Digital Authentication |
| | | 24 | E-Learning |
| | | 25 | Location-Based Services & Tracking |
| | | 26 | Voice Communication |
| | | 27 | Electronic Messaging |
| | | 28 | Workflow Management |
| | | 29 | Cloud Storage & Data Security |
| | | 30 | Information Processing |
| | | 31 | Cloud Computing |
| | | 32 | Recommender Systems |
| | | 33 | Social Networking & Media Platforms |
| | | 34 | Digital Media Content |
| F8 | Computer Vision | 35 | Augmented and Virtual Reality (AR/VR) |
| | | 36 | Machine Learning & Neural Networks |
| | | 37 | Medical Imaging & Image Processing |
| F9 | HealthTech | 38 | Health Monitoring |
| | | 39 | Medical Information |
| | | 40 | E-Healthcare |

*Notes*: This table lists the 40 emerging digital technologies along with their respective emerging technology families. Emerging digital technologies are obtained by clustering the embeddings using the k–means algorithm, where the embeddings are derived with the sentence transformer all-mpnet-base-v2. For a short description of these technologies, refer to Tables A.1 to A.3 in Appendix A.1. Technologies are grouped by families, where a family comprises technologies whose occupation structure of semantic links is highly correlated.

# 3 Semantic–based Exposure

In this section, we present the methodology for computing the exposure scores of industries and occupations to emerging digital technologies. First, we describe how we compute the cosine similarity scores of industries and occupations with patents using textual data and filtering for relevant pairs. We describe each step in detail, for both industries and occupations. Then, we describe their aggregation from the patent to the technology level to obtain the semantic-based exposure scores.

These scores denote the *relevance* of each technology to a given industry or occupation rather than their actual adoption. For industries, the relevance is determined by whether a technology is integrated into the production process or if the technology itself constitutes an enhanced output of an industry. Regarding occupations, the relevance pertains to the significance of technology in the execution of tasks and functions inherent to an occupation.

## 3.1 Industry Cosine Similarity Scores

**Industry Descriptions.** We select the 3-digit NACE Rev.2 classification as the most detailed level at which to consider industries' descriptions. This selection is based on two primary considerations. First, this allows us to incorporate titles and descriptions from the 4-digit into the 3-digit industry descriptions—providing a more extensive text corpus for matching. Second, industry subsets under the same 3-digit category do not exhibit substantial differences in their connections to patents, allowing for a merger without significant loss of information.

For each industry $i \in \mathcal{I}$, we break the industrial descriptions (both the 3-digit and their nested 4-digit children) into individual sentences and concatenate each sentence with its corresponding title. We represent these composite sentences as $s \in S_i \subset \mathcal{S}_\mathcal{I}$, where $S_i$ denotes the set of composite sentences (i.e. title combined with one description sentence) corresponding to industry $i$. This results in 271 industries at the 3-digit level, each represented by 11 composite sentences on average.

**Embeddings.** We produce the embeddings of these composite sentences using the same pre-trained sentence transformer as in Section 2, namely all-mpnet-base-v2. The embedding of a composite sentence $s$ for an industry $i$ is denoted as $Emb_{s,i}$.

**Cosine Similarity.** For each patent $p \in \mathcal{P}$, we compute the cosine similarity of all composite sentences $s \in \mathcal{S}_\mathcal{I}$ with both parts of the patent titles, namely $p_1$ (representing the invention's description) and $p_2$ (representing its function). Specifically, the cosine similarities are computed as:

$$C_{s,i}^{p_1} = \frac{Emb_{s,i} \cdot Emb_{p_1}}{||Emb_{s,i}|| \, ||Emb_{p_1}||}, \tag{1}$$

$$C_{s,i}^{p_2} = \frac{Emb_{s,i} \cdot Emb_{p_2}}{||Emb_{s,i}|| \, ||Emb_{p_2}||}, \tag{2}$$

which quantify the semantic relationship between $p_1$, respectively $p_2$, and $s$. Nevertheless, similarity can be discerned through different nuances of meaning. In our context, this could pertain to aspects such as an application, a technical domain, or specified functions, whether central or ancillary. This data is encapsulated into a scalar, whose magnitude *approximates* the degree of similarity between an aspect of the industry (as described in its NACE 4-digit nomenclature) and an aspect of the invention (as described in the patent).

To reduce the noise and capture the most relevant meaning of the similarity between an invention and an industry, for each $(i, p_1)$ and $(i, p_2)$ combinations, we retain the composite sentence $s$ that exhibits the highest cosine similarity score. Formally,

$$C_i^{p_1} := \arg\max_{s \in S_i} C_{s,i}^{p_1}, \tag{3}$$

$$C_i^{p_2} := \arg\max_{s \in S_i} C_{s,i}^{p_2}, \tag{4}$$

where $C_{s,i}^{p_1}$ and $C_{s,i}^{p_2}$ are, respectively, given by Equations (1) and (2). These scalars summarize the quality of the semantic match between an industry $i$ and the description ($p_1$) or the function ($p_2$) of the patent.

**Redundancy.** To enhance the quality of the matching and filter out irrelevant matches, we incorporate *redundancy* in the calculation of cosine similarity of industry–patent pairs $(i, p)$. For industry–patent combinations $(i, p)$, we separately rank the sub-pairs $(i, p_1)$ and $(i, p_2)$ based on their respective cosine similarity scores $C_i^{p_1}$ and $C_i^{p_2}$. We then identify the industry-patent combinations $(i, p)$ as relevant (denoted as $(i, p)^\star$) if *both* sub-pairs $(i, p_1)$ and $(i, p_2)$ are within the top 10 of their respective rankings. This methodology results in the exclusion of certain pairs that do not rank simultaneously in the top 10 for both components.[11] Thus, we retain inventions for which both the description of the invention and its function are relevant to the industry.

For the identified relevant pairs, we calculate the harmonic mean with both cosine similarity scores, for the description of the invention and its function. This yields the composite

[11]In addition, we manually exclude three very specific connections to improve our exposure scores; see Appendix A.2 for more details.

Table 2: Example of Redundancy Filtering of Industries for Targeted TV Advertising

| | | Cosine Similarity | | |
|---|---|---|---|---|
| Code | NACE Industry | $C_i^{p_1}$ | $C_i^{p_2}$ | $C_i^{p}$ |
| 60.2 | Television programming and broadcasting activities | 0.391 | 0.445 | 0.416 |
| 73.1 | Advertising | 0.458 | 0.373 | 0.411 |
| 73.2 | Market research and public opinion polling | 0.295 | 0.272 | 0.283 |
| 59.1 | Motion picture, video and television programme activities | 0.271 | 0.263 | 0.267 |
| 61.2 | Wireless telecommunications activities | 0.290 | 0.229 | 0.256 |
| 26.3 | Manufacture of communication equipment | 0.257 | 0.240 | 0.249 |
| 78.1 | Activities of employment placement agencies | 0.265 | | |
| 47.9 | Retail trade not in stores, stalls or markets | 0.263 | | |
| 56.3 | Beverage serving activities | 0.261 | | |
| 80.1 | Private security activities | 0.253 | | |
| 61.3 | Satellite telecommunications activities | | 0.294 | |
| 61.1 | Wired telecommunications activities | | 0.237 | |
| 97.0 | Activities of households as employers of domestic personnel | | 0.231 | |
| 58.1 | Publishing of books, periodicals and other publishing activities | | 0.223 | |

*Notes*: This table presents the redundancy filtering of industries for the Patent ID 2013B87254. It displays the cosine similarity of distinct 3-digit NACE Rev.2 industry descriptions with the patent description "Method for targeting television advertisement based on profile linked to online device" (Column 3) and the function principle "selecting television advertisement to be directed to set-top box based on profile information pertaining to the user or online activity" (Column 4). Industries are ranked according to Column 3 in decreasing order. Cosine similarity scores in Columns 3 and 4 are displayed only for sub-pairs belonging to their respective top 10. Column 5 shows the composite patent-industry cosine similarity score, which corresponds to the harmonic mean of Columns 3 and 4. Cosine similarity scores in Column 5 are displayed only for pairs that rank simultaneously in both top 10.

cosine similarity score for industry–patent pairs $(i, p)^\star$ as follows:

$$C_i^p = 2\left(\frac{1}{C_i^{p_1}} + \frac{1}{C_i^{p_2}}\right)^{-1}, \tag{5}$$

where $C_i^{p_1}$ and $C_i^{p_2}$ are, respectively, given by Equations (3) and (4). As a result of the calculation presented in Equation (5), we establish a connection between an invention identified in a single patent $p \in \mathcal{P}$ and a set of relevant industries in which that patent can be used to improve the process product or organization.

Table 2 illustrates the redundancy principle at work, considering the first patent example provided in Section 2. This example describes a targeted TV advertising method based on user profile information. For this patent, redundancy helps filter out industries irrelevant to the technology. The redundancy filtering for the other two patent examples mentioned in Section 2 is presented in Tables A.4 and A.5 in the appendix.

### 3.2 Occupation Cosine Similarity Scores

**Occupation Descriptions.** We choose the 4-digit ISCO-08 as the most detailed level at which to consider the textual description of occupations. Unlike industries, the 4-digit level of ISCO-08 comprises a set of distinct occupations that are informative for our analysis. Each ISCO-08 occupation is associated with a specific set of tasks, although some tasks may overlap across different occupations.

For each occupation $o \in \mathcal{O}$, we consider two components of the occupation description: the occupation title $o_1$ and the task description $o_2$. We divide the task description into individual tasks $s \in S_o \subset \mathcal{S}_{\mathcal{O}}$, where $S_o$ is the set of tasks for occupation $o$. This results in 433 occupations at the 4-digit level, each represented by one title and, on average, 7.5 tasks.

**Embeddings.** Similar to industries, we produce the embeddings using the same sentence transformer model. We represent the embeddings of the occupation title as $Emb_{o_1}$ and the embeddings of a task $s$ as $Emb_{s,o_2}$.

**Cosine Similarity.** For each patent $p \in \mathcal{P}$, we compute the cosine similarity of the patent title (in its entirety) with both parts that describe the occupations, namely, the title $o_1$ and all the tasks separately $o_{s,2}$. More specifically, we compute the cosine similarities as:

$$C^p_{o_1} = \frac{Emb_{o_1} \cdot Emb_p}{||Emb_{o_1}||\,||Emb_p||}, \tag{6}$$

$$C^p_{s,o_2} = \frac{Emb_{s,o_2} \cdot Emb_p}{||Emb_{s,o_2}||\,||Emb_p||}, \tag{7}$$

which express the idea of the semantic connection between $o_1$, respectively $s$, and $p$.

For each $(o_2, p)$ combination, as above for industries, we retain the composite sentence with the highest cosine similarity score. More formally,

$$C^p_{o_2} := \arg\max_{s \in S_o} C^p_{s,o_2}, \tag{8}$$

where $C^p_{s,o_2}$ is the cosine similarity between patent $p$ and task $s$ given by Equation (7). There is no need to aggregate in the case of the title part $o_1$ as each occupation has only one title. The quality of the semantic match between an occupation and a patent is summarised in both of these scalars, either through the title of the occupation or the tasks performed in that latter.

**Redundancy.** We employ the same methodology as with industries, designating the occupation–patent combinations $(o, p)$ as relevant (denoted as $(o, p)^\star$) if *both* sub-combinations $(o, p)_1$

and $(o,p)_2$ rank within the top 10 of their respective rankings. Thus, we retain inventions for which the invention is relevant to the occupation.[12]

For the identified relevant pairs, we calculate the harmonic mean with both cosine similarity scores. This yields the composite cosine similarity score for industry–patent pairs $(o,p)^{\star}$ as follows:

$$C_o^p = 2\left(\frac{1}{C_{o_1}^p} + \frac{1}{C_{o_2}^p}\right)^{-1}, \tag{9}$$

where $C_{o_1}^p$ and $C_{o_2}^p$ are, respectively, given by Equations (6) and (8). As a result of the calculation presented in Equation (9), we establish a connection between an invention identified in a single patent $p \in \mathcal{P}$ and a set of occupations to which that patent is relevant. Tables A.6 to A.8, in the appendix, illustrate redundancy filtering of occupations for our patent examples from Section 2.

### 3.3 Aggregation by Technology

We aggregate cosine similarity scores $C_p^i$ and $C_p^o$ obtained at the patent level in Equations (5) and (9), to the technology level. To this end, we implement a weighting scheme based on the number of citations that a patent receives from other patents to proxy for the relevance of each patent and the likelihood that it is used in industries and occupations. Given the heterogeneity in patent impact, it is pertinent that their weighting reflects this (Hall et al. 2005, OECD 2009).

We assign a weight to the cosine similarity score of each patent proportional to the number of citations it has received relative to the total number of citations accrued by all patents associated with the same occupation/industry–technology pair within the same year.[13] The specific weight assigned to a patent is computed as:

$$\omega_d^p = \frac{m_p}{\sum_{p \in \mathcal{P}_{dt}^k} m_p}, \tag{10}$$

where $m_p$ is the number of citations received by patent $p$, $\mathcal{P}_{dt}^k$ represents the set of patents associated with emerging digital technology $k$, filed in year $t$, and relevant to industry/occupation $d = \{i,o\}$.

We implement this weighting scheme to aggregate the cosine similarity scores at the patent

[12] In addition, we manually exclude three very specific connections to improve our exposure scores; see Appendix A.2 for more details.

[13] Approximately 41% of patents in our sample have not received any citations. This includes 1,733 patents, or 0.91%, which had an indeterminable citation count and are treated as having zero citations. Similarly, there are 77,307 patents, or 40.54%, patents with no citations. Figure A.3 in the appendix shows the distribution of patents of undetermined-count and non-cited patents across technologies. Figure A.2 in the appendix shows the distribution of patent citations across technologies.

level to the technology level. The cosine similarity of a technology $k$ to an industry/occupation is then computed as:

$$C_{dt}^{k} = |\mathcal{P}_{dt}^{k}| \times \sum_{p \in \mathcal{P}_{dt}^{k}} \omega_{d}^{p} C_{d}^{p}, \tag{11}$$

where $C_d^p$ denotes the cosine similarity score of the pair $(d,p)$ as derived from Equations (5) and (9), $\omega_d^p$ represents the weight from Equation (10), and $|\mathcal{P}_{dt}^k|$ is the total number of patents assigned to industry/occupation–technology pair $(d,k)$ for $d = \{i,o\}$ in year $t$. This results in the cosine similarity score of industry/occupation $i/o$ with technology $k$ for the year $t$.[14] Lastly, we aggregate cosine similarity scores across all years to obtain a cumulative measure for the period 2012–2021. The equation for this aggregation is as follows:

$$C_{d}^{k} = \sum_{t} C_{dt}^{k}, \text{ with } d = \{i,o\}. \tag{12}$$

### 3.4 Exposure Scores

To obtain our final measure of the exposure of 3-digit NACE Rev.2 industries and 4-digit ISCO-08 occupations to emerging digital technologies $X_d^k$, we apply inverse hyperbolic sine transformation, which helps address the right skewness in the distribution of cosine similarity scores. Formally,

$$X_{d}^{k} = \sinh^{-1}\left(C_{d}^{k}\right), \tag{13}$$

where $C_d^k$ is the cosine similarity score for industry/occupation–technology pair $(d,k)$ between 2012 and 2021 as described in Equation (12).

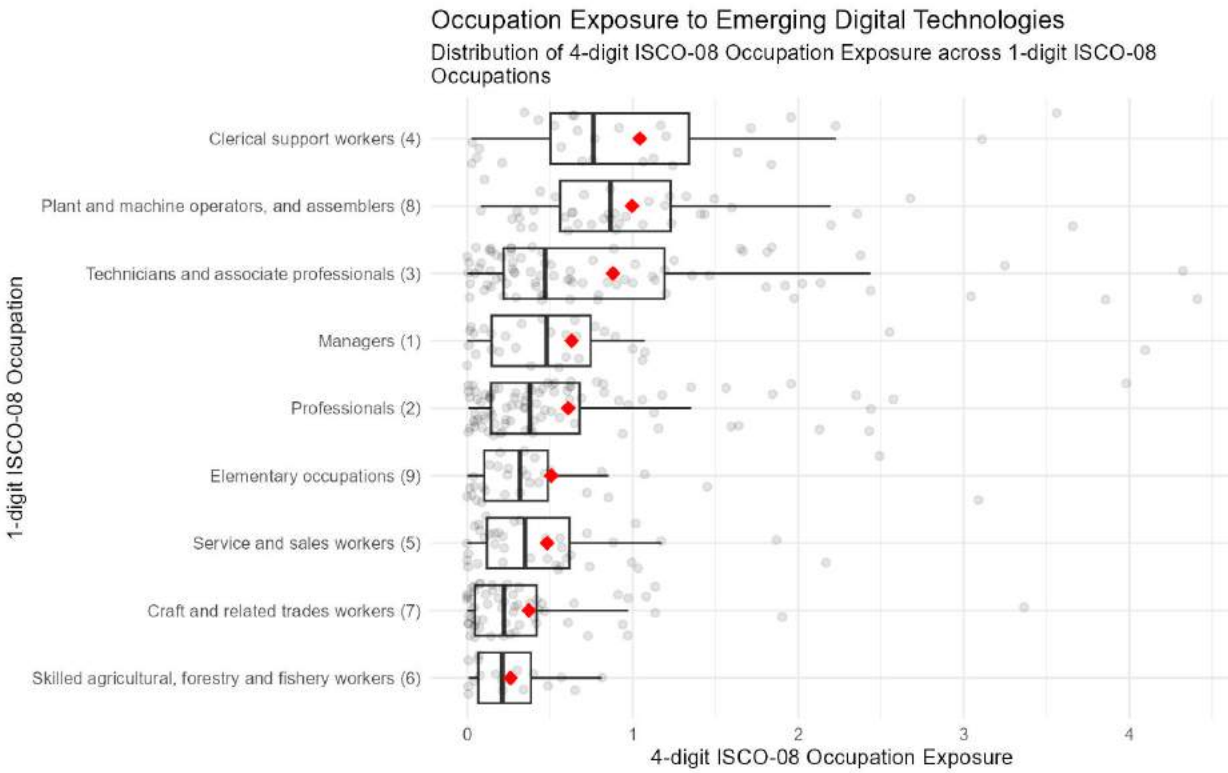
We deliver these data as an open–access database, the 'TechXposure' database. In this database, we also provide measures of exposure at higher levels of aggregation, such as the 1-digit and 2-digit levels for industries, and from the 1-digit to the 3-digit levels for occupations. For details on the derivation of these measures, see Appendix A.6.

## 4 Descriptive Analysis

In this section, we describe the exposure of both occupations and industries to emerging digital technologies. We start with occupations and then look at industries.

[14]Note that aggregating without weighting by citations results in yearly cosine similarity scores very similar to those obtained with the weighting scheme. Figure A.4, in the appendix, displays the correlation between the weighted and unweighted yearly cosine similarity scores. The Pearson correlations between scores derived from both methods are approximately 0.99 for both industries and occupations. The Spearman rank correlation yields a value of about 0.89.

Figure 1: Overall Occupation Exposure by 1-digit ISCO-08 Occupation



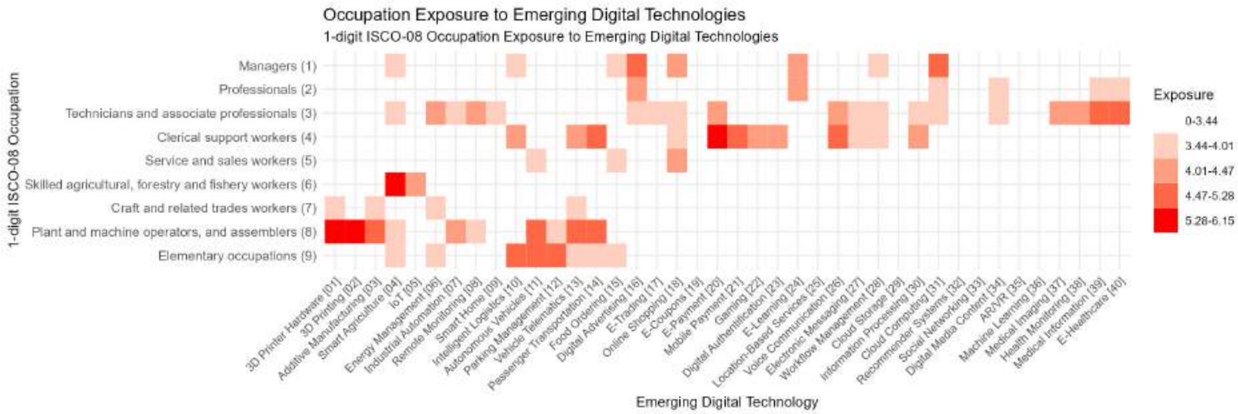

*Notes*: This figure presents the distribution of exposure to emerging digital technologies across 4-digit ISCO-08 occupations, with each 1-digit occupation displayed separately in boxplots. Vertical bars indicate the median exposure for all 4-digit occupations within the same 1-digit occupation, and diamond points represent the average exposure for these 4-digit occupations.

## 4.1 Occupation Exposure to Emerging Digital Technologies

We start by examining the overall exposure of occupations, which we define as the average exposure across all technologies. This corresponds to $X_o = \frac{1}{40}\sum_k X_o^k$, where $X_o^k$ is defined by Equation (13). Figure 1 presents the distribution of overall exposure to emerging digital technologies across ISCO-08 occupations. In this figure, 4-digit occupations are grouped into their respective 1-digit categories, and their distribution is presented as a boxplot. Occupation groups are ranked by their average overall exposure to emerging digital technologies, indicated by the diamond point.

We observe that Clerical Support Workers (ISCO-08 Group 4) and Plant and Machine Operators, and Assemblers (8) are the most exposed to emerging digital technologies. The occupations in these ISCO groups typically involve a higher proportion of routine tasks associated with information handling and production equipment supervision, respectively. Despite having already experienced a significant impact from earlier waves of ICT development (Goos and Manning 2007, Goos et al. 2009, Goos et al. 2014), these middle-paying jobs continue to be strongly related to newer ICT vintages, especially emerging digital technologies that enable

Figure 2: Occupation Exposure by Emerging Digital Technologies (1-digit ISCO-08)




*Notes*: Each cell shows the exposure of a 1-digit ISCO-08 occupation (row) to a given emerging digital technology (column). Exposure scores below the 80th percentile (0-3.44) are transparent, whereas the four other groups represent respectively the 80th (3.44-4.01), 90th (4.01-4.47), 95th (4.47-5.28), and 99th (5.28-6.15) percentile of the distribution. Figure B.1, in the appendix, presents the same figure at the 2-digit level.

handling of information and production equipment in semi- or unsupervised manner.

High-paying occupations, including Managers (1), Professionals (2), and Technicians and Associate Professionals (3), are the next most exposed to the emerging digital technologies. The tasks performed in these occupations are predominantly non-routine and cognitive, often involving the use of a variety of digital technologies. As technologies advance and new vintages appear, new tasks may also emerge, leading to changes in the task structure of these occupations.

Conversely, we observe that low-paying occupations, such as Service and Sales Workers (5), Skilled Agricultural, Forestry and Fishery Workers (6), Craft and Related Trades Workers (7), and Elementary Occupations (9), are less exposed to emerging digital technologies. These occupations involve more interactive and non-routine tasks, which are less reliant on these technologies.

Lastly, we observe greater heterogeneity in exposure to emerging digital technologies within high-paying occupations (1, 2, and 3) compared to middling occupations (4 and 8). This suggests that only a subset of the former group is related to emerging technologies, while the latter group exhibits more generalized exposure.

We break down the overall exposure of 1-digit ISCO Groups by examining at their exposure to each of the 40 emerging digital technologies. Figure 2 presents 1-digit occupation exposure as a heatmap, where the exposure levels are indicated at the intersections of 1-digit occupations (rows) and emerging digital technologies (columns). This visualization reveals two distinct patterns.

First, we observe a distinct divide between *tangible* and *intangible* technologies in terms of their relevance to different occupations. On the one hand, *tangible* technology families, such as 3D Printing, Embedded Systems, and Smart Mobility, are more relevant to manual occupations within ISCO Groups 6 to 9. On the other hand, *intangible* technology families, such as Computer Vision, E-Commerce, Payment Systems, HealthTech, and Digital Services, are more relevant to cognitive occupations, specifically within ISCO Groups 1 to 4.

Second, we note that both Technicians and Associate Professionals (3) and Clerical Support Workers (4) exhibit exposure to a wide range of emerging digital technologies. In contrast, Managers (1) and Professionals (2) appear to have a more limited scope of relevant technologies, primarily concentrated in the realm of intangible technologies. Similarly, exposure of ISCO Groups 6 to 9 is exclusively focused on tangible technologies. It is important to note that this aggregated mapping conceals some heterogeneity in exposure within 1-digit ISCO-08 occupations due to aggregation; see Figure B.1 in the appendix for a more detailed mapping at the 2-digit level.

## 4.2 Industry Exposure to Emerging Digital Technologies

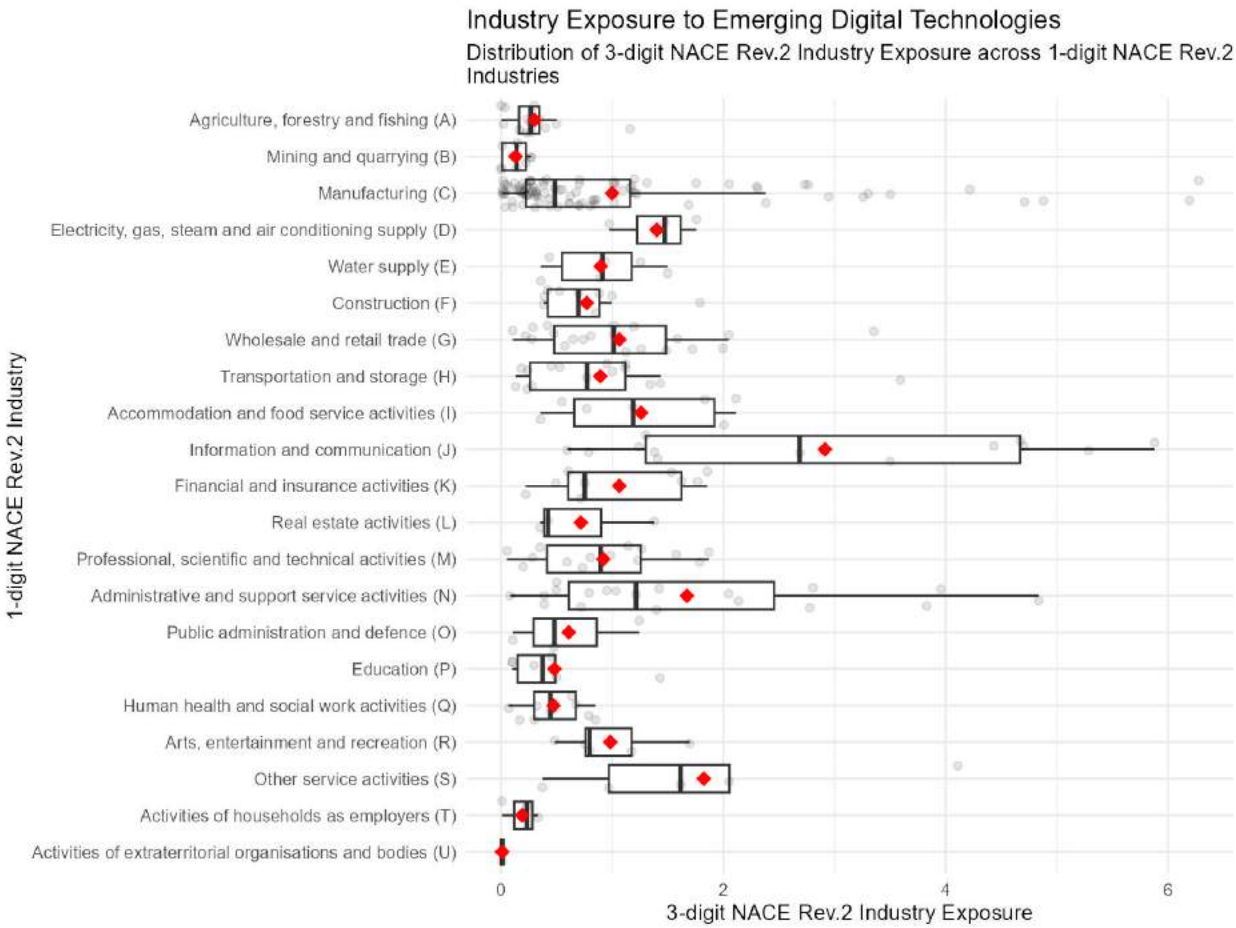
For industries, we examine their overall exposure, which we define as the average exposure across all technologies. This corresponds to $X_i = \frac{1}{40} \sum_k X_i^k$, where $X_i^k$ is defined by Equation (13). Figure 3 presents the distribution of overall exposure to emerging digital technologies across NACE Rev.2 industries. In this figure, 3-digit industries are grouped into their respective 1-digit sectors, and their distribution is presented as a boxplot.

We observe that the Information and Communication (J) and Manufacturing (C) sectors host the most exposed 3-digit industries. This finding is notable due to the significant heterogeneity of industry exposure within these 1-digit sectors. Such differences in exposure may indicate the industries' roles as either producers or intensive users, as opposed to light users, of emerging digital technologies. More specifically, industries within the Information and Communication (J) sector are likely to produce intangible technologies, while a specific subset of the Manufacturing (C) sector is likely to produce tangible technologies.

Interestingly, the Administrative and Support Service Activities (N) sector exhibits a high average level of exposure to emerging digital technologies. Several 3-digit industries within this sector achieve overall exposure levels comparable to those in Sectors C and J. This observation is consistent with the findings presented in Section 4.1, as Sector N is a significant employer of Clerical Support Workers (ISCO Group 4), identified as the most exposed 1-digit ISCO Group (see Fig. 1).

We disaggregate the overall exposure of 1-digit NACE sectors into their exposure to each of the 40 emerging digital technologies. Figure 4 replicates the exposure heatmap for 1-digit sectors; see Figure B.2 in the appendix for a more detailed mapping at the 2-digit level.

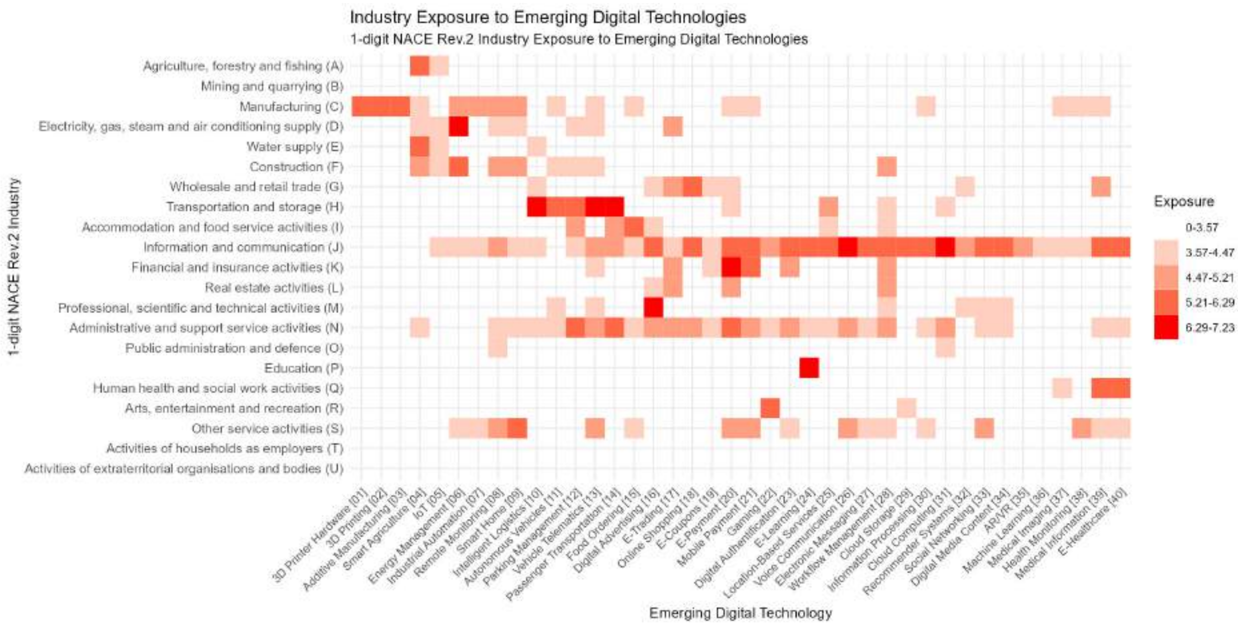
Figure 3: Overall Industry Exposure by 1-digit NACE Rev.2 Industry




*Notes*: This figure presents the distribution of exposure to emerging digital technologies across 3-digit NACE Rev.2 industries, with each 1-digit industry displayed separately in boxplots. Vertical bars indicate the median exposure for all 3-digit industries within the same 1-digit industry, and diamond points represent the average exposure for these 4-digit industries.

Similar to occupations, we observe a divide between tangible and intangible emerging digital technologies. In the figure, exposure cells follow a top-left to bottom-right diagonal pattern, thereby associating tangible technologies with sectors like Agriculture (A), Mining and Quarrying (B), and Manufacturing (C), and aligning intangible technologies with service sectors from Financial and Insurance Activities (K) to Other Service Activities (S). In between these extremes, we find sectors ranging from Electricity, Gas and Air Conditioning Supply (D) to Information and Communication (J) operate physical infrastructures and are thus exposed to more tangible but distributed technology families, such as Embedded Systems and Smart Mobility.

Figure 4: Industry Exposure by Emerging Digital Technologies (1-digit NACE Rev.2)




*Notes*: Each cell shows the exposure of a 1-digit NACE Rev.2 industry (row) to a given emerging digital technology (column). Exposure scores below the 80th percentile (0-3.57) are transparent, whereas the four other groups represent respectively the 80th (3.57-4.47), 90th (4.47-5.21), 95th (5.21-6.29), and 99th (6.29-7.23) percentile of the distribution. Figure B.2, in the appendix, presents the same figure at the 2-digit level.

# 5 Impact on Employment

In this section, we estimate the causal effect of emerging digital technologies on regional employment using an instrumental variable (IV) shift-share approach based on industry exposure scores and the baseline employment shares of these industries in the region. Our analysis is conducted using a setup similar to that of Acemoglu and Restrepo (2020), who estimate the impact of robots on US regional employment.

Throughout this section, we interpret our estimates using the canonical task-based framework of Acemoglu and Restrepo (2018, 2019). This framework identifies three mechanisms by which (emerging digital) technologies can affect labor demand and, consequently, employment. The initial premise is that new technologies enable capital to substitute for labor in a range of tasks. This substitution has three impacts on labor. First, new technologies might change the task content of production, reducing the role of labor and hence labor demand, leading to lower employment; this is known as the *displacement effect*. Second, new technologies may enhance worker productivity by enabling a more flexible allocation of tasks, thereby increasing labor demand and employment; this is termed the *productivity effect*. Third, new technologies may create new tasks, consequently increasing labor demand and employment; this is identified as the *reinstatement effect*.

We begin by describing our IV shift-share approach to estimate the overall impact of emerging digital technologies on regional employment. Then, we assess the employment effects of technology families. Lastly, we proceed to a more granular level, estimating the effects of each technology on regional employment.

## 5.1 Overall Impact of Emerging Digital Technologies

We use employment data from the Regional European Labour Force Survey (EU-LFS), which provides information on the number of employees and population at the NUTS-2 level from 322 regions in 32 European countries.[15] Additionally, the EU-LFS provides information on the number of employees in 1-digit NACE industries that are grouped into 10 distinct sectors.[16] Our period of analysis begins in 2012, which marks the starting year for our patent sample, hence, our measure of exposure to emerging digital technologies. We conclude our analysis in 2019 to avoid confounding factors related to employment and population fluctuations caused by the COVID-19 pandemic.[17]

We estimate the impact of emerging digital technologies on regional employment using a long–difference approach. We focus on the regional employment-to-population ratio as our outcome variable, specifically examining the change in this ratio between 2012 and 2019.

Estimating the causal impact of technology on employment presents two main challenges: reverse causality and omitted variable bias. Reverse causality suggests that technological advancements may also result from labor shortages or rising labor costs. Additionally, unobserved factors, such as changes in the organisation of industries or investments in infrastructures, could simultaneously affect both technological change and employment levels.

To address these concerns, we adopt a shift-share strategy, leveraging recent advancements in this literature (Adão et al. 2019; Goldsmith-Pinkham et al. 2020; Borusyak et al. 2021). Specifically, we use the Bartik instrument by interacting regional industry shares from the baseline year of 2010 with their corresponding exposure to emerging digital technologies during the period 2012–2019. In this context, industry shares serve as a measure of differential exposure to these technologies at the regional level.

The identification relies on two key assumptions. First, we assume that regional industry

[15]The list of countries includes (in alphabetical order): Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, and the United Kingdom.

[16]These sectors are Agriculture (A); Industry (B-E); Construction (F); Market Services (G-I); Information and Communication (J); Financial and Insurance Activities (K); Real Estate Activities (L); Professional, Scientific, Technical, Administration and Support Service Activities (M-N); Public Administration, Defence, Education, Human Health and Social Work Activities (O-Q); and Other Services (R-U).

[17]Although our exposure metrics from Section 3 cover the period 2012–2021, we recompute them for the subperiod 2012–2019 to maintain consistency with our period of analysis in this section.

shares are exogenous when conditioned on observable factors, such as country fixed effects and regional characteristics, in line with the argument presented by Goldsmith-Pinkham et al. (2020). Although direct testing of this assumption is not feasible, we show in Figure C.1 in the appendix that there is no correlation between regional employment shares in 2010 and the subsequent change in the regional employment-to-population ratio from 2012 to 2019, conditional on country fixed effects and regional demographics. The second identifying assumption, inspired by the methodology in Acemoglu and Restrepo (2020), posits that regions more exposed to emerging digital technologies are not disproportionately affected by other labor market shocks or trends.

To reinforce the validity of our shift-share strategy, we argue that the industrial exposure to emerging digital technologies, which presents the shock in our shift-share design, is quasi-exogenous to changes in regional employment within Europe. Our metrics for industrial exposure, as derived in Section 3, are based on the semantic similarity between novel patents and industry descriptions. Yet, only 9.6% of the patents in our sample originate from Europe, suggesting that the advancement of these technologies is predominantly a global phenomenon.[18] Global technological trends are unlikely to be only influenced by regional labor markets in Europe.

The regional exposure to all emerging digital technologies is determined by the interaction between sectoral employment shares in the baseline year and sectoral exposure to all these technologies. More precisely, the exposure of a region $X_r$ is calculated as follows:

$$X_r = \sum_j l_{rj} X_j, \tag{14}$$

where $l_{rj}$ is the employment share of sector $j$ in region $r$ in the baseline year 2010, and $X_j$ is defined as the average exposure of sector $j$ to emerging technologies from 2012 to 2019, calculated as

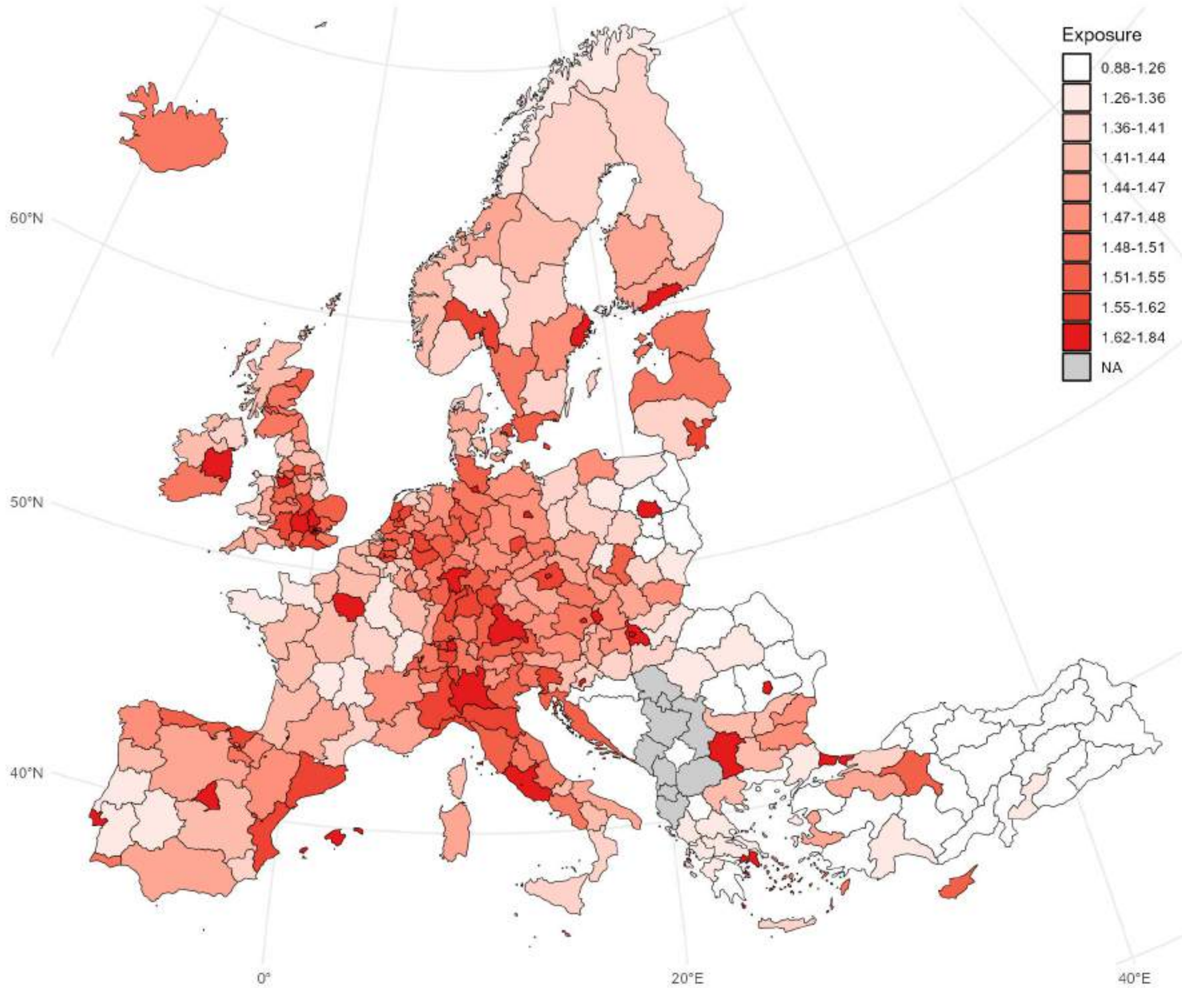
$$X_j \equiv \frac{1}{40} \times \sum_{k \in \mathcal{K}} X_j^k,$$

where $X_j^k$ represents the average exposure of sector $j$ to each technology $k$ across all 1-digit NACE industries $i \in j$ during this period.

Table C.1 in the Appendix details the average employment share by economic sector across European regions in 2010. The three largest sectors are the Public Sector (with an average employment share of 25.7%), Market Services (24%), and Industry (17.1%). The Information and Communication sector, which is the most exposed to emerging digital technologies, accounts for only 2.3% of employment on average.

[18]Among patents originating from Europe, roughly 28% did not receive any citations.

Figure 5: Geographic Distribution of Regional Exposure to Emerging Digital Technologies across Europe from 2012 to 2019

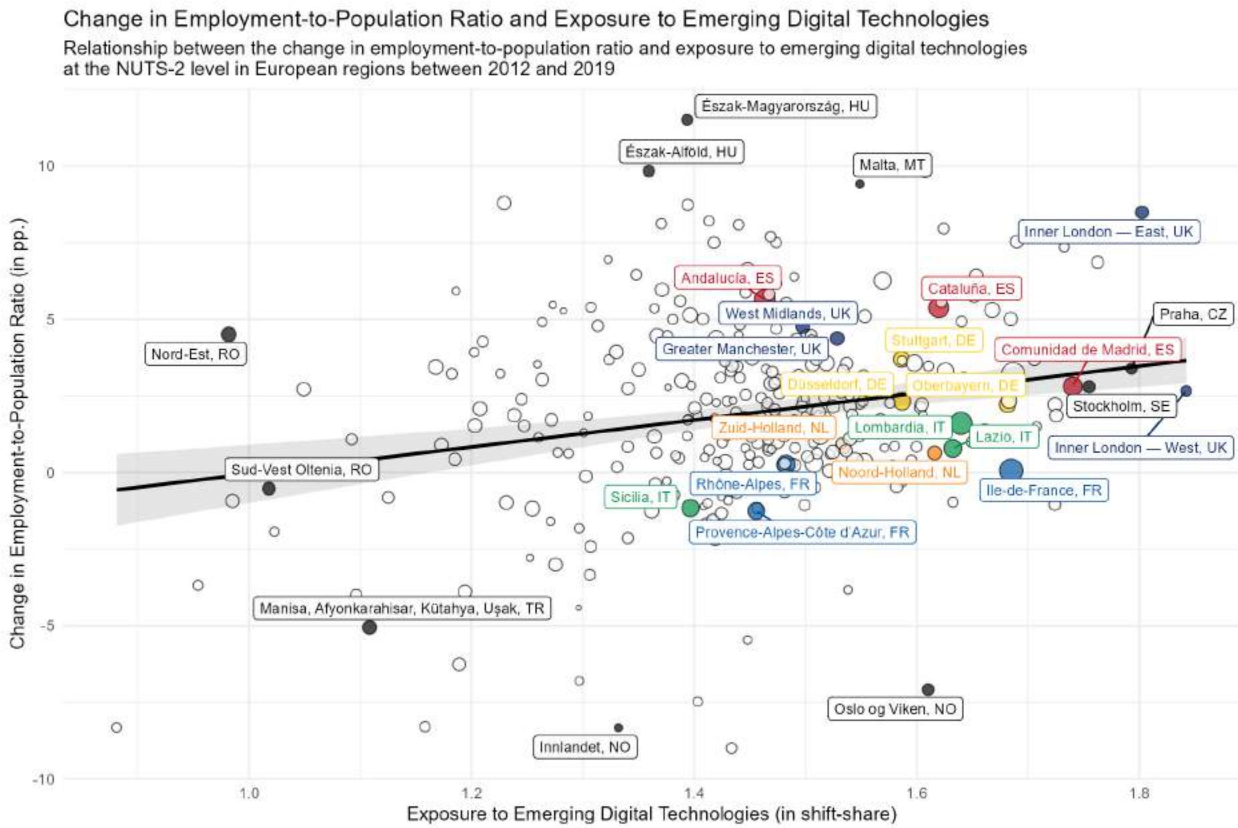



*Notes*: This figure illustrates the geographic distribution of exposure to emerging digital technologies for NUTS-2 regions. Regional exposure is constructed as a shift-share variable by interacting the sectoral employment shares in the baseline year 2010 and sectoral exposure to these technologies from 2012 to 2019. Regions are categorized into deciles. Regions are shaded according to their exposure level, with the legend indicating the range of exposure. Areas not applicable (NA) are marked in grey.

Figure 5 shows the geographic distribution of exposure across European regions. Emerging digital technologies are more prevalent in industries concentrated in European capital cities, which typically have larger service sectors compared to more peripheral regions. Beyond capital cities, regions with the highest levels of exposure levels are predominantly found in Western Europe, specifically in countries such as Germany, Italy, Spain, Switzerland, and the UK.

Figure 6 depicts a positive relationship between the change in the employment-to-population

Figure 6: Change in Employment-to-Population Ratio and Exposure to Emerging Digital Technologies




*Notes*: This figure shows the relationship between the change in the employment-to-population ratio and the exposure to emerging digital technologies in European NUTS-2 regions between 2012 and 2019. Each point represents a region. The size of the point is proportional to the population in 2010. The horizontal axis measures the exposure to emerging technologies calculated by the shift-share method, while the vertical axis represents the change in the employment-to-population ratio in percentage points (pp.). The solid line indicates a positive correlation between regional exposure to emerging technologies and employment growth. The grey shaded area indicates the 95% confidence interval.

ratio from 2012 to 2019 and the regional exposure to emerging digital technologies.[19] Although the observed correlation is statistically significant, it is not adjusted for country fixed effects and regional demographic characteristics, which are crucial for our identifying assumption regarding the exogeneity of the shares conditional on observables.

We estimate the impact of regional exposure to emerging digital technologies on the regional employment-to-population ratio change using the following empirical specification:

$$\Delta Y_r = \alpha + \beta X_r + Z\delta + \phi_{c(r)} + u_r, \tag{15}$$

where $\Delta Y_r$ represents the change in the employment-to-population ratio (in pp.) for region

[19] In the Online Appendix, we show that this positive association persists even after excluding regions with exceptionally low exposure levels — specifically, those with exposure below -2 standard deviations (i.e. below 1.149), which typically includes rural areas in Romania, Turkey, and overseas French territories.

Table 3: Effect of Emerging Digital Technologies on Regional Employment

| | $\Delta$ Emp-to-pop. ratio (2012-2019) $\times$ 100 | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Exposure to Emerging Technologies | $0.634^{**}$ | $0.957^{***}$ | $1.069^{***}$ |
| | (0.212) | (0.132) | (0.116) |
| Country FE | ✓ | ✓ | ✓ |
| Demographics | | ✓ | ✓ |
| Industry share | | | ✓ |
| $R^2$ | 0.666 | 0.695 | 0.697 |
| Adj. $R^2$ | 0.629 | 0.655 | 0.656 |
| Num. obs. | 322 | 322 | 322 |

*Notes*: This table presents the estimates of exposure to emerging digital technologies on regional employment. It presents the coefficients measuring the effect of regional exposure to emerging technologies, constructed as shift-shares and standardized, on changes in the employment-to-population ratio between 2012 and 2019 in European regions, expressed in percentage points. Regressions are weighted by population in 2010. Column (1) includes country fixed effects; Column (3) adds demographics controls in 2010, including the logarithm of population, the proportion of females, the proportion of the population aged over 65, and the proportions of the population with secondary and tertiary education levels; Column (3) adds the share of employment in the industry sector. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors between parentheses are derived following the AKM0 inference procedure from Adão et al. (2019).

$r$ between 2012 and 2019, $X_r$ denotes the regional exposure to emerging digital technologies as defined in Equation (14) and standardized, $Z$ is a set of covariates which capture regional characteristics,[20] $\phi_{c(r)}$ are country fixed effects, and $u_r$ is the error term.

Table 3 presents the estimates of the effect of regional exposure to emerging digital technologies on the change in the employment-to-population ratio (2012–2019). As the exposure is standardized across regions, the estimated coefficient of interest $\hat{\beta}$ can be interpreted as the effect of a one-standard-deviation increase in regional exposure on the employment-to-population ratio, expressed in percentage points. Following the recent literature on shift–share designs, we report the AKM0 shift–share standard errors which account for arbitrary cross-regional correlation in the regression residuals (Adão et al. 2019).

The positive relationship observed in Figure 6 remains robust upon including fixed effects, and various covariates, such as demographic characteristics of the region and the industry share. In the specification encompassing all covariates, in the last column, a one-standard-deviation increase in regional exposure implies a 1.069 pp. change, equivalent to 2.1%, in the employment-to-population ratio from 2012 to 2019.

The latter estimation reveals the overall impact of emerging digital technologies on em-

[20]Our set of control variables, fixed at their 2010 values to avoid endogeneity, is similar to Acemoglu and Restrepo (2020). This set includes the log of population (in thousands), the proportion of females, the proportion of the population aged over 65, the proportion of the population with secondary and tertiary education levels, and the proportion of employment in the industry sector.

ployment is positive at the regional level. However, it remains to be determined whether this positive relationship between emerging digital technologies and employment is uniform across all technologies, or if certain technologies might negatively affect employment — as documented in the literature (e.g. Acemoglu and Restrepo 2020).

### 5.2 Assessing the Impact of Emerging Digital Technology Families

We conduct the analysis at the level of technology families. We employ the same shift-share strategy to calculate the regional exposure to technology family $X_r^K$, defined as

$$X_r^K = \sum_j l_{rj} X_j^K,$$

where $l_{rj}$ is the employment share of sector $j$ in region $r$, and $X_j^K$ is the exposure of sector $j$ to technology family $K$, which is computed as the average sectoral exposure across technologies within the same family (i.e., $X_j^K = \frac{1}{|K|}\sum_{k \in K} X_j^k$).

Figure 7 illustrates the geographic distribution of exposure to the 9 families of emerging digital technology.[21] Regional exposure is standardized at the family level to facilitate comparisons and account for variations in exposure magnitudes across different technology families.

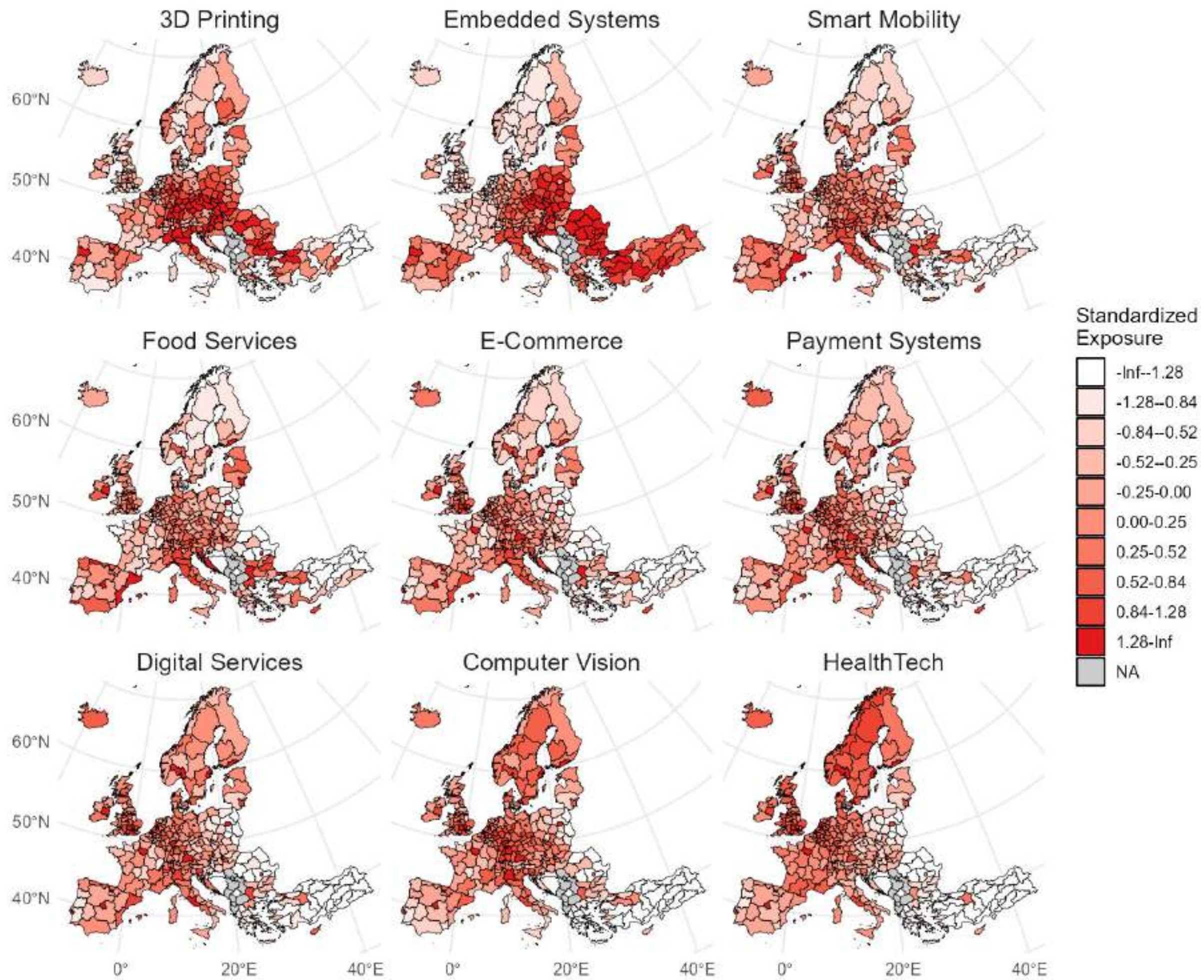
Exposure to emerging digital technologies exhibits significant variation across European regions and between technology families. For instance, regions with the highest exposure to tangible technologies, such as 3D Printing and Embedded Systems, are predominantly located in Central and Eastern European countries, as well as in certain areas of Southern Europe, including Northern Portugal and Turkey. These are the regions with the highest manufacturing shares. Conversely, Western and Northern European countries show greater exposure to Computer Vision and HealthTech, which correlates with their more service-oriented economies and digitized healthcare systems.

Furthermore, spatial differences in exposure are also evident within countries, characterized by disparities between rural and urban areas. Exposure to E-Commerce, Payment Systems, and Digital Services is predominantly concentrated in capital cities and financial hubs. In contrast, exposure to Smart Mobility and Food Services is relatively more pronounced in the rural regions of Western countries, such as France, Italy, Spain, and the United Kingdom.

We estimate the impact of the regional exposure to a specific emerging technology family on the employment-to-population ratio using an empirical specification analogous to that of Equation (15). However, instead of using the exposure to all technologies $X_r$, we focus on the

[21] In Appendix C.2, we provide the geographic distribution of exposure for all 40 individual emerging digital technologies in Figures C.2 through C.6.

Figure 7: Geographic Distribution of Regional Exposure to Families of Emerging Digital Technologies across Europe from 2012 to 2019




*Notes*: This figure illustrates the geographic distribution of exposure to families of emerging digital technologies for NUTS-2 regions. Regional exposure is constructed as a shift-share variable by interacting the sectoral employment shares in the baseline year 2010 and sectoral exposure to these technology families from 2012 to 2019. Regions are categorized into deciles. Regions are shaded according to their exposure level, with the legend indicating the range of exposure. Areas not applicable (NA) are marked in grey.

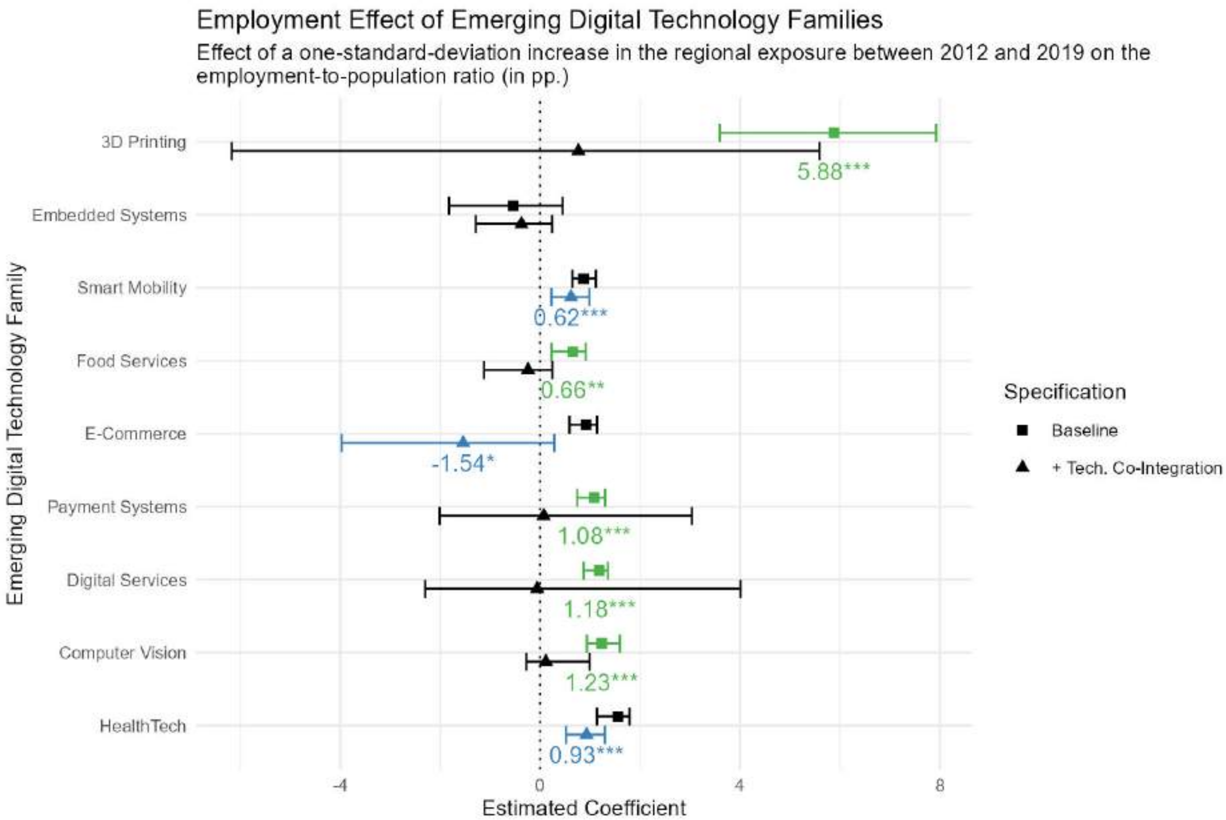
regional exposure to a particular family $X_r^K$. More specifically, the empirical specification is:

$$\Delta Y_r = \alpha + \beta_K X_r^K + \gamma_K X_r^{-K} + Z\delta + \phi_{c(r)} + u_r, \tag{16}$$

where $X_r^{-K}$ is regional exposure to all *other* emerging digital technologies. This latter variable is constructed as a shift–share variable, similar to that of Equation (14), but specifically excluding the exposure from the technology family of interest $K$. For interpretability, we standardize our variable of interest $X_r^K$.

Our estimated coefficient of interest, denoted as $\hat{\beta}_K$, represents the employment effect, measured in pp. change, of a one-standard-deviation increase in the regional exposure to a

Figure 8: Employment Effect of Emerging Digital Technology Families




*Notes*: This figure the coefficients measuring the effect of regional exposure to emerging digital technology families, constructed as shift-shares and standardized, on changes in the employment-to-population ratio, expressed in percentage points (pp.), between 2012 and 2019 in European regions. The confidence intervals are reported at the 5% significance level using the AKM0 inference procedure from Adão et al. (2019). Regressions are weighted by population in 2010. The analysis includes two empirical specifications: the first is the baseline, incorporating country fixed effects and regional controls in 2010 (including the logarithm of population, the proportion of females, the proportion of the population aged over 65, the proportions of the population with secondary and tertiary education levels, and the share of employment in the industry sector); the second specification accounts for technology co-integration by controlling for the regional exposure to all other emerging technologies, also constructed as a shift-share.

specific emerging technology family $K$, conditional on the regional exposure to all *other* families of emerging digital technologies. This empirical approach allows us to identify the causal effect of technology family $K$ on employment at the regional level by accounting for any potential co-integration of that family with other emerging digital technologies. Our approach is consistent with methodologies applied in the recent literature, which assess the impact of a particular technology — such as robots — on employment, while also accounting for exposure to complementary technologies — such as ICT (see, for example, Acemoglu and Restrepo 2020; Dauth et al. 2021).

Figure 8 displays the estimated coefficients along with their corresponding 95% AKM0 confidence intervals for the employment effects of emerging digital technology families. Before delving into the results, it is crucial to explain how to interpret the figure, especially since this layout will be employed again in the discussion of individual emerging technologies in the following subsection.

The figure is interpreted as follows. The vertical axis lists the technology families, while the horizontal axis depicts the estimated coefficients. Each technology family is associated with two coefficients, each representing a distinct empirical specification. The baseline specification includes country fixed effects $\phi_{c(r)}$ and regional controls $Z$. In the second specification, we also control for the co-integration with all other emerging digital technologies by adding $X_r^{-K}$ as a control variable. The figure reports the point estimates as well as their 95% confidence intervals derived from the AKM0 inference procedure which are, by construction, not symmetric around the point estimate; refer to Adão et al. (2019) for more details.

Embedded Systems, which include technologies related to Industrial Automation and IoT, is the only technology family exhibiting a negative coefficient in the baseline specification. When we control for the exposure to all other technologies with which Embedded Systems may co-integrate, the coefficient remains negative and insignificant. However, we observe that standard errors decrease and the confidence interval shrinks, suggesting that co-integration of Embedded Systems with other emerging technologies is important. Conversely, the other tangible technology, namely 3D Printing, demonstrates a significantly positive effect on employment, which does not persist when we account for the exposure to all other technologies.

We observe a positive and significant effect on regional employment from exposure to all other technology families in the baseline specification. However, when controlling for the exposure to other technologies, the coefficients for Computer Vision, Food Services, Payment Systems, and Digital Services become insignificant. That is, the effect of regional exposure to technologies in these families is balanced by the exposure to other emerging technologies that may complement them.

In contrast, the impact of E-Commerce, Smart Mobility, and HealthTech on the employment-to-population ratio remains significant when accounting for exposure to all other technologies. However, their signs differ. On the one hand, the E-Commerce coefficient turns negative (and is significant at the 10% level), suggesting a negative impact of E-Commerce on regional employment. On the other hand, both Smart Mobility and HealthTech coefficients remain positive and significant, indicating a positive impact on regional employment.

These findings prompt a more detailed investigation at the individual technology level, raising questions such as which specific technologies within technology families might be driving these results. In the following section, we delve into the most granular layer of our analysis, focusing on the regional exposure to individual emerging digital technologies.

### 5.3 Disentangling the Individual Effects of Emerging Digital Technologies

To estimate the individual effects of regional exposure to each emerging digital technology on employment, we use the same shift-share strategy previously described, applying it independently to each technology (see Table 1 for the full list). The regional exposure to a specific technology $k$ is represented by:

$$X_r^k = \sum_j l_{rj} X_j^k,$$

where $l_{rj}$ denotes the employment share of the sector $j$ in region $r$, and $X_j^k$ is the exposure of sector $j$ to technology $k$.

We proceed to estimate the causal impact of regional exposure to an individual emerging digital technology on the regional employment-to-population ratio with the following empirical specification:

$$\Delta Y_r = \alpha + \beta_k X_r^k + \gamma_{1k} X_r^{K\setminus\{k\}} + \gamma_{2k} X_r^{-K} + Z\delta + \phi_{c(r)} + u_r, \tag{17}$$

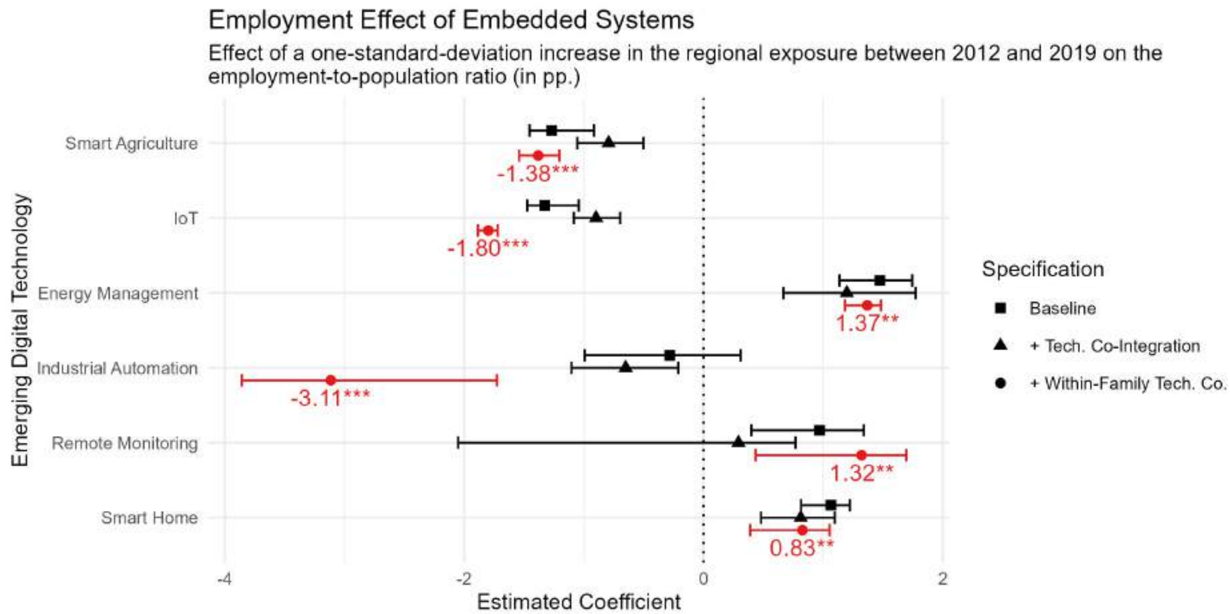
where $X_r^k$ is our variable of interest, $X_r^{K\setminus\{k\}}$ represents the regional exposure to all other technologies within the same family (excluding the one of interest), $X_r^{-K}$ indicates the regional exposure to all remaining emerging digital technologies,[22] and $Z$ includes the same set of covariates as in Equation (15).

Our estimated coefficient of interest, denoted as $\hat{\beta}_k$, represents the employment effect, measured in pp. change, of a one-standard-deviation increase in the regional exposure to a specific emerging digital technology $k$, conditional on the regional exposure to both its technology family and all other emerging technologies.

This specification enables us to address three confounding factors associated with technology co-integration, which could otherwise bias our estimation of the causal effect on regional employment. Firstly, by including $X_r^{K\setminus\{k\}}$ as a control, we mitigate the confounding influence of closely related emerging technologies within the same family. Secondly, incorporating $X_r^{-K}$ helps control for the overall impact of emerging digital digital technologies. Thirdly, by simultaneously accounting for both $X_r^{K\setminus\{k\}}$ and $X_r^{-K}$, we address potential complementarities between the technology family and all other emerging digital technologies, considering their co-integration at the regional level. Consequently, $\hat{\beta}_k$ reflects the causal effect of regional exposure to a specific technology, independent of any co-integration effects with other emerg-

[22] Both $X_r^{K\setminus\{k\}}$ and $X_r^{-K}$ are calculated as shift-share variables. The former is computed for the corresponding technology family $K$, excluding the technology under consideration $k$; the latter is computed for all remaining emerging digital technologies, thereby excluding both the technology of interest as well as its family.

Figure 9: Employment Effect of Emerging Digital Technologies in Embedded Systems




*Notes*: This figure presents the coefficients measuring the effect of regional exposure to emerging digital technology, constructed as shift-shares and standardized, on changes in the employment-to-population ratio, expressed in percentage points (pp.), between 2012 and 2019 in European regions. The confidence intervals are reported at the 5% significance level using the AKM0 inference procedure from Adão et al. (2019). Regressions are weighted by population in 2010. The analysis includes three empirical specifications: the first is the baseline, incorporating country fixed effects and regional controls in 2010 (including the logarithm of population, the proportion of females, the proportion of the population aged over 65, the proportions of the population with secondary and tertiary education levels, and the share of employment in the industry sector); the second specification accounts for technology co-integration by controlling for the regional exposure to all other emerging technologies, also constructed as a shift-share; the third specification accounts for within-family technology co-integration by controlling for the regional exposure to all other emerging technologies belonging to the same family, also constructed as a shift-share.

ing digital technologies or combinations thereof. We present our results by technology family.

**Embedded Systems.** Figure 9 presents the estimated coefficients along with their 95% confidence intervals for technologies within the Embedded Systems family. This figure includes the same initial two specifications detailed in Figure 8, along with a third specification as presented in Equation (17).

Industrial Automation, encompassing robots, shows no significant effect on employment in the baseline specification. However, the inclusion of technology co-integration in the second specification turns the coefficient negative, showing an adverse impact on employment. A more in-depth consideration of technology co-integration in the third specification reveals a more markedly negative impact of Industrial Automation on employment.

The patterns for other technologies are less dependent on exposure to potentially complementary technologies within and outside the family of Embedded Systems. Both Smart Agriculture and IoT demonstrate a significant negative impact on employment. In contrast, Remote Monitoring, Energy Management, and Smart Home have a positive impact on employment. The fluctuation of the coefficient associated with Remote Monitoring, becoming insignificant in the second specification and then positive and significant again in the third, suggests that its positive impact on employment occurs when co-integrated with other Embedded Systems.

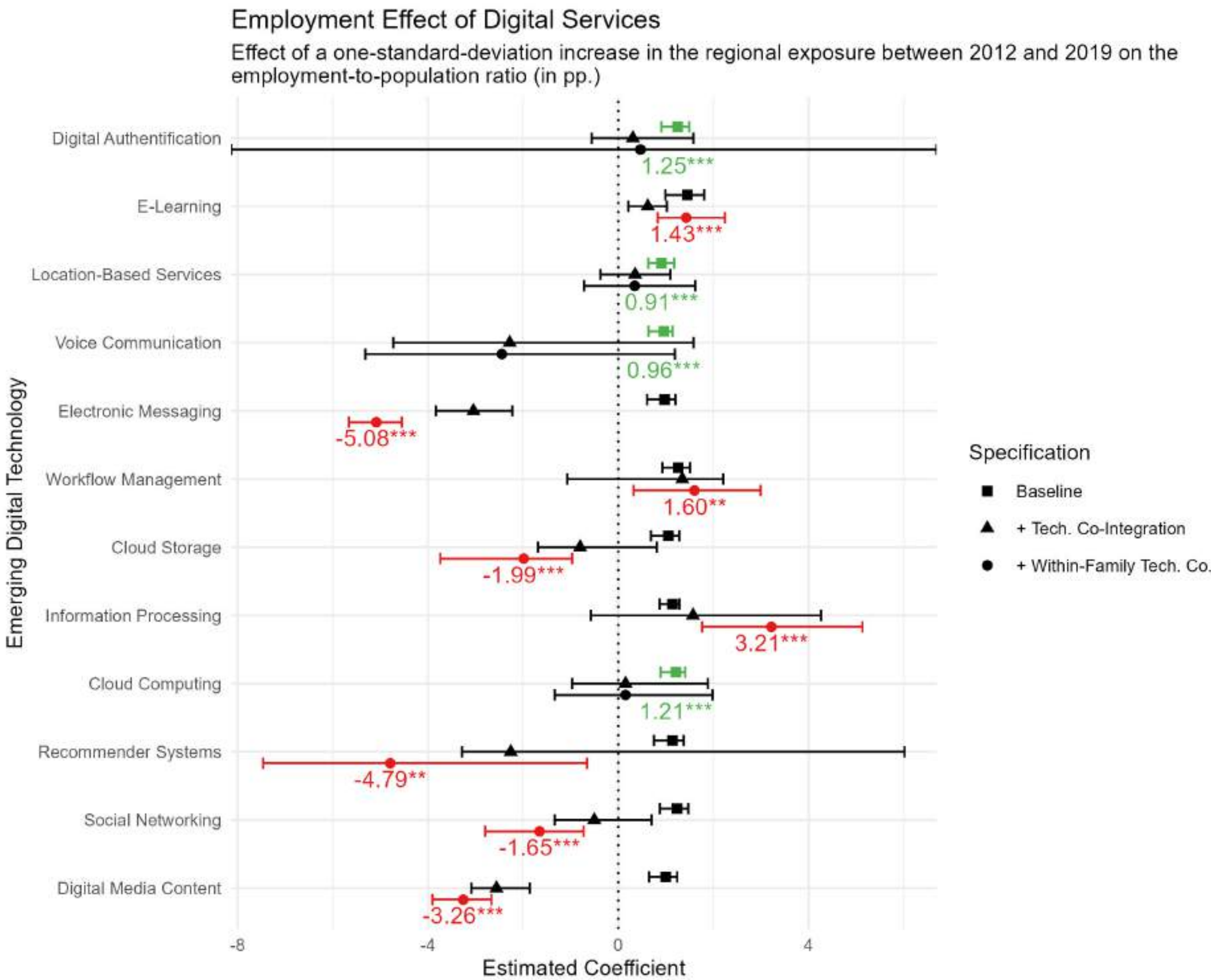
Taken together, the insignificant impact of regional exposure to Embedded Systems technologies, as shown in Figure 8, reflects the substantial heterogeneity in labor market impacts of individual technologies belonging to Embedded Systems. These results might be influenced by the mix of industries exposed to each technology. For example, manufacturing is predominantly exposed to Industrial Automation, while agriculture is the primary sector exposed to both Smart Agriculture and IoT. Our findings indicate that, in these sectors, these physical technologies tend to replace workers, with the impact observable at the regional level. In contrast, it is primarily the service sectors, construction, and information and communication that are most exposed to Remote Monitoring. Our analysis suggests that Remote Monitoring and Energy Management tend to complement workers rather than replace them. This is particularly true for the information and communication sector, which plays a role in producing some of these technologies.

**Digital Services.** Figure 10 presents the effects of exposure to individual emerging technologies within Digital Services on regional employment. In the baseline specification, a positive and significant impact is observed for all technologies.

However, when we account for the co-integration of Digital Services with all other emerging digital technologies in the second specification, we observe a negative impact on employment for Electronic Messaging and Digital Media Content and a positive one for E-Learning. Yet, coefficients associated with all other technologies become insignificant. This aligns with the results obtained at the technology family level (Fig. 8). These findings suggest that the co-integration of Digital Services with other emerging digital technologies significantly influences their impact on the labor market.

Results from the third specification, which accounts for technology co-integration within Digital Services, reveal the impact of these emerging digital technologies on employment. We find that Workflow Management, Information Processing, and E-Learning have a positive impact on employment, suggesting that the productivity effect of these technologies on labor is dominant. Conversely, we find that Cloud Storage, Recommender Systems, and Social Networking now have a negative impact on employment, along with Electronic Messaging and Digital Media Content, suggesting that the displacement effect of these is dominant. Lastly, we find no impact of Location-Based Services, Cloud Computing, and Voice Communication.

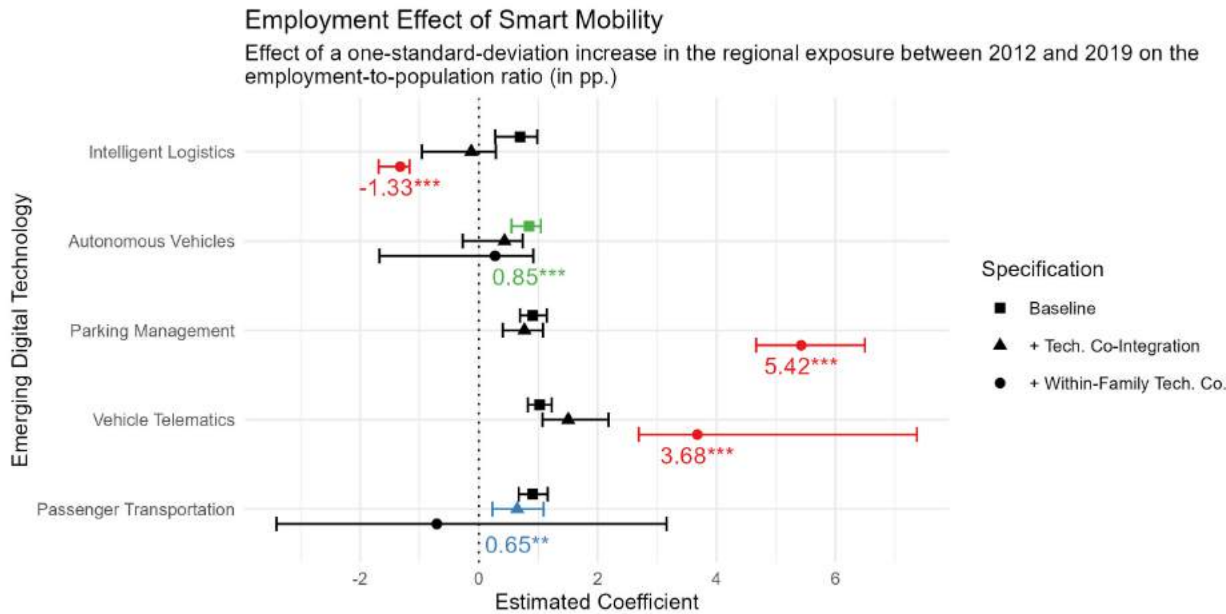
Figure 10: Employment Effect of Emerging Digital Technologies in Digital Services




*Notes*: This figure presents the coefficients measuring the effect of regional exposure to emerging digital technology, constructed as shift-shares and standardized, on changes in the employment-to-population ratio, expressed in percentage points (pp.), between 2012 and 2019 in European regions. The confidence intervals are reported at the 5% significance level using the AKM0 inference procedure from Adão et al. (2019). Regressions are weighted by population in 2010. The analysis includes three empirical specifications: the first is the baseline, incorporating country fixed effects and regional controls in 2010 (including the logarithm of population, the proportion of females, the proportion of the population aged over 65, the proportions of the population with secondary and tertiary education levels, and the share of employment in the industry sector); the second specification accounts for technology co-integration by controlling for the regional exposure to all other emerging technologies, also constructed as a shift-share; the third specification accounts for within-family technology co-integration by controlling for the regional exposure to all other emerging technologies belonging to the same family, also constructed as a shift-share.

**Smart Mobility.** Figure 11 presents the effects of Smart Mobility technologies on employment. All coefficients are positive and significant in the first specification, suggesting a positive employment effect of Smart Mobility technologies. When accounting for co-integration with all other emerging digital technologies at the regional level, in the second specification, the coefficients of Intelligent Logistics and Autonomous Vehicles turn insignificant. This indicates that the positive impact of Smart Mobility on employment, as shown in Figure 8, is driven by the three other technologies, namely, Vehicle Telematics, Passenger Transportation, and Parking Management.

Figure 11: Employment Effect of Emerging Digital Technologies in Smart Mobility



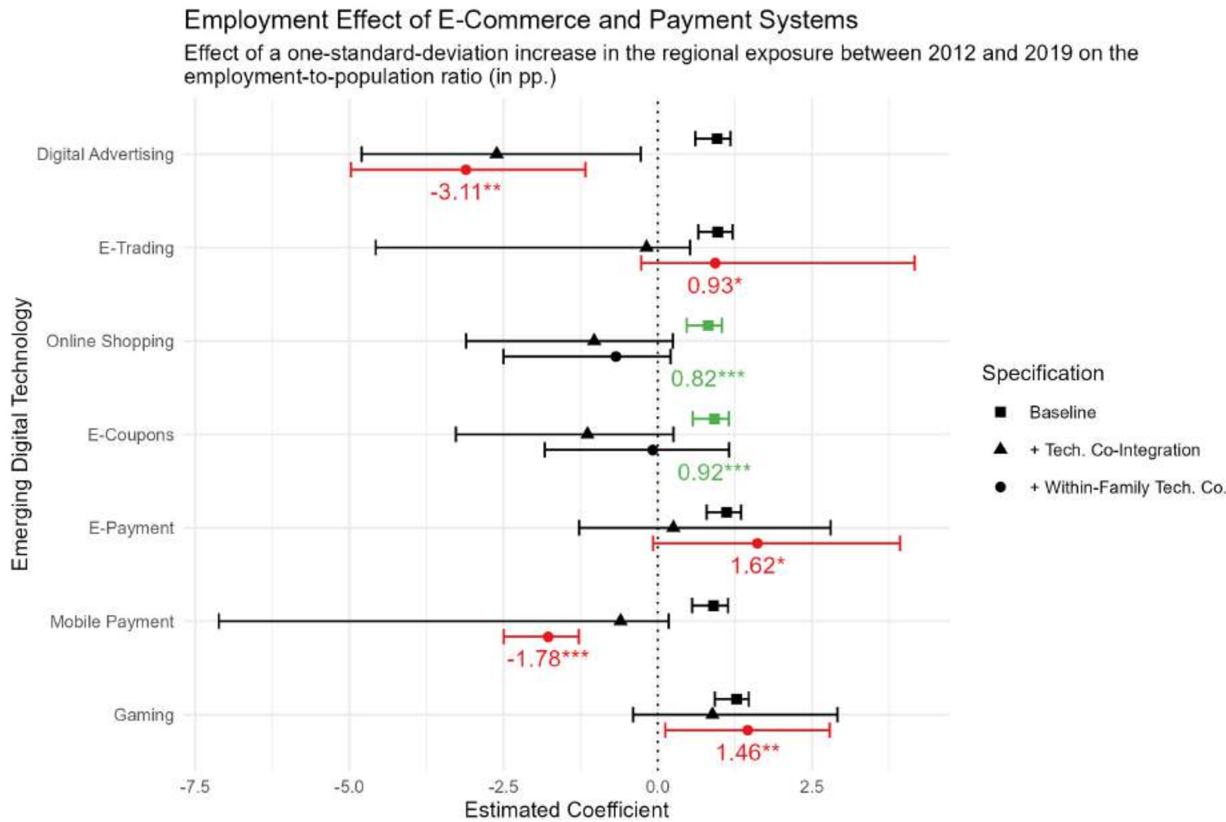

*Notes*: This figure presents the coefficients measuring the effect of regional exposure to emerging digital technology, constructed as shift-shares and standardized, on changes in the employment-to-population ratio, expressed in percentage points (pp.), between 2012 and 2019 in European regions. The confidence intervals are reported at the 5% significance level using the AKM0 inference procedure from Adão et al. (2019). Regressions are weighted by population in 2010. The analysis includes three empirical specifications: the first is the baseline, incorporating country fixed effects and regional controls in 2010 (including the logarithm of population, the proportion of females, the proportion of the population aged over 65, the proportions of the population with secondary and tertiary education levels, and the share of employment in the industry sector); the second specification accounts for technology co-integration by controlling for the regional exposure to all other emerging technologies, also constructed as a shift-share; the third specification accounts for within-family technology co-integration by controlling for the regional exposure to all other emerging technologies belonging to the same family, also constructed as a shift-share.

The third specification demonstrates that Intelligent Logistics, encompassing mobile robots, negatively impacts employment. The adverse displacement effect of this latter technology is counterbalanced by the beneficial productivity and reinstatement effects of other Smart Mobility technologies on employment throughout the entire technology family, specifically, Vehicle Telematics and Parking Management.

**E-Commerce and Payment Systems.** Figure 12 presents the employment effects of technologies from the E-Commerce and Payment Systems technology families. We group these two families due to their closely related purposes and the relatively small number of emerging digital technologies they encompass. In the first specification, all coefficients are significantly positive, suggesting a positive impact on employment.

However, when accounting for co-integration with other emerging technologies in the second specification, all coefficients turn insignificant, except for the one associated with Digital Advertising, which becomes negatively significant. This aligns with the estimates at the technology family level, as shown in Figure 8, indicating no substantial impact of both technology

Figure 12: Employment Effect of Emerging Digital Technologies in E-Commerce and Payment Systems



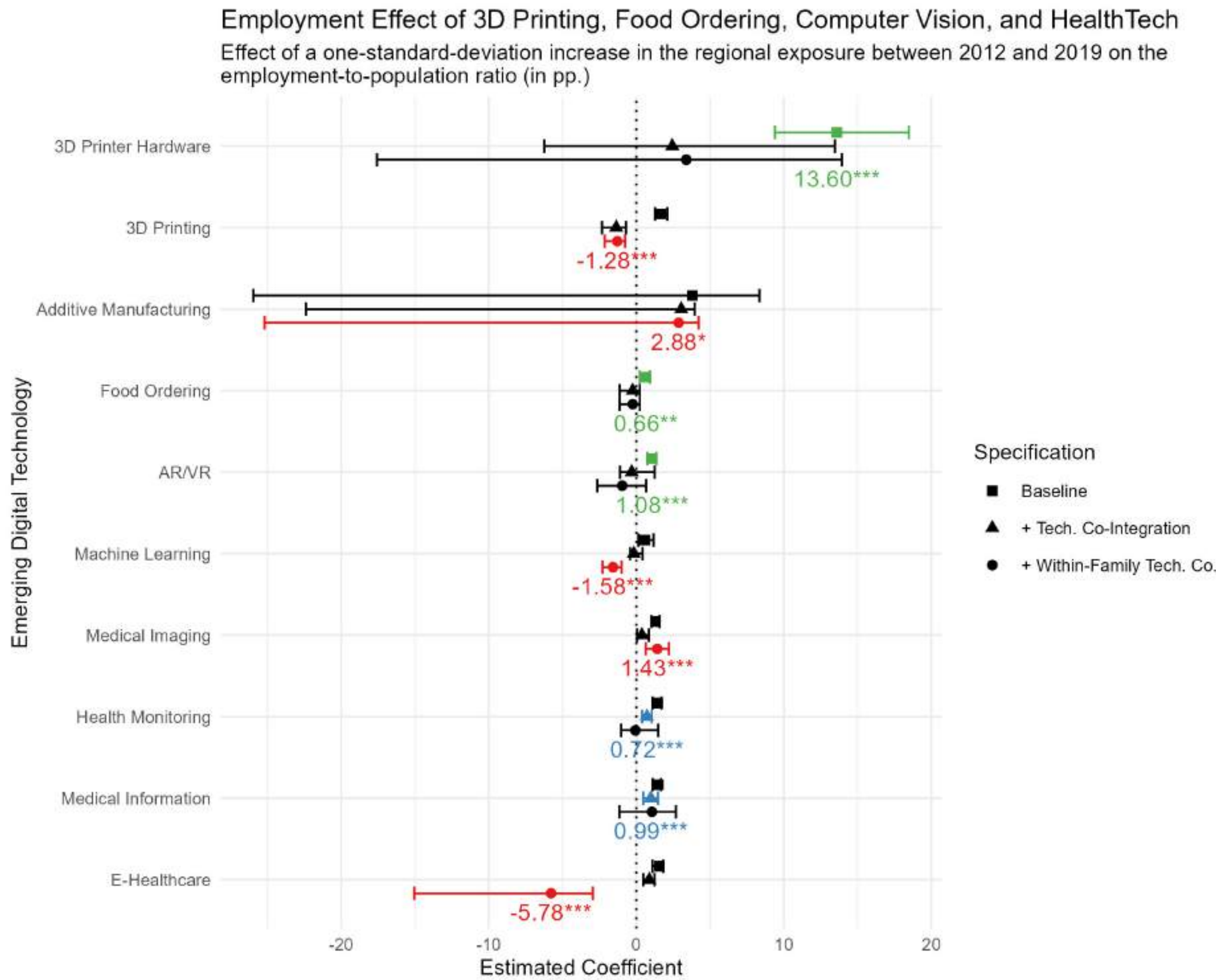

*Notes*: This figure presents the coefficients measuring the effect of regional exposure to emerging digital technology, constructed as shift-shares and standardized, on changes in the employment-to-population ratio, expressed in percentage points (pp.), between 2012 and 2019 in European regions. The confidence intervals are reported at the 5% significance level using the AKM0 inference procedure from Adão et al. (2019). Regressions are weighted by population in 2010. The analysis includes three empirical specifications: the first is the baseline, incorporating country fixed effects and regional controls in 2010 (including the logarithm of population, the proportion of females, the proportion of the population aged over 65, the proportions of the population with secondary and tertiary education levels, and the share of employment in the industry sector); the second specification accounts for technology co-integration by controlling for the regional exposure to all other emerging technologies, also constructed as a shift-share; the third specification accounts for within-family technology co-integration by controlling for the regional exposure to all other emerging technologies belonging to the same family, also constructed as a shift-share.

families on regional employment.

The third specification indicates that Mobile Payment negatively affects employment, thereby demonstrating that the displacement effect predominates for this technology. Conversely, the same specification reveals that Gaming and E-Trading positively influence employment, with the latter being significant at the 10% level.

**3D Printing.** Figure 13 presents the employment effects of emerging digital technologies within the 3D Printing, Computer Vision, and HealthTech technology families. We start by focusing on 3D Printing technologies, identified as the three technologies at the top of the figure. The first specification suggests a positive and significant impact on employment for

Figure 13: Employment Effect of Emerging Digital Technologies in 3D Printing, Computer Vision, and HealthTech




*Notes*: This figure presents the coefficients measuring the effect of regional exposure to emerging digital technology, constructed as shift-shares and standardized, on changes in the employment-to-population ratio, expressed in percentage points (pp.), between 2012 and 2019 in European regions. The confidence intervals are reported at the 5% significance level using the AKM0 inference procedure from Adão et al. (2019). Regressions are weighted by population in 2010. The analysis includes three empirical specifications: the first is the baseline, incorporating country fixed effects and regional controls in 2010 (including the logarithm of population, the proportion of females, the proportion of the population aged over 65, the proportions of the population with secondary and tertiary education levels, and the share of employment in the industry sector); the second specification accounts for technology co-integration by controlling for the regional exposure to all other emerging technologies, also constructed as a shift-share; the third specification accounts for within-family technology co-integration by controlling for the regional exposure to all other emerging technologies belonging to the same family, also constructed as a shift-share.

3D Printer Hardware and 3D Printing. However, in the second specification, which accounts for co-integration with other emerging technologies, the employment impact of 3D Printing shifts to negative, while the impact of the hardware component becomes insignificant. This observation is consistent with the results shown in Figure 8, where the overall employment effect of 3D Printing is initially positive but becomes insignificant upon considering technology co-integration.

In the third specification, which accounts for the within-family co-integration, the coefficient for 3D Printing, representing the systems that produce 3D objects, remains significantly negative. This indicates that such technology predominantly contributes to the displacement effect in 3D Printing.

**Computer Vision.** The subsequent three emerging digital technologies presented in Figure 13 belong to the Computer Vision family. The baseline specification reveals their positive effects on employment when considered as stand-alone technologies.

Upon accounting for technology co-integration in the second specification, the coefficients for both Machine Learning and AR/VR turn insignificant, while the Medical Imaging coefficient remains positive. This aligns with the findings at the technology family level in Figure 8, which indicate no employment effect for Computer Vision when considering technology co-integration with other emerging digital technologies at the regional level.

In the third specification, which additionally considers co-integration with other Computer Vision technologies, the estimates suggest that Medical Imaging positively impacts employment, while Machine Learning shows a negative effect.

**HealthTech.** Lastly, we examine the three HealthTech technologies situated at the bottom of Figure 13. We observe that incorporating technology co-integration with all other emerging digital technologies in the second specification does not alter the positive impact of HealthTech on employment, as established in the first specification. This corresponds with the estimate in Figure 8. However, in the third specification, the coefficients for Medical Information and Health Monitoring turn non-significant, while E-Healthcare exhibits a substantial negative effect on employment, emphasizing the displacement effect of this particular technology on labor.

# 6 Conclusion

Recent developments in digital technologies, notably AI, have raised public and academic interest in the impact of emerging digital technologies on future employment. Determining whether these technologies will create more jobs than they eliminate is a crucial issue for both individuals and policymakers. However, prior research has largely focused on analyzing either very specific technologies, such as industrial robots or certain applications of AI, or a diverse array of digital technologies commonly labeled as "automation technologies".

In this paper, we measure the exposure of industries and occupations to 40 digital technologies that have emerged over the past decade and investigate their effects on European employment. Using state-of-the-art NLP tools, such as sentence transformers, we introduce a novel methodology to measure the exposure of industries and occupations at a granular level. We have made our pioneering data available as an open–access resource, named the 'TechXposure' database. Using this new data source, we estimate the employment impact of these emerging digital technologies. Our main findings reveal that emerging digital technologies have an overall positive impact on the employment-to-population ratio, thereby creating employment opportunities rather than destroying jobs. However, when examining the specific effects of these technologies, we find considerable heterogeneity in the employment impact of these technologies, along with a significant role of emerging technology co-integration at the regional level. Yet, our paper does not address the question of the quality of these employment opportunities, which is a research question we intend to investigate in the future.

We highlight the advantages and limitations of our exposure scores present in the 'TechXposure' database. First, since our exposure scores are based on text data from standard European classifications, they are universal and not influenced by any specific European country. Second, our method does not rely on keywords (or tokens) and therefore only requires a set of relevant patents, making it replicable in other contexts, such as for green technologies or using future ISCO/NACE classifications. However, our exposure scores do not account for the augmentation or automation effect on occupations and industries; they solely reflect the relevance of technologies to a given industry or occupation. This limitation in capturing their employment effects allows us to make fewer assumptions in data construction, leaving the question open as some technologies may have positive effects on employment in one context and negative ones in another. Additionally, our set of technologies does not include recent developments in Large Language Models (LLM), such as ChatGPT, as our analysis period focuses on technologies that emerged until 2021. However, our set does include several other applications of AI, specifically in areas such as Machine Learning (for computer vision), Information Processing, and Workflow Management. Lastly, our exposure metrics do not measure the adoption of these emerging digital technologies, which is a topic we intend to address in future research.

We regard our paper as a foundational contribution to new avenues for future research on technological change and labor markets. By constructing this open–access database, we anticipate that future studies will greatly benefit from its use. It offers an unprecedented level of detail in analyzing the exposure of occupations and industries to emerging technologies, encompassing not only those frequently discussed in economic literature, such as robots and AI, but also less-studied technologies like social networks, cloud technologies, and health technologies. Given that our database is based on European classifications of occupations and industries, it presents a valuable opportunity for research focused on Europe. This research could provide deeper insights into the impact of emerging digital technologies on the economy, particularly considering Europe's rich diversity in institutional contexts that may significantly influence technology development, adoption, and labor market effects. We believe our database is user-friendly and accessible for both researchers and policymakers.

# References

Acemoglu, D., Autor, D., Hazell, J., and Restrepo, P. (2022). Artificial Intelligence and Jobs: Evidence from Online Vacancies. *Journal of Labor Economics*, 40(S1):S293–S340.

Acemoglu, D. and Restrepo, P. (2018). The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment. *American Economic Review*, 108:1488–1542.

Acemoglu, D. and Restrepo, P. (2019). Automation and New Tasks: How Technology Displaces and Reinstates Labor. *Journal of Economic Perspectives*, 33:3–30.

Acemoglu, D. and Restrepo, P. (2020). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy*, 128:2188–2244.

Adão, R., Kolesár, M., and Morales, E. (2019). Shift-Share Designs: Theory and Inference. *The Quarterly Journal of Economics*, 134(4):1949–2010.

Aghion, P., Antonin, C., Bunel, S., and Jaravel, X. (2021). What are the Labor and Product Market Effects of Automation? New Evidence from France.

Akerman, A., Gaarder, I., and Mogstad, M. (2015). The Skill Complementarity of Broadband Internet. *The Quarterly Journal of Economics*, 130(4):1781–1824.

Alekseeva, L., Azar, J., Giné, M., Samila, S., and Taska, B. (2021). The Demand for AI Skills in the Labor Market. *Labour Economics*, 71:102002.

Autor, D., Chin, C., Salomons, A., and Seegmiller, B. (2022). New Frontiers: The Origins and Content of New Work, 1940–2018. *The Quarterly Journal of Economics*, Forthcoming.

Autor, D. H., Katz, L. F., and Kearney, M. S. (2006). The Polarization of the U.S. Labor Market. *American Economic Review*, 96(2):189–194.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, 118:1279–1333.

Borusyak, K., Hull, P., and Jaravel, X. (2021). Quasi-Experimental Shift-Share Research Designs. *The Review of Economic Studies*, 89(1):181–213.

Chaturvedi, S., Ciarli, T., Nomaler, O., and Prytkova, E. (2023). What's the Future of Automation? Using Semantic Analysis to Identify Emerging Technologies. Working paper.

Dauth, W., Findeisen, S., Suedekum, J., and Woessner, N. (2021). The Adjustment of Labor Markets to Robots. *Journal of the European Economic Association*, 19:3104–3153.

Davies, D. L. and Bouldin, D. W. (1979). A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1(2):224–227.

Dechezleprêtre, A., Hémous, D., Olsen, M., and Zanella, C. (2021). Induced Automation: Evidence from Firm-level Patent Data. *SSRN Electronic Journal*.

Felten, E., Raj, M., and Seamans, R. (2021). Occupational, Industry, and Geographic Exposure to Artificial Intelligence: A Novel Dataset and its Potential Uses. *Strategic Management Journal*, 42(12):2195–2217.

Felten, E. W., Raj, M., and Seamans, R. (2018). A Method to Link Advances in Artificial Intelligence to Occupational Abilities. *AEA Papers and Proceedings*, 108:54–57.

Frey, C. B. and Osborne, M. A. (2017). The Future of Employment: How Susceptible are Jobs to Computerisation? *Technological Forecasting and Social Change*, 114:254–280.

Gentzkow, M., Kelly, B., and Taddy, M. (2019). Text as Data. *Journal of Economic Literature*, 57(3):535–74.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik Instruments: What, When, Why, and How. *American Economic Review*, 110(8):2586–2624.

Goos, M. and Manning, A. (2007). Lousy and Lovely Jobs: The Rising Polarization of Work in Britain. *Review of Economics and Statistics*, 89:118–133.

Goos, M., Manning, A., and Salomons, A. (2009). Job Polarization in Europe. *American Economic Review*, 99:58–63.

Goos, M., Manning, A., and Salomons, A. (2014). Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review*, 104:2509–2526.

Graetz, G. and Michaels, G. (2018). Robots at work. *Review of Economics and Statistics*, 100:753–768.

Hall, B., Jaffe, A., and Trajtenberg, M. (2005). Market Value and Patent Citations: A First Look. *The RAND Journal of Economics*, 36(1):16–38.

Kelly, B., Papanikolaou, D., Seru, A., and Taddy, M. (2021). Measuring Technological Innovation over the Long Run. *American Economic Review: Insights*, 3(3):303–320.

Kogan, L., Papanikolaou, D., Schmidt, L., and Seegmiller, B. (2021). Technology-Skill Complementarity and Labor Displacement: Evidence from Linking Two Centuries of Patents with Occupations. Technical report.

Mann, K. and Püttmann, L. (2023). Benign Effects of Automation: New Evidence from Patent Texts. *Review of Economics and Statistics*, 105(3):562–579.

Michaels, G., Natraj, A., and Reenen, J. V. V. (2014). Has ICT Polarized Skill Demand? Evidence from Eleven Countries over Twenty-Five Years. *Review of Economics and Statistics*, 96:60–77.

OECD (2009). OECD Patent Statistics Manual.

Reimers, N. and Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *arXiv preprint arXiv:1908.10084*.

Vries, G. J. d., Gentile, E., Miroudot, S., and Wacker, K. M. (2020). The Rise of Robots and the Fall of Routine Jobs. *Labour Economics*, 66:101885.

Webb, M. (2019). The Impact of Artificial Intelligence on the Labor Market. Working paper.

# Appendices

## A Data Appendix

In this Appendix, we provide additional information on the set of emerging digital technologies and the derivation of exposure scores. Appendix A.1 describes the set of technologies. Appendix A.2 presents the manual exclusions implemented to enhance our exposure metrics. Appendix A.4 offers further details about the citation-based weighting scheme. Appendix A.5 explains the semantic co-occurrence of technologies across occupations, which is instrumental in classifying technologies into families. Appendix A.6 outlines the methodology for computing exposure scores at higher levels of aggregation in the ISCO and NACE classifications.

### A.1 Description of Emerging Digital Technologies

Tables A.1 to A.3 present the 40 emerging digital technologies from the TechXposure database as well as their descriptions.

### A.2 Manual Exclusions

**Industry.** For industries, we make the following manual adjustments:

- We exclude the exposure scores that relate to 'Printing and service activities related to printing' (18.1) due to the persistent conflation of its intended meaning (i.e. printing products with text, symbols (e.g. musical notation), and imagery (e.g. maps, engraving, etc.)) with emerging digital technologies.
- We exclude the sentence "*manufacture of computer printout paper ready for use*" (Sentence ID 17.2_11) from the industry description text of 'Manufacture of articles of paper and paperboard' (17.2) when combining tasks with patents belonging to the technologies within the 3D Printing family.
- We exclude the sentence "*units giving this type of instructions might be named "schools", "studios", "classes" etc.*" (Sentence ID 85.5_17) from the industry description text of 'Other education' (85.5) when combining tasks with patents belonging to the technology Machine Learning.

**Occupation.** For occupations, we make the following manual adjustments:

- Analogously with industry 18.1, we exclude the exposure scores that relate to 'Printing trades workers' (732) and its nested occupations (7321, 7322, 7323) due to the persistent conflation of its intended meaning with emerging digital technologies.

- We exclude the task "*creating the blueprint or pattern pieces for a particular apparel design with the aid of a computer;*" (Task ID 7532_2) from the occupation description text of 'Printers' (7532) when combining tasks with patents belonging to the technology Machine Learning.
- We exclude the task "*preparing and developing instructional training material and aids such as handbooks, visual aids, online tutorials, demonstration models and supporting training reference documentation;*" (Task ID 2424_3) from the occupation description text of 'Training and staff development professionals' (2424) when combining tasks with patents belonging to the technology Machine Learning.

### A.3 Redundancy Filtering Examples

Tables A.4 and A.5 present additional examples of redundancy filtering for industries. Tables A.6 to A.8 present examples of redundancy filtering for occupations.

### A.4 Distribution of Patents and Citation-based Weighting Scheme

Figure A.1 presents the distribution of patents across emerging digital technologies. Figure A.3 presents the distribution of non-cited and undetermined-count patents across emerging digital technologies. Figure A.2 presents the log distribution of patent citations across emerging digital technologies. Figure A.4 presents the correlation between citation-weighted and unweighted yearly cosine similarity scores for both industries and occupations.

### A.5 Technology Co-Occurrence

Using our cosine similarity scores, we examine the semantic co-occurrence of emerging digital technologies across occupations. Let $C^k_{\mathcal{O}} = (C^k_1, \dots, C^k_o, \dots, C^k_O)$ represent the vector of cosine similarity scores for all occupations related to technology $k$. We define the pairwise semantic-based technology co-occurrence as the correlation between $C^k_{\mathcal{O}}$ and $C^{k'}_{\mathcal{O}}$ for each pair of technologies $(k, k')$. These pairwise correlations are computed for all technologies using semantic similarity scores at the 3-digit occupational level.

Figure A.5 presents the result of technology grouping based on cosine semantic scores. We observe a distinct segmentation within the figure, categorized as 'technology families'. Starting from the top-left corner and moving along the diagonal, the first group encountered includes technologies related to 3D Printing. Subsequent to this, the range from Smart Agriculture to Smart Home falls within the Embedded Systems family. A significant block then emerges, spanning from Intelligent Logistics to Passenger Transportation, and encompasses

Smart Mobility technologies. Following this, a standalone block dedicated to Food Ordering appears. The next two blocks represent E-Commerce and Payment Systems, respectively. This sequence is succeeded by the most extensive block, which includes 12 technologies and relates to Digital Services. Afterward, AR/VR, Machine Learning, and Medical Imaging are grouped under Computer Vision technologies. Finally, the figure concludes with HealthTech technologies.

### A.6 Exposure Scores at Higher Levels of Aggregation

To calculate exposure scores at higher levels of aggregation within the ISCO and NACE classifications, we apply the inverse hyperbolic sine transformation to the average cosine similarity score aggregated across all industries/occupations from the most granular classification level up to the level of interest.

For example, consider the derivation of the exposure score for a 1-digit NACE industry $I \subset \mathcal{I}$ to an emerging digital technology $k$. We begin with the cosine similarity score, aggregating it to a higher level of classification as follows:

$$C_I^k = \frac{1}{|I|} \sum_{i \in I} C_i^k,$$

where $C_i^k$ is cosine similarity score between a 3–digit industry $i$ (belonging to the 1-digit industry $I$) and technology $k$, as obtained in Equation (12). We then apply the inverse hyperbolic sine transformation to obtain the exposure score, namely, $X_I^k = \sinh^{-1}\left(C_I^k\right)$. This methodology is similarly employed to derive exposure scores for 2-digit industries, as well as for occupation exposures at higher levels of aggregation.

Table A.1: Description of the Emerging Digital Technologies (1/3)

| | Technology | Description |
|---|---|---|
| 1 | 3D Printer Hardware | Three-dimensional printers and their components, such as printing heads, pens, nozzles, platforms, and devices for printing, extruding, cleaning, recycling, heating, and cooling. |
| 2 | 3D Printing | Printing systems for creating three-dimensional objects using a variety of materials and techniques, like photocuring and powder spreading. |
| 3 | Additive Manufacturing | Technologies and processes for additive manufacturing, with applications such as prostheses and building materials. |
| 4 | Smart Agriculture & Water Management | Various Internet of Things (IoT) technologies for intelligent and remote management in agriculture, and water and sewage systems. |
| 5 | Internet of Things (IoT) | Systems and devices interconnected via IoT for data collection, remote control, and real-time monitoring in diverse applications, including agriculture, home automation, and environmental monitoring. |
| 6 | Predictive Energy Management and Distribution | A combination of network, data management, and AI technologies for monitoring, distribution, and efficient use of electrical power and energy, including renewable energy sources, and for consumption prediction in intelligent power management. |
| 7 | Industrial Automation & Robot Control | Industrial process automation, including robots, programmable logic controllers, and related control apparatuses such as remote control and fault diagnosis. |
| 8 | Remote Monitoring & Control Systems | Real-time remote monitoring and management technologies for factories, building management, warehouses, intelligent homes, disaster management, and network security. |
| 9 | Smart Home & Intelligent Household Control | Various IoT technologies for the intelligent control of homes and buildings, including household appliances, home environments, and smart home integrations, often utilizing wireless communication and monitoring. |
| 10 | Intelligent Logistics | A combination of monitoring, remote control technologies, data acquisition, and mobile robot technologies for logistics and delivery applications, including supply chain management, warehouse operations, package tracking, and courier services. |
| 11 | Autonomous Vehicles & UAVs | Developments in unmanned aerial vehicles (UAVs), drones, and autonomous driving technologies, with an emphasis on vehicle control, navigation, and sensor integration. |
| 12 | Parking & Vehicle Space Management | Networking technologies for parking management, including systems for monitoring available spaces and intelligent parking solutions. |
| 13 | Vehicle Telematics & Electric Vehicle Management | Technologies for intra-vehicle information management, especially in electric vehicles, including aspects of real-time monitoring, traffic information, and vehicle diagnostics. |
| 14 | Passenger Transportation | Technologies for ride-sharing, taxi hailing, and public transportation reservations using real-time information, electronic ticketing, and route optimization. |

*Notes*: This table provides descriptions of emerging digital technologies ranging from 1 to 14.

Table A.2: Description of the Emerging Digital Technologies (2/3)

| | Technology | Description |
|---|---|---|
| 15 | Food Ordering & Vending Systems | Wireless infrastructures, encryption, monitoring, and remote control technologies for food order management, such as automatic vending, self-service ordering, meal preparation, and delivery. |
| 16 | Digital Advertising | Automated tracing and tagging, and AI technologies for digital advertisements, including targeted delivery on mobile devices. |
| 17 | Electronic Trading and Auctions | Online trading platforms, financial instrument exchanges, and auction mechanisms, focusing on real-time bidding, trading, and market data. |
| 18 | Online Shopping Platforms | Wireless technologies (e.g., RFID and mobile terminals), encryption (e.g., blockchain), and AI technologies for e-commerce transactions, and digital tools related to the purchase, sale, and display of product information, including recommendation systems. |
| 19 | E-Coupons & Promotion Management | Data management platforms for electronic coupon distribution, management, redemption, and associated loyalty programs. |
| 20 | Electronic Payments & Financial Transactions | A combination of wireless (e.g., mobile) and encryption (e.g., blockchain) technologies for processing electronic payments (e.g., credit card transactions) and interfacing with financial institutions. |
| 21 | Mobile Payments | A combination of mobile technologies for processing electronic payments. |
| 22 | Gaming & Wagering Systems | A combination of user interface and data management technologies for gaming, both online and physical, including gambling and gaming machines. |
| 23 | Digital Authentication | Encryption and robotic processing technologies for verifying user identities, securing transactions, and safeguarding data through various authentication mechanisms, such as biometrics and cryptographic methods. |
| 24 | E-Learning | A combination of AI and data management technologies for digital platforms and systems in education, including teaching, learning, and classroom management. |
| 25 | Location-Based Services & Tracking | Technologies that provide location-based content and services, often relying on global positioning and navigation systems and related communication technology. |
| 26 | Voice Communication | Technologies focusing on voice communication, including communication protocols and user interfaces. |
| 27 | Electronic Messaging | Digital communication methods, infrastructure, and user interfaces for services such as email and conferences. |
| 28 | Workflow Management | A combination of AI and network technologies for management applications, including workflow automation, recruitment, event scheduling, and building and property management. |

*Notes*: This table provides descriptions of emerging digital technologies ranging from 15 to 28.

Table A.3: Description of the Emerging Digital Technologies (3/3)

| | Technology | Description |
|---|---|---|
| 29 | Cloud Storage & Data Security | Cloud-based data storage, distributed data management, encryption, and backup, often integrated with blockchain technology. |
| 30 | Information Processing | Systems for managing, processing, and delivering data and information across various domains, potentially including content generation, transmission, and verification. |
| 31 | Cloud Computing | Cloud computing and virtual machines, focusing on cloud platforms and resource allocation in cloud environments. |
| 32 | Recommender Systems | Algorithms and systems for providing recommendations and personalized content delivery based on user behavior, search queries, and similarity metrics. |
| 33 | Social Networking & Media Platforms | User interfaces for online social networking services, content sharing, and recommendation systems. |
| 34 | Digital Media Content | Tools and platforms for digital media content creation, management, distribution, and access. |
| 35 | Augmented and Virtual Reality (AR/VR) | Augmented reality (AR) and virtual reality (VR) models, devices, interfaces, and experiences, including head-mounted displays and interactions in virtual environments. |
| 36 | Machine Learning & Neural Networks | Machine learning training techniques, model architectures, and data processing for computer vision applications. |
| 37 | Medical Imaging & Image Processing | Diverse applications for acquiring and analyzing medical images from various modalities, such as computed tomography (CT), ultrasound, magnetic resonance imaging (MRI), and virtual reality (VR), for purposes including diagnosis, surgical planning, and the design of prostheses. |
| 38 | Health Monitoring | Wearable and implantable devices and systems for real-time health monitoring that track vital signs such as blood pressure, heart rate, and temperature, coupled with comprehensive medical data management. |
| 39 | Medical Information | A combination of data sharing, encryption, and Natural Language Processing (NLP) technologies for the storage, retrieval, and management of medical and patient information, encompassing electronic medical records, prescription management, and remote healthcare services. |
| 40 | E-Healthcare | An integration of data sharing, wireless communication, monitoring, and user interface technologies for healthcare and health management systems, including those used in hospitals and cloud-based platforms. |

*Notes*: This table provides descriptions of emerging digital technologies ranging from 29 to 40.

Table A.4: Example of Redundancy Filtering of Industries for Intelligent Vehicular Control Device

| Code | NACE Industry | Cosine Similarity | | |
|---|---|---|---|---|
| | | $C_i^{p_1}$ | $C_i^{p_2}$ | $C_i^{p}$ |
| 52.2 | Support activities for transportation | 0.531 | 0.454 | 0.489 |
| 49.4 | Freight transport by road and removal services | 0.371 | 0.418 | 0.393 |
| 29.1 | Manufacture of motor vehicles | 0.409 | 0.371 | 0.389 |
| 27.9 | Manufacture of other electrical equipment | 0.358 | 0.375 | 0.366 |
| 30.9 | Manufacture of transport equipment n.e.c. | 0.452 | | |
| 29.2 | Manufacture of bodies (coachwork) for motor vehicles; manufacture of trailers and semi-trailers | 0.389 | | |
| 33.1 | Repair of fabricated metal products, machinery and equipment | 0.379 | | |
| 45.3 | Sale of motor vehicle parts and accessories | 0.377 | | |
| 49.1 | Passenger rail transport, interurban | 0.371 | | |
| 47.3 | Retail sale of automotive fuel in specialised stores | 0.362 | | |
| 26.5 | Manufacture of instruments and appliances for measuring, testing and navigation; watches and clocks | | 0.472 | |
| 26.3 | Manufacture of communication equipment | | 0.434 | |
| 26.2 | Manufacture of computers and peripheral equipment | | 0.410 | |
| 56.1 | Restaurants and mobile food service activities | | 0.392 | |
| 61.2 | Wireless telecommunications activities | | 0.378 | |
| 49.3 | Other passenger land transport | | 0.369 | |

*Notes*: This table presents the redundancy filtering of industries for the Patent ID 201713859U. It displays the cosine similarity of distinct 3-digit NACE Rev.2 industry descriptions with the patent description “Vehicle intelligent logistics control device” (Column 3) and the function principle “GPS locating module for obtaining position information of transport vehicle through main control chip, RFID reader for reading RFID tag information, and 4G module connected with server” (Column 4). Industries are ranked according to Column 3 in decreasing order. Cosine similarity scores in Columns 3 and 4 are displayed only for sub-pairs belonging to their respective top 10. Column 5 shows the composite patent-industry cosine similarity score, which corresponds to the harmonic mean of Columns 3 and 4. Cosine similarity scores in Column 5 are displayed only for pairs that rank simultaneously in both top 10.

Table A.5: Example of Redundancy Filtering of Industries for Speech Recognition System

| Code | NACE Industry | Cosine Similarity | | |
|---|---|---|---|---|
| | | $C_i^{p_1}$ | $C_i^{p_2}$ | $C_i^{p}$ |
| 26.3 | Manufacture of communication equipment | 0.256 | 0.333 | 0.289 |
| 28.2 | Manufacture of other general-purpose machinery | 0.246 | 0.344 | 0.286 |
| 82.9 | Business support service activities n.e.c. | 0.279 | 0.285 | 0.282 |
| 26.4 | Manufacture of consumer electronics | 0.250 | 0.295 | 0.271 |
| 63.9 | Other information service activities | 0.245 | 0.269 | 0.257 |
| 62.0 | Computer programming, consultancy and related activities | 0.276 | | |
| 85.5 | Other education | 0.250 | | |
| 61.9 | Other telecommunications activities | 0.225 | | |
| 58.1 | Publishing of books, periodicals and other publishing activities | 0.224 | | |
| 26.5 | Manufacture of instruments and appliances for measuring, testing and navigation; watches and clocks | | 0.303 | |
| 28.9 | Manufacture of other special-purpose machinery | | 0.294 | |
| 72.1 | Research and experimental development on natural sciences and engineering | | 0.276 | |
| 18.2 | Reproduction of recorded media | | 0.265 | |

*Notes*: This table presents the redundancy filtering of industries for the Patent ID 202048118D. It displays the cosine similarity of distinct 3-digit NACE Rev.2 industry descriptions with the patent description "System for recognizing training speech" (Column 3) and the function principle "process or which is configured to increment counter associated with word sequences, and train language model of automatic transcription system using word sequences and counter" (Column 4). Industries are ranked according to Column 3 in decreasing order. Cosine similarity scores in Columns 3 and 4 are displayed only for sub-pairs belonging to their respective top 10. Column 5 shows the composite patent-industry cosine similarity score, which corresponds to the harmonic mean of Columns 3 and 4. Cosine similarity scores in Column 5 are displayed only for pairs that rank simultaneously in both top 10.

Table A.6: Example of Redundancy Filtering of Occupations for Targeted TV Advertising

| Code | ISCO Occupation | Cosine Similarity | | |
|---|---|---|---|---|
| | | $C^p_{o_1}$ | $C^p_{o_2}$ | $C^p_o$ |
| 2431 | Advertising and marketing professionals | 0.413 | 0.502 | 0.453 |
| 1222 | Advertising and public relations managers | 0.308 | 0.420 | 0.356 |
| 3521 | Broadcasting and audio-visual technicians | 0.274 | 0.380 | 0.318 |
| 3322 | Commercial sales representatives | 0.250 | 0.394 | 0.306 |
| 2434 | ICT sales professionals | 0.297 | | |
| 7422 | ICT installers and servicers | 0.282 | | |
| 4227 | Survey and market research interviewers | 0.279 | | |
| 2656 | Announcers on radio, television and other media | 0.278 | | |
| 1330 | ICT service managers | 0.262 | | |
| 3512 | ICT user support technicians | 0.252 | | |
| 5242 | Sales demonstrators | | 0.396 | |
| 1420 | Retail and wholesale trade managers | | 0.393 | |
| 3432 | Interior designers and decorators | | 0.388 | |
| 2153 | Telecommunications engineers | | 0.374 | |
| 3323 | Buyers | | 0.358 | |
| 9520 | Street vendors (excluding food) | | 0.357 | |

*Notes*: This table presents the redundancy filtering of occupations for the Patent ID 2013B87254 (i.e., "Method for targeting television advertisement based on profile linked to online device, involves selecting television advertisement to be directed to set-top box based on profile information pertaining to user or online activity"). It displays the cosine similarity of the patent title with the 4-digit ISCO-08 title (Column 3) and the task with the highest cosine similarity (Column 4). Occupations are ranked according to Column 3 in decreasing order. Cosine similarity scores in Columns 3 and 4 are displayed only for sub-pairs belonging to their respective top 10. Column 5 shows the composite patent-occupation cosine similarity score, which corresponds to the harmonic mean of Columns 3 and 4. Cosine similarity scores in Column 5 are displayed only for pairs that rank simultaneously in both top 10.

Table A.7: Example of Redundancy Filtering of Occupations for Intelligent Vehicular Control Device

| | | Cosine Similarity | | |
|---|---|---|---|---|
| Code | ISCO Occupation | $C^p_{o_1}$ | $C^p_{o_2}$ | $C^p_o$ |
| 8322 | Car, taxi and van drivers | 0.354 | 0.525 | 0.423 |
| 4323 | Transport clerks | 0.333 | 0.440 | 0.379 |
| 9333 | Freight handlers | 0.333 | 0.420 | 0.371 |
| 9621 | Messengers, package deliverers and luggage porters | 0.308 | 0.412 | 0.353 |
| 8332 | Heavy truck and lorry drivers | 0.301 | 0.405 | 0.345 |
| 7422 | ICT installers and servicers | 0.371 | | |
| 8341 | Mobile farm and forestry plant operators | 0.332 | | |
| 1330 | ICT service managers | 0.314 | | |
| 1324 | Supply, distribution and related managers | 0.298 | | |
| 8160 | Food and related products machine operators | 0.273 | | |
| 8344 | Lifting truck operators | | 0.496 | |
| 9329 | Manufacturing labourers not elsewhere classified | | 0.481 | |
| 4321 | Stock clerks | | 0.420 | |
| 9520 | Street vendors (excluding food) | | 0.409 | |
| 8331 | Bus and tram drivers | | 0.405 | |

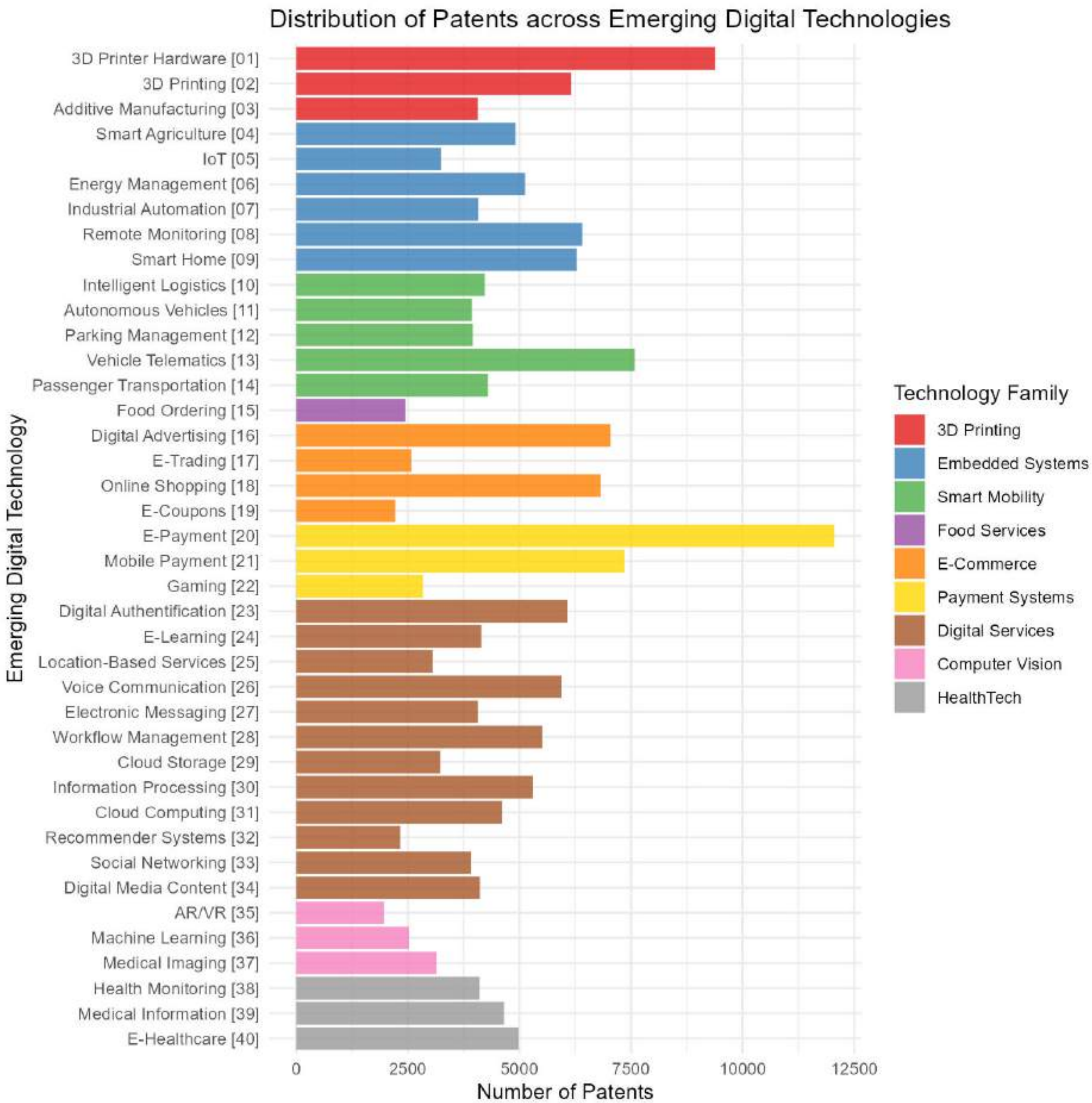
*Notes*: This table presents the redundancy filtering of occupations for the Patent ID 201713859U (i.e., "Vehicle intelligent logistics control device, has GPS locating module for obtaining position information of transport vehicle through main control chip, RFID reader for reading RFID tag information, and 4G module connected with server"). It displays the cosine similarity of the patent title with the 4-digit ISCO-08 title (Column 3) and the task with the highest cosine similarity (Column 4). Occupations are ranked according to Column 3 in decreasing order. Cosine similarity scores in Columns 3 and 4 are displayed only for sub-pairs belonging to their respective top 10. Column 5 shows the composite patent-occupation cosine similarity score, which corresponds to the harmonic mean of Columns 3 and 4. Cosine similarity scores in Column 5 are displayed only for pairs that rank simultaneously in both top 10.

Table A.8: Example of Redundancy Filtering of Occupations for Speech Recognition System

| Code | ISCO Occupation | Cosine Similarity | | |
|---|---|---|---|---|
| | | $C^p_{o_1}$ | $C^p_{o_2}$ | $C^p_o$ |
| 4131 | Typists and word processing operators | 0.309 | 0.452 | 0.367 |
| 2643 | Translators, interpreters and other linguists | 0.245 | 0.379 | 0.298 |
| 4413 | Coding, proofreading and related clerks | 0.232 | 0.343 | 0.277 |
| 2266 | Audiologists and speech therapists | 0.218 | 0.363 | 0.273 |
| 8153 | Sewing machine operators | 0.214 | | |
| 7532 | Garment and related patternmakers and cutters | 0.209 | | |
| 4223 | Telephone switchboard operators | 0.198 | | |
| 8143 | Paper products machine operators | 0.197 | | |
| 8131 | Chemical products plant and machine operators | 0.193 | | |
| 7422 | ICT installers and servicers | 0.193 | | |
| 4110 | General office clerks | | 0.396 | |
| 3252 | Medical records and health information technicians | | 0.339 | |
| 4120 | Secretaries (general) | | 0.329 | |
| 4132 | Data entry clerks | | 0.324 | |
| 4311 | Accounting and bookkeeping clerks | | 0.304 | |
| 2152 | Electronics engineers | | 0.302 | |

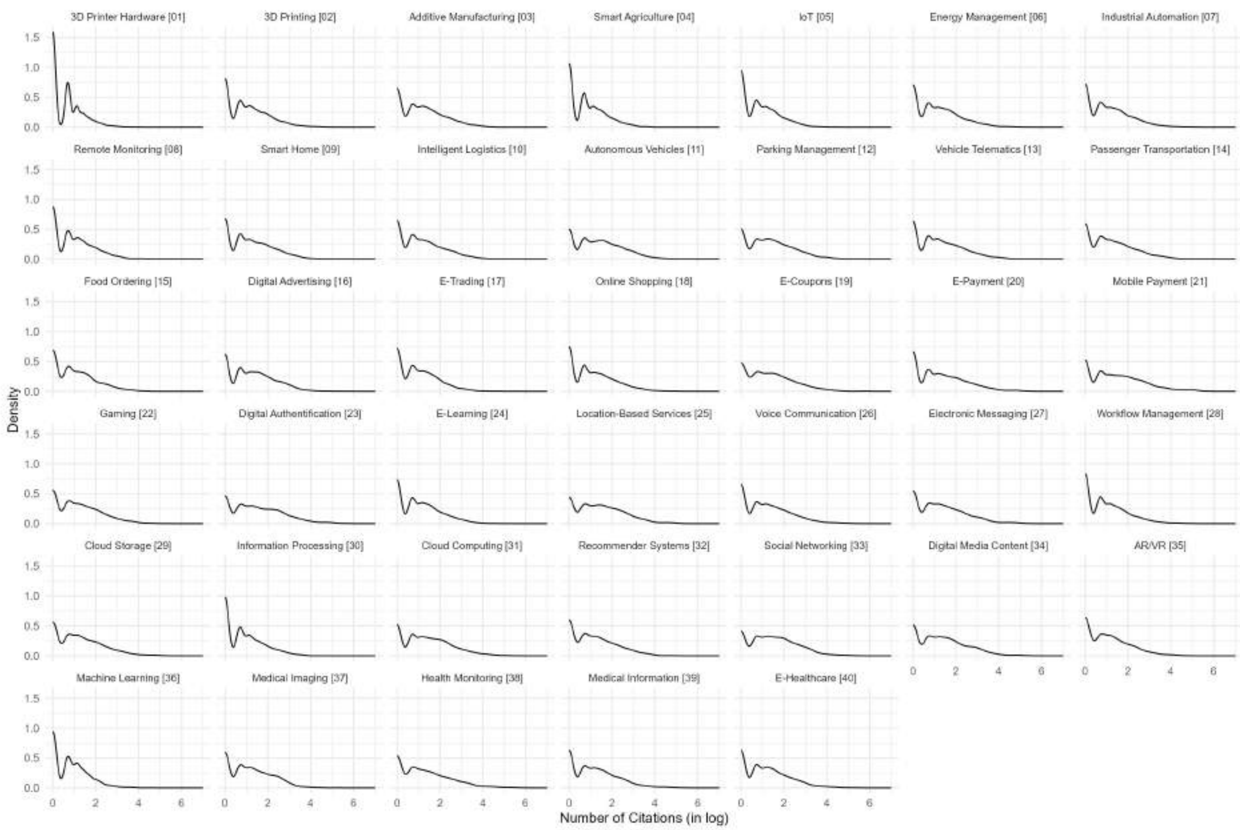
*Notes*: This table presents the redundancy filtering of occupations for the Patent ID 202048118D (i.e., "System for recognizing training speech, has process or which is configured to increment counter associated with word sequences, and train language model of automatic transcription system using word sequences and counter"). It displays the cosine similarity of the patent title with the 4-digit ISCO-08 title (Column 3) and the task with the highest cosine similarity (Column 4). Occupations are ranked according to Column 3 in decreasing order. Cosine similarity scores in Columns 3 and 4 are displayed only for sub-pairs belonging to their respective top 10. Column 5 shows the composite patent-occupation cosine similarity score, which corresponds to the harmonic mean of Columns 3 and 4. Cosine similarity scores in Column 5 are displayed only for pairs that rank simultaneously in both top 10.

Figure A.1: Distribution of Patents across Emerging Digital Technologies



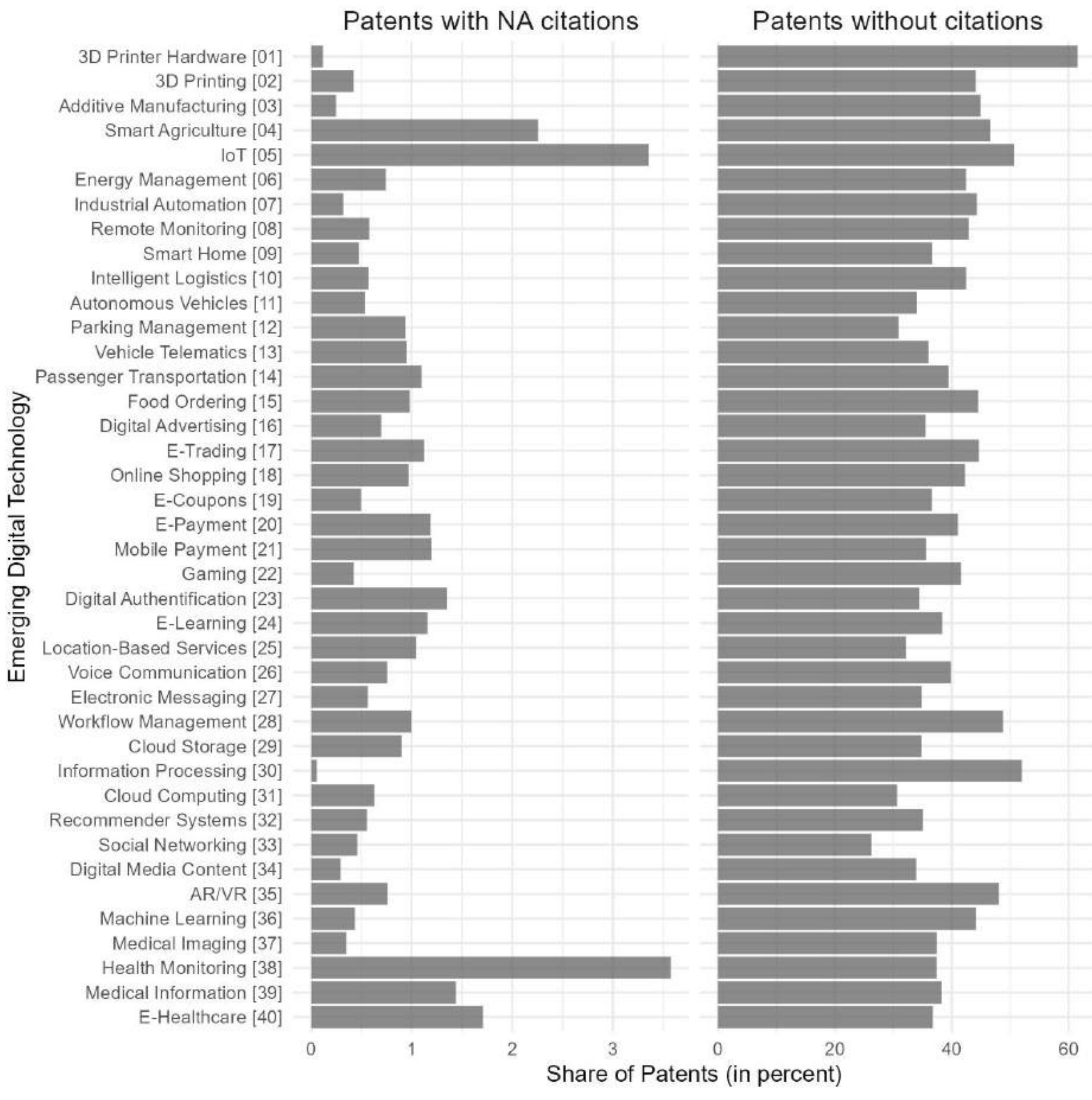

*Notes*: This figure presents the distribution of patents across emerging digital technologies. The set of patents includes 190,714 Derwent patents, filed between 2012 and 2021. This patent set constructed by Chaturvedi et al. (2023) comprises the most novel patents related to digital innovations, together with the patents that follow their semantic trajectory, that is, the most similar patents filed in subsequent years.

Figure A.2: Log Distribution of Patent Citations across Emerging Digital Technologies



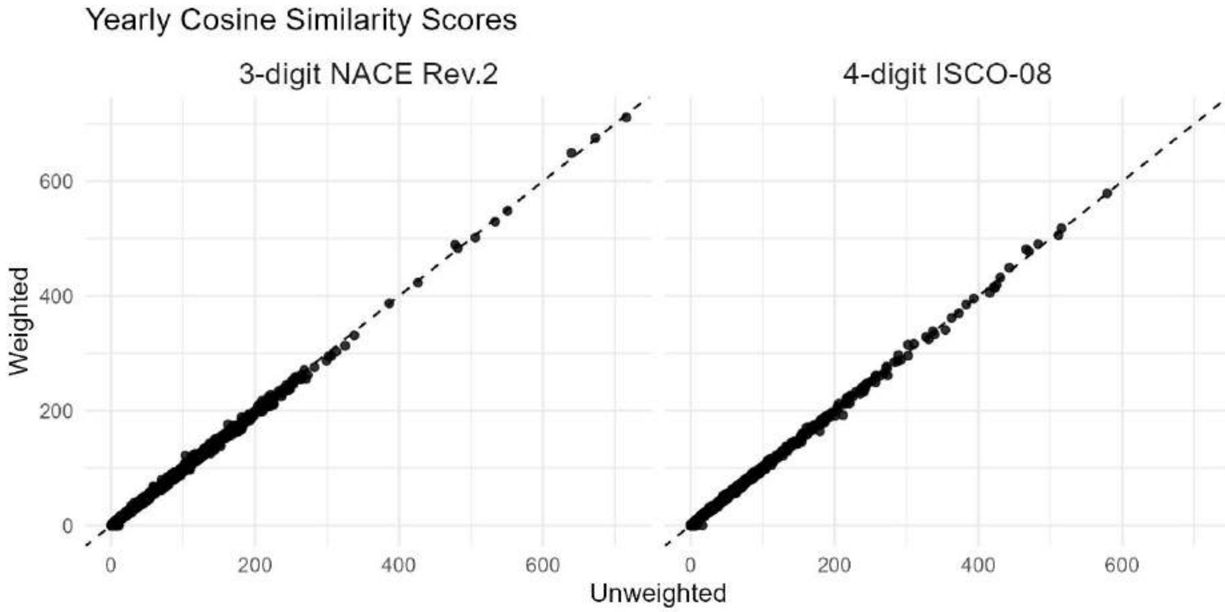

*Notes*: This figure presents the log distribution of patent citations across emerging digital technologies.

Figure A.3: Distribution of Non-Cited and Undetermined-Count Patents across Emerging Digital Technologies



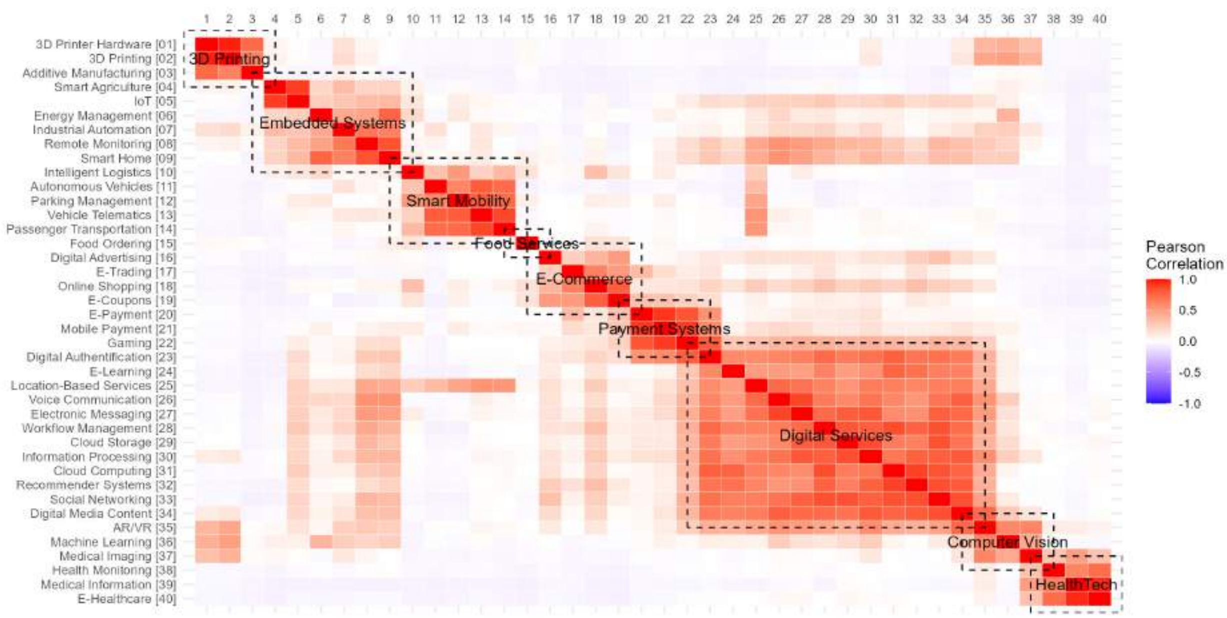

*Notes*: This figure presents the distribution of non-cited and undetermined-count patents across emerging digital technologies.

Figure A.4: Weighted versus Unweighted Yearly Cosine Similarity Scores




*Notes*: This figure presents the correlation between citation-weighted and unweighted yearly cosine similarity scores for both industries and occupations.

Figure A.5: Semantic Co-Occurrence of Technologies in 3–digit ISCO-08 Occupations




*Notes*: This figure shows all pairwise semantic-based technology co-occurrences as a correlation matrix, which is symmetric with diagonal values of 1. The matrix categorizes technologies into blocks, grouping them according to their semantic associations with occupations.

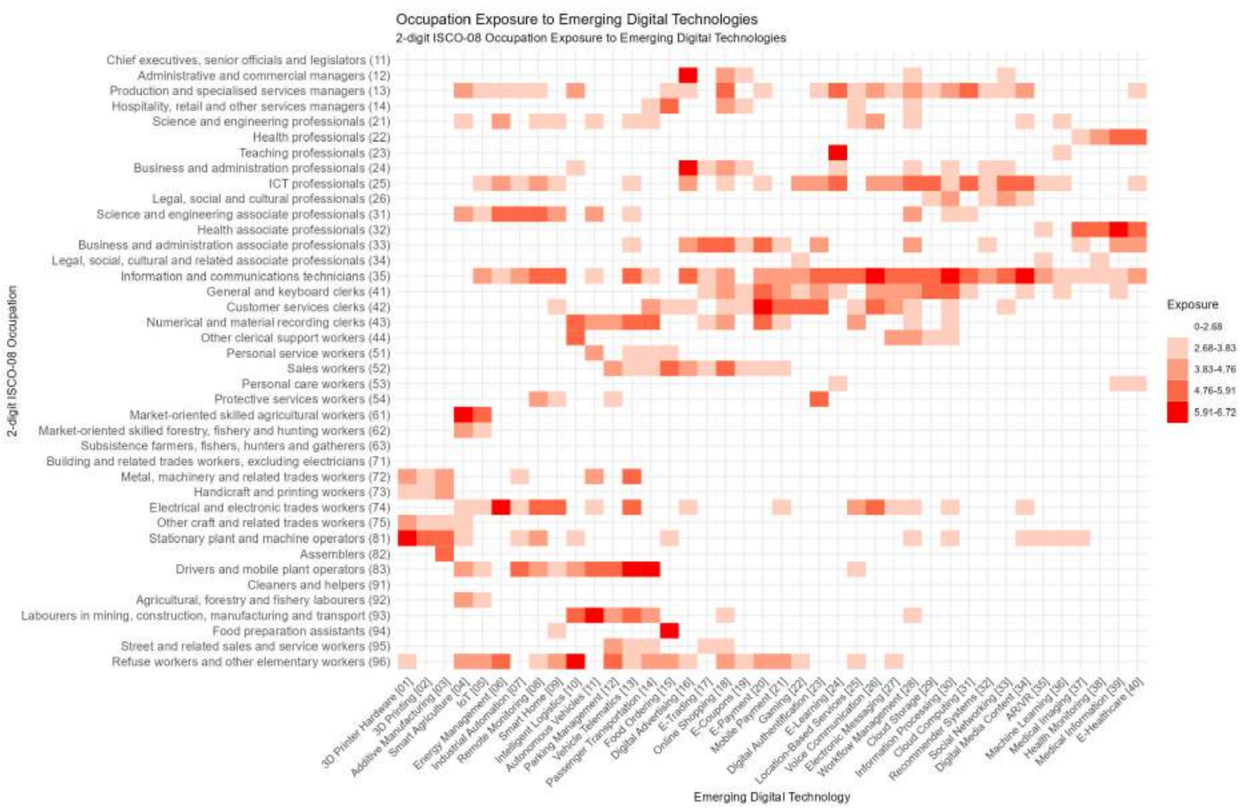
# B Descriptive Statistics Appendix

In this Appendix, we provide additional descriptive statistics on the exposure of industries and occupations to emerging digital technologies.

Tables B.1 and B.2 display the top 30 exposed 4-digit ISCO-08 occupations and 3-digit NACE Rev.2 industries, respectively, according to their average exposure to all emerging digital technologies, denoted as $X_o = \frac{1}{40}\sum_k X_o^k$, where $X_o^k$ is the exposure of occupation $o$ to technology $k$ given by Equation (13) and $X_i = \frac{1}{40}\sum_k X_i^k$, where $X_i^k$ is the exposure of industry $i$ to technology $k$ also given by Equation (13). Tables also include their top-3 concentration ratio (CR3) expressed in percent.

Figures B.1 and B.2 present the exposure of 2-digit ISCO-08 occupations and 2-digit NACE Rev.2 industries, respectively, to the 40 emerging digital technologies.

Figure B.1: Occupation Exposure by Emerging Digital Technologies (2-digit ISCO-08)




*Notes*: Each cell shows the exposure of a 2-digit ISCO-08 occupation (row) to a given emerging digital technology (column). Exposure scores below the 80th percentile (0-2.68) are transparent, whereas the four other groups represent respectively the 80th (2.68-3.83), 90th (3.83-4.76), 95th (4.76-5.91), and 99th (5.91-6.72) percentile of the distribution.

Table B.1: Top 30 Exposed 4-digit ISCO-08 Occupations

| Code | ISCO Occupation | $X_o$ | $CR3_o$ |
|---|---|---|---|
| 3513 | Computer network and systems technicians | 4.41 | 11.7 |
| 3511 | ICT operations technicians | 4.32 | 12.4 |
| 1330 | ICT service managers | 4.10 | 13.1 |
| 2523 | Computer network professionals | 3.98 | 12.7 |
| 3512 | ICT user support technicians | 3.86 | 12.4 |
| 8132 | Photographic products machine operators | 3.66 | 15.9 |
| 4223 | Telephone switchboard operators | 3.56 | 14.6 |
| 7422 | ICT installers and servicers | 3.36 | 14.3 |
| 3514 | Web technicians | 3.25 | 13.3 |
| 4132 | Data entry clerks | 3.11 | 15.6 |
| 9623 | Meter readers and vending-machine collectors | 3.09 | 16.9 |
| 3133 | Chemical processing plant controllers | 3.04 | 18.0 |
| 8322 | Car, taxi and van drivers | 2.68 | 21.8 |
| 2153 | Telecommunications engineers | 2.57 | 17.1 |
| 1324 | Supply, distribution and related managers | 2.55 | 19.8 |
| 9621 | Messengers, package deliverers and luggage porters | 2.49 | 19.7 |
| 2513 | Web and multimedia developers | 2.44 | 19.5 |
| 3311 | Securities and finance dealers and brokers | 2.44 | 22.7 |
| 2521 | Database designers and administrators | 2.43 | 17.8 |
| 3252 | Medical records and health information technicians | 2.38 | 25.3 |
| 8183 | Packing, bottling and labelling machine operators | 2.36 | 18.0 |
| 2622 | Librarians and related information professionals | 2.35 | 20.9 |
| 4323 | Transport clerks | 2.23 | 24.5 |
| 8312 | Railway brake, signal and switch operators | 2.20 | 21.0 |
| 5244 | Contact centre salespersons | 2.17 | 20.7 |
| 3522 | Telecommunications engineering technicians | 2.13 | 19.4 |
| 2529 | Database and network professionals n.e.c. | 2.13 | 20.7 |
| 3135 | Metal production process controllers | 2.03 | 20.2 |
| 3114 | Electronics engineering technicians | 1.98 | 19.7 |
| 2522 | Systems administrators | 1.96 | 17.6 |

*Notes*: This table presents the top 30 4-digit ISCO-08 occupations ranked by exposure to all emerging digital technologies. Columns (from left to right) correspond to occupation code, occupation title, average exposure to emerging digital technologies, top-3 concentration ratio which represents the sum of top-3 technology exposure shares (in percent).

Table B.2: Top 30 Exposed 3-digit NACE Rev.2 Industries

| Code | NACE Industry | $X_i$ | $\text{CR3}_i$ |
|---|---|---|---|
| 26.3 | Manufacture of communication equipment | 6.28 | 9.7 |
| 26.2 | Manufacture of computers and peripheral equipment | 6.19 | 9.5 |
| 63.1 | Data processing, hosting and related activities | 5.88 | 10.0 |
| 62.0 | Computer programming, consultancy and related activities | 5.28 | 10.6 |
| 26.5 | Manufacture of instruments and appliances for measuring | 4.88 | 11.8 |
| 82.9 | Business support service activities n.e.c. | 4.83 | 11.5 |
| 28.2 | Manufacture of other general-purpose machinery | 4.71 | 12.7 |
| 63.9 | Other information service activities | 4.70 | 11.8 |
| 61.2 | Wireless telecommunications activities | 4.67 | 11.7 |
| 61.9 | Other telecommunications activities | 4.43 | 12.2 |
| 33.1 | Repair of fabricated metal products, machinery and equipment | 4.22 | 12.1 |
| 95.1 | Repair of computers and communication equipment | 4.11 | 12.2 |
| 79.9 | Other reservation service and related activities | 3.96 | 13.4 |
| 80.2 | Security systems service activities | 3.83 | 14.2 |
| 52.2 | Support activities for transportation | 3.59 | 16.0 |
| 27.9 | Manufacture of other electrical equipment | 3.50 | 15.1 |
| 61.1 | Wired telecommunications activities | 3.50 | 13.8 |
| 47.4 | Retail sale of information and communication equipment | 3.35 | 15.2 |
| 26.4 | Manufacture of consumer electronics | 3.30 | 13.1 |
| 28.9 | Manufacture of other special-purpose machinery | 3.26 | 18.3 |
| 27.1 | Manufacture of electric motors, generators and transformers | 2.95 | 18.3 |
| 82.2 | Activities of call centres | 2.81 | 16.2 |
| 80.1 | Private security activities | 2.78 | 16.6 |
| 26.1 | Manufacture of electronic components and boards | 2.76 | 16.8 |
| 17.2 | Manufacture of articles of paper and paperboard | 2.73 | 14.8 |
| 58.1 | Publishing of books, periodicals and other publishing activities | 2.69 | 17.4 |
| 27.3 | Manufacture of wiring and wiring devices | 2.38 | 17.7 |
| 18.2 | Reproduction of recorded media | 2.31 | 20.0 |
| 33.2 | Installation of industrial machinery and equipment | 2.30 | 20.6 |
| 82.1 | Office administrative and support activities | 2.14 | 18.8 |

*Notes*: This table presents the top 30 3-digit NACE Rev.2 industries ranked by exposure to all emerging digital technologies. Columns (from left to right) correspond to industry code, industry title, average exposure to emerging digital technologies, top-3 concentration ratio which represents the sum of top-3 technology exposure shares (in percent).

Figure B.2: Industry Exposure by Emerging Digital Technologies (2-digit NACE Rev.2)


*Notes*: Each cell shows the exposure of a 2-digit NACE Rev.2 industry (row) to a given emerging digital technology (column). Exposure scores below the 80th percentile (0-2.92) are transparent, whereas the four other groups represent respectively the 80th (2.92-4.50), 90th (4.50-5.47), 95th (5.47-6.79), and 99th (6.79-8.01) percentile of the distribution.

# C Employment Impact Appendix

In this Appendix, we provide additional information on the regional employment analysis performed in Section 5. Appendix C.1 presents additional statistics on the employment shares used in the regional exposure constructed as a shift-share variable. Appendix C.2 presents the geographic distribution of exposure to individual emerging digital technologies.

## C.1 Employment Shares

Table C.1 presents the employment shares of our 10 sectors of activities averaged across all the European regions in 2010. The three largest sectors in Europe are the Public Sector (O-Q), accounting for an average of 25.7% of employment, Market Services (G-I), with an average share of 24%, and Industry (B-E), representing 17.1% on average. Subsequently, there is a group of sectors each contributing between 7% and 8% on average to employment, comprising Agriculture (A), Construction (F), and Professional, Scientific, Technical, Administration, and Support Service Activities (M-N). The remaining four sectors collectively account for 10.5% of employment. Notably, the Information and Communication sector (J), pivotal to emerging digital technologies, comprises only 2.3% of average regional employment in Europe. This figure is comparable to the Financial and Insurance Activities sector (K), which averages 2.6%.

To provide evidence on the exogeneity of shares conditional on observables, we estimate the effect of shares on the change in the outcome variable at the regional level using OLS, conditional on observables. We weight our observations by the population in 2010. The estimated regression is:

$$\Delta Y_r = \alpha + \sum_j \beta_j l_{rj} + Z\gamma + \phi_{c(r)} + u_r,$$

where $\Delta Y_r$ is the change in the employment-to-population ratio for region $r$ between 2012 and 2019 in percentage points, $l_{rj}$ is the share of employment in region $r$ in sector $j$ in 2010, $Z$ is the set of covariates which capture regional characteristics, $\phi_{c(r)}$ are the country fixed effects, and $u_r$ is the error term.

Figure C.1 summarizes the estimated coefficients of interest $\widehat{\beta_j}$ along with their 95% confidence interval. Each row corresponds to the correlation between the employment share in sector $j$ and the change in regional employment-to-population ratio (in pp.). The first specification only accounts for country fixed effects whereas the second specification also includes regional demographics. We find no correlation between shares and changes in the outcome variable, conditional on observables.

## C.2 Geographic Distribution of Exposure

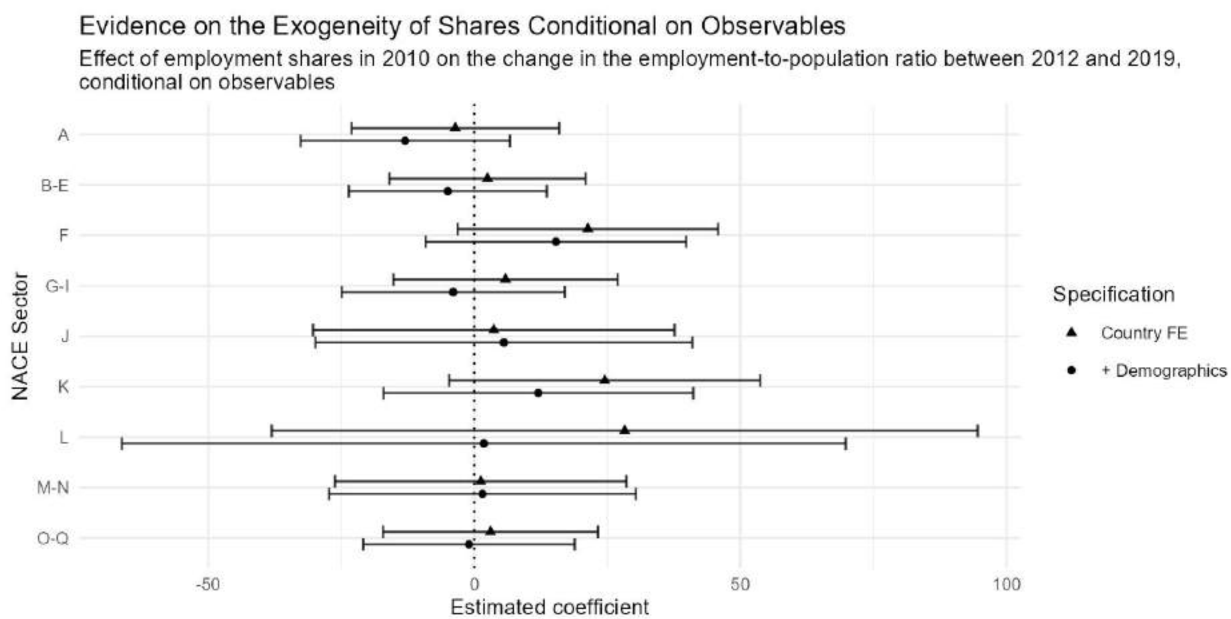
Figures C.2 to C.6 present the geographic distribution of exposure to individual emerging digital technologies, constructed as a shift-share. Regional exposure scores are standardized to allow comparability between technologies.

Table C.1: Average Employment Share by Sector of Activities in 2010

| | NACE Sector | Mean | SD |
|---|---|---|---|
| A | Agriculture | 0.071 | 0.099 |
| B-E | Industry, excluding Construction | 0.171 | 0.074 |
| F | Construction | 0.078 | 0.019 |
| G-I | Market Services, excluding Information and Communication | 0.240 | 0.040 |
| J | Information and Communication | 0.023 | 0.017 |
| K | Financial and Insurance Activities | 0.026 | 0.018 |
| L | Real Estate Activities | 0.005 | 0.005 |
| M-N | Professional, Scientific, Technical, Administration and Support Service Activities | 0.079 | 0.032 |
| O-Q | Public Administration, Defence, Education, Human Health and Social Work Activities | 0.256 | 0.075 |
| R-U | Other Services | 0.051 | 0.018 |

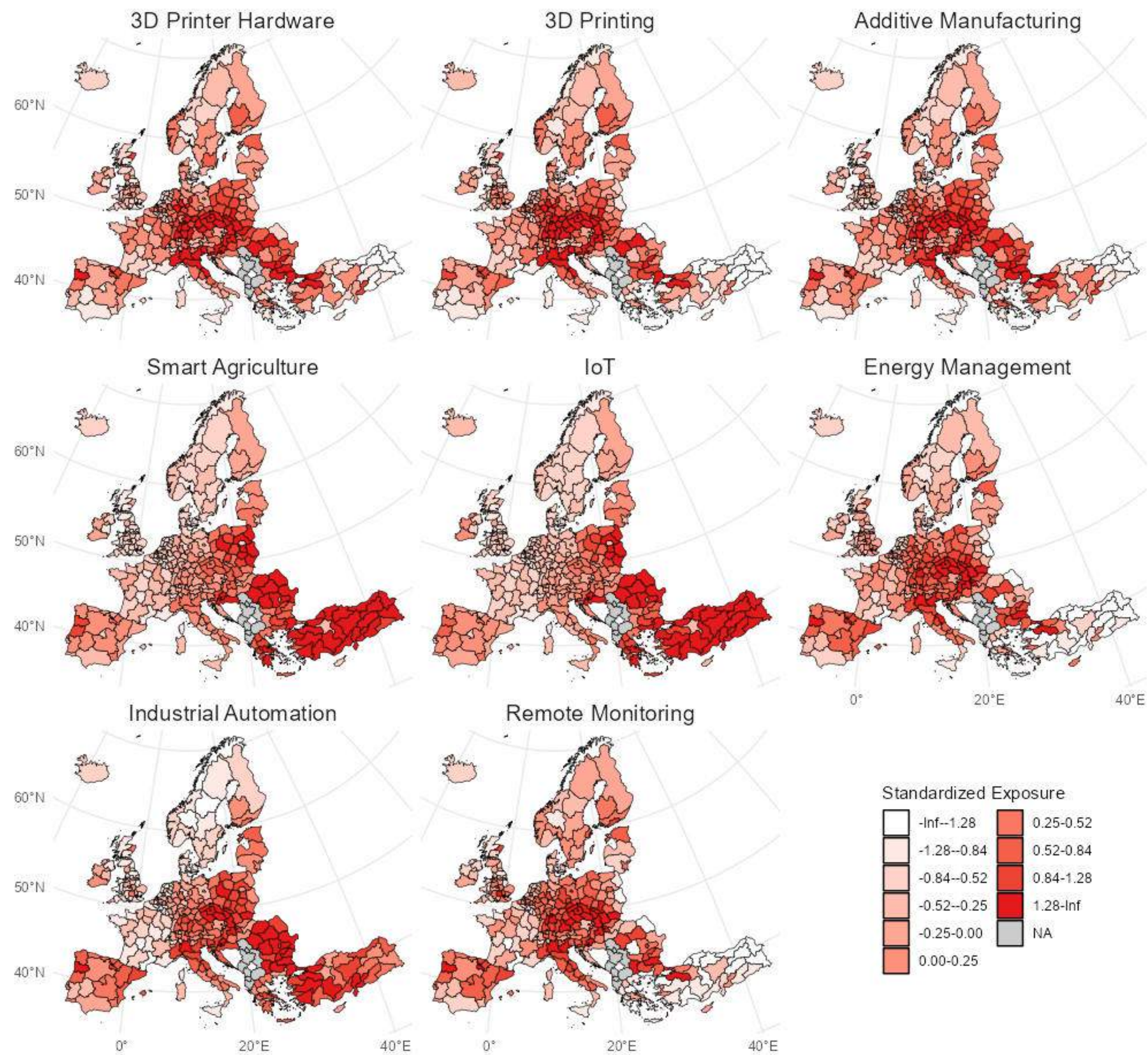
*Notes*: This table presents the employment share by sector of activities averaged across all the European regions in 2010. The first column indicates the 1-digit NACE codes, the second column is the name of the NACE sector, the third column is the average employment share in 2010, and the fourth column gives the standard errors.

Figure C.1: Evidence on the Exogeneity of Shares Conditional on Observables



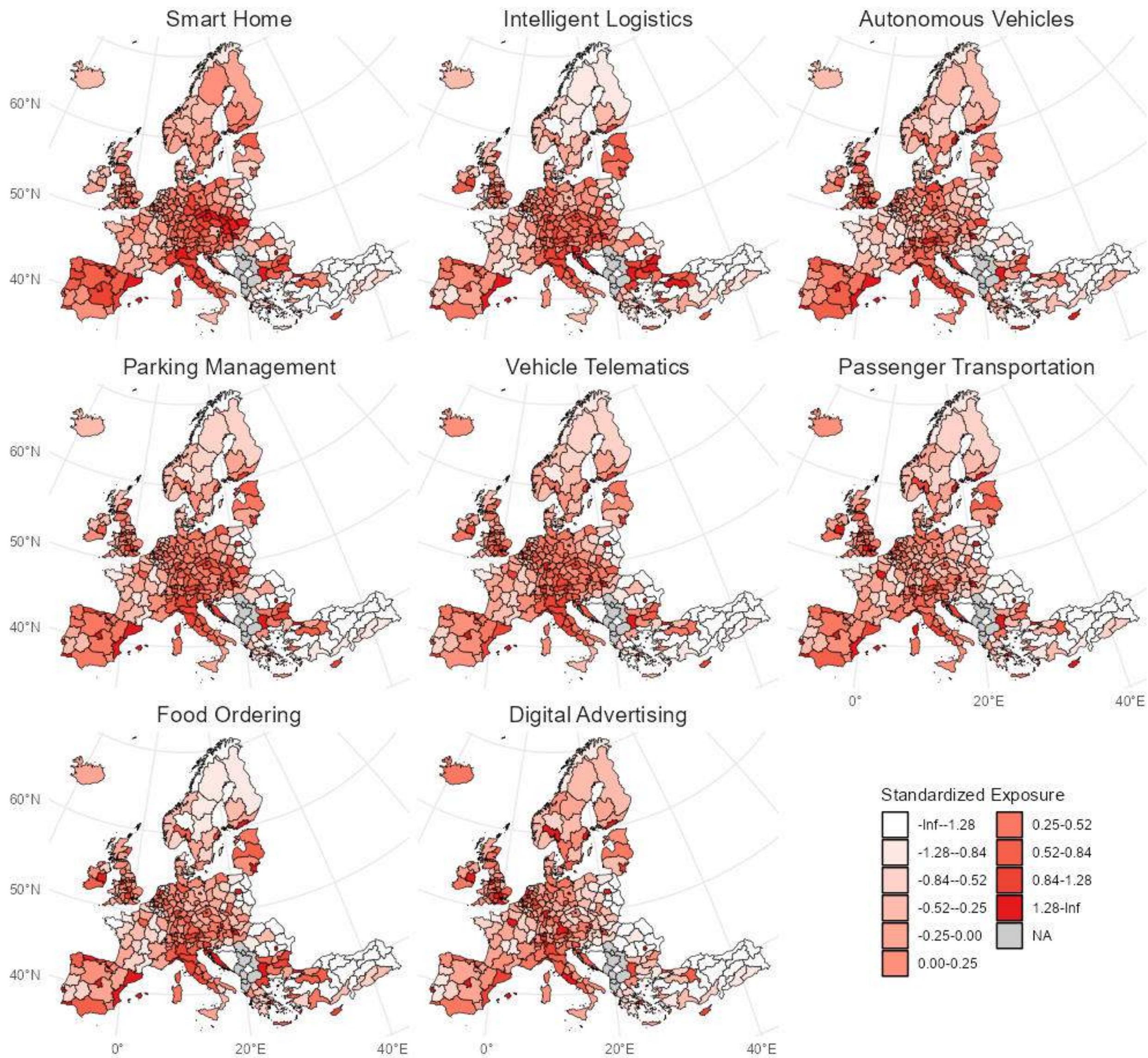

*Notes*: This figure presents the estimated coefficients of the effect of employment shares in 2010 on the change in the employment-to-population ratio between 2012 and 2019 (in pp.), conditional on observables. The estimates are reported with their confidence interval at the 95% level. The first specification includes only country fixed effects. The second specification adds demographics controls in 2010, including the logarithm of population, the proportion of females, the proportion of the population aged over 65, and the proportions of the population with secondary and tertiary education levels.

Figure C.2: Geographic Distribution of Exposure to Emerging Digital Technologies (1/5)



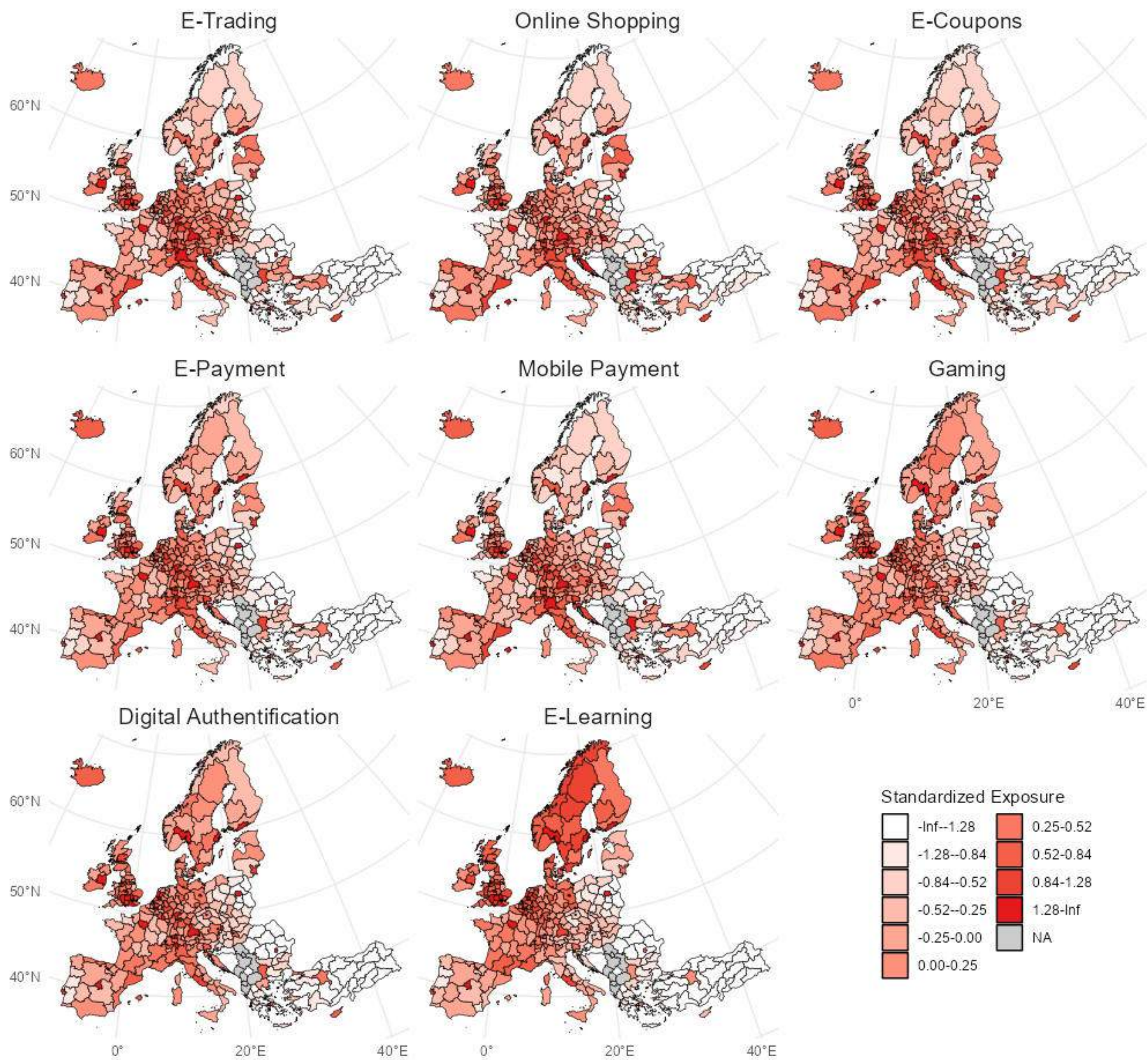

*Notes*: This figure illustrates the geographic distribution of exposure to emerging digital technologies for NUTS-2 regions. Regional exposure is constructed as a shift-share variable by interacting the sectoral employment shares in the baseline year 2010 and sectoral exposure to these technologies from 2012 to 2019. Regions are categorized into deciles. Regions are shaded according to their exposure level, with the legend indicating the range of exposure. Areas not applicable (NA) are marked in grey.

Figure C.3: Geographic Distribution of Exposure to Emerging Digital Technologies (2/5)



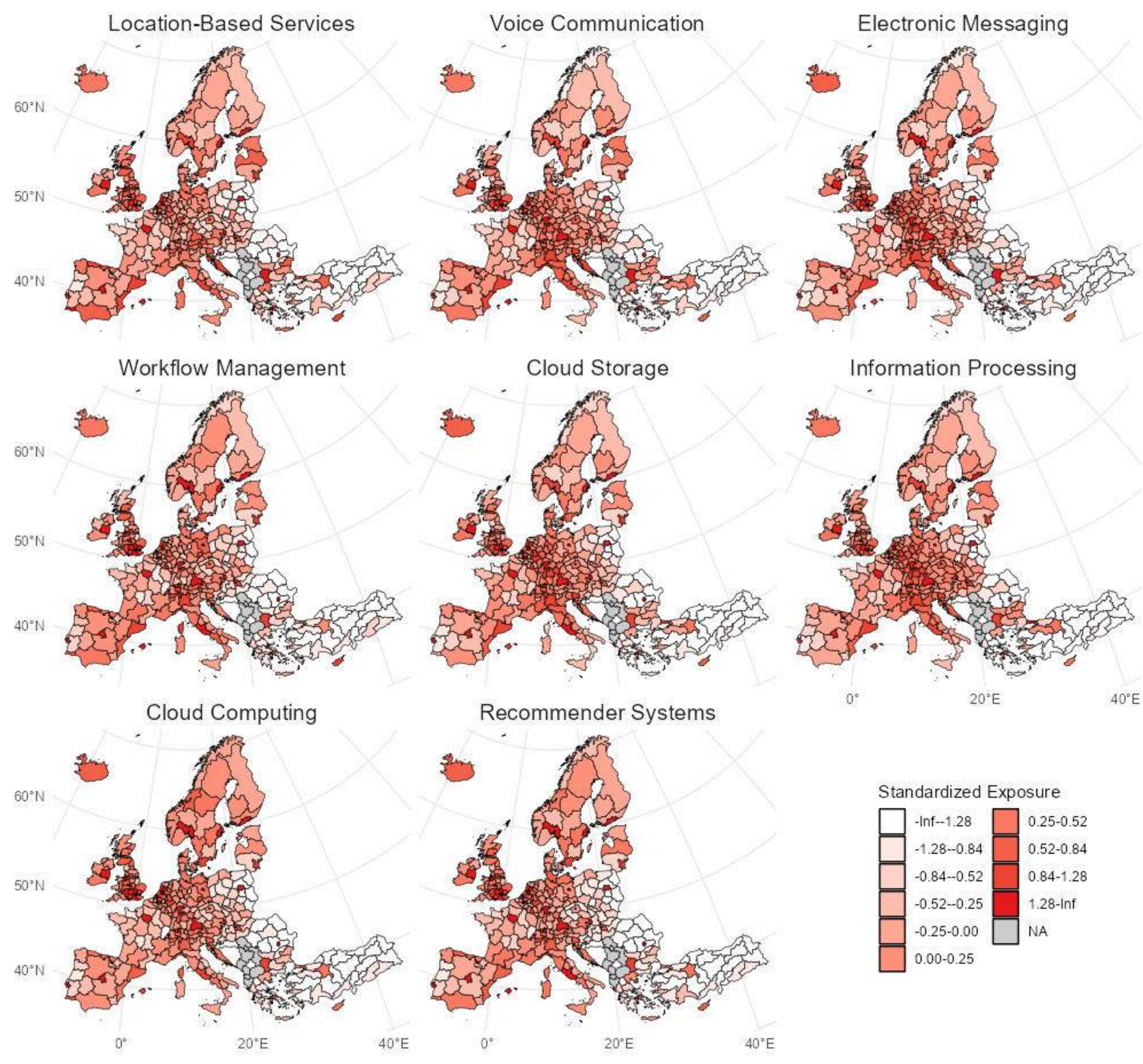

*Notes*: This figure illustrates the geographic distribution of exposure to emerging digital technologies for NUTS-2 regions. Regional exposure is constructed as a shift-share variable by interacting the sectoral employment shares in the baseline year 2010 and sectoral exposure to these technologies from 2012 to 2019. Regions are categorized into deciles. Regions are shaded according to their exposure level, with the legend indicating the range of exposure. Areas not applicable (NA) are marked in grey.

Figure C.4: Geographic Distribution of Exposure to Emerging Digital Technologies (3/5)



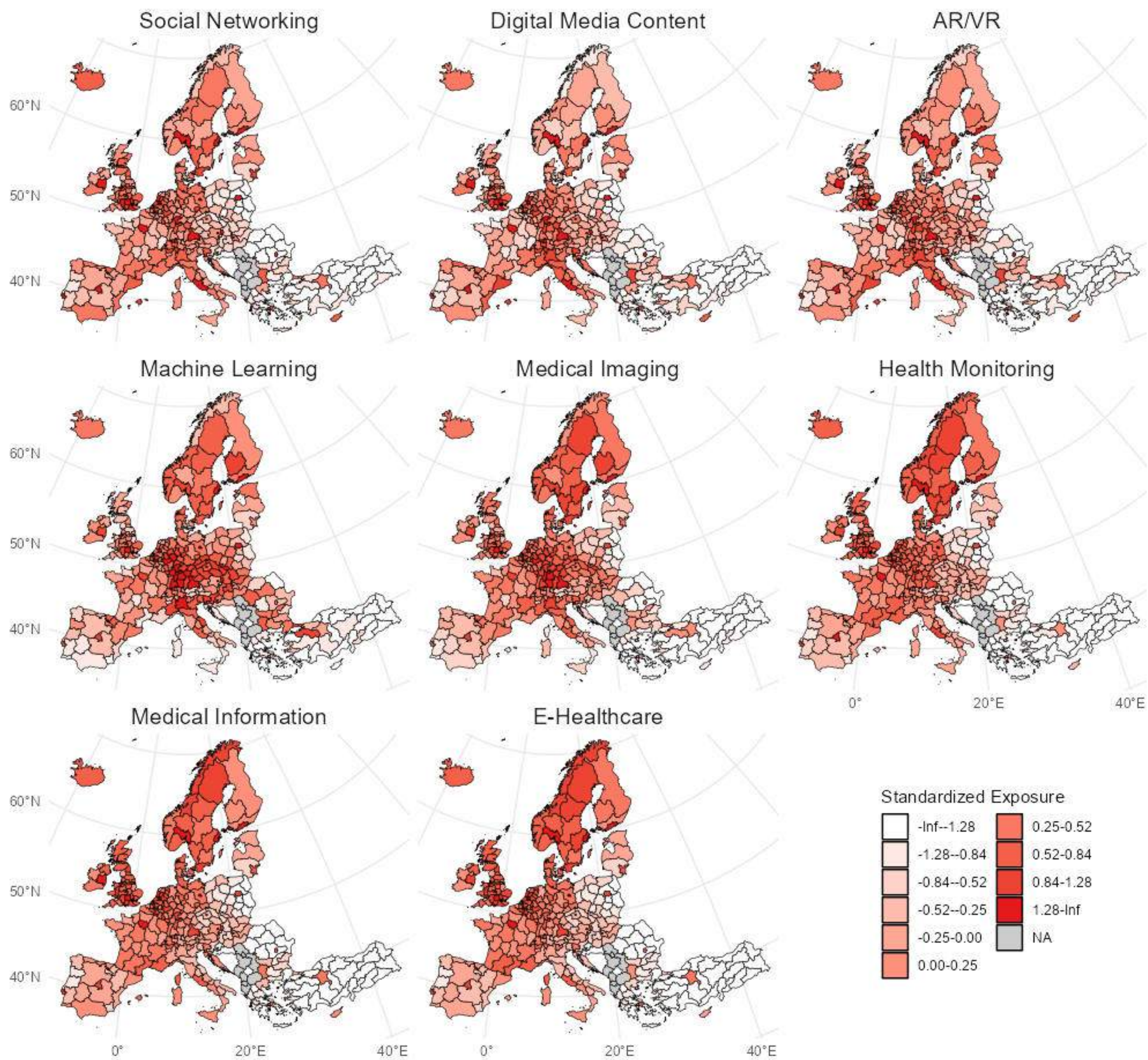

*Notes*: This figure illustrates the geographic distribution of exposure to emerging digital technologies for NUTS-2 regions. Regional exposure is constructed as a shift-share variable by interacting the sectoral employment shares in the baseline year 2010 and sectoral exposure to these technologies from 2012 to 2019. Regions are categorized into deciles. Regions are shaded according to their exposure level, with the legend indicating the range of exposure. Areas not applicable (NA) are marked in grey.

Figure C.5: Geographic Distribution of Exposure to Emerging Digital Technologies (4/5)




*Notes*: This figure illustrates the geographic distribution of exposure to emerging digital technologies for NUTS-2 regions. Regional exposure is constructed as a shift-share variable by interacting the sectoral employment shares in the baseline year 2010 and sectoral exposure to these technologies from 2012 to 2019. Regions are categorized into deciles. Regions are shaded according to their exposure level, with the legend indicating the range of exposure. Areas not applicable (NA) are marked in grey.

Figure C.6: Geographic Distribution of Exposure to Emerging Digital Technologies (5/5)




*Notes*: This figure illustrates the geographic distribution of exposure to emerging digital technologies for NUTS-2 regions. Regional exposure is constructed as a shift-share variable by interacting the sectoral employment shares in the baseline year 2010 and sectoral exposure to these technologies from 2012 to 2019. Regions are categorized into deciles. Regions are shaded according to their exposure level, with the legend indicating the range of exposure. Areas not applicable (NA) are marked in grey.

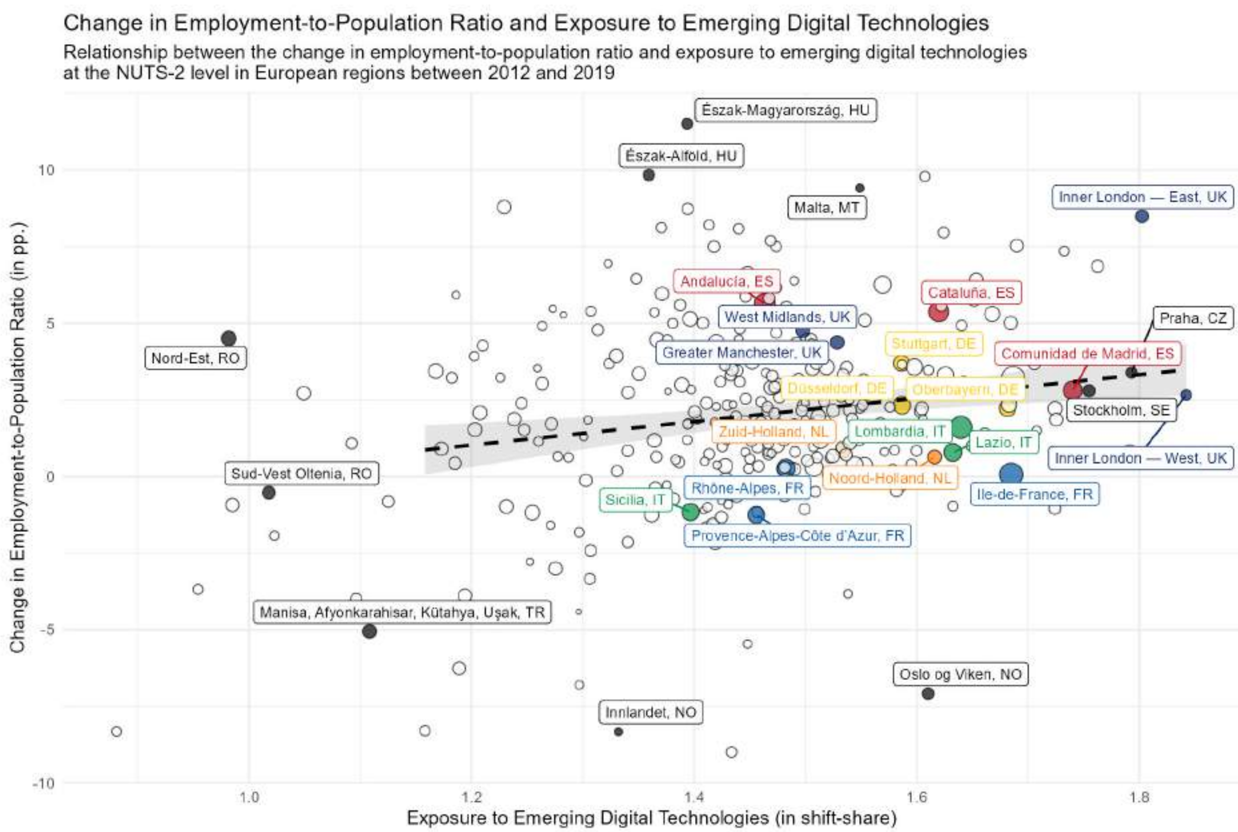
# Online Appendix

# The Employment Impact of Emerging Digital Technologies

Ekaterina Prytkova, Fabien Petit, Deyu Li, Sugat Chaturvedi, Tommaso Ciarli

Figure OA.1 depicts a positive relationship between the change in the employment-to-population ratio from 2012 to 2019 and the regional exposure to emerging digital technologies, after excluding regions with exceptionally low exposure levels — specifically, those with an exposure index below -2 standard deviations (i.e. below 0.929), which typically includes rural areas in Romania, Turkey, and overseas French territories.

Figure OA.1: Change in Employment-to-Population Ratio and Exposure to Emerging Digital Technologies




*Notes*: This figure shows the relationship between the change in the employment-to-population ratio and the exposure to emerging technologies in European NUTS-2 regions between 2012 and 2019. Each point represents a region, with select regions labeled for emphasis. The size of the point is proportional to the population in 2010. The horizontal axis measures the exposure to emerging technologies calculated by the shift-share method, while the vertical axis represents the change in the employment-to-population ratio. The solid line indicates a positive correlation between increased regional exposure to emerging technologies and employment growth. Regressions lines are weighted by population in 2010. Data points are color-coded by country. Outliers are highlighted and labeled for clarity.