The Employment Impact of Emerging Digital Technologies: Evidence from US Labor Markets*

Ekaterina Prytkova† Fabien Petit‡

February 20, 2025

Abstract

This paper estimates the exposure of US occupations and industries to emerging digital technologies and their impact on US commuting zone (CZ) employment. Building upon the natural language processing approach introduced by Prytkova et al. (2024), we estimate the exposure of O*NET-SOC occupations and NAICS industries, thereby extending the open–access 'TechXposure' database to the US context. Using this new data source, we apply a shift-share design to instrument the CZ exposure to emerging digital technologies and estimate their employment impact across CZs between 2012 and 2019. We find that digital technologies have an overall positive net impact on US employment. However, the impact varies among different worker demographics: while there is a noticeable decline in employment for core working-age (25–44) and non-college-educated workers in more exposed CZs, we observe employment increases for younger (16–24) and older (45–64) workers, as well as for those with a college education.

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 Sugat Chaturvedi, Tommaso Ciarli, Cecilia García Peñalosa, Susan Houseman, Deyu Li, Anne Polivka, Morgan Raux, Aysegul Sahin, as well as to audiences at the CRIW Conference on The Changing Nature of Work for useful comments and suggestions.

†University Côte d'Azur. Email: ekaterina.prytkova@univ-cotedazur.fr.

‡University of Barcelona, University College London, CEPEO. Email: f.petit@ub.edu. Website: fabien-petit.com.

1 Introduction

Over the last decades, emerging digital technologies—including artificial intelligence (AI), electric vehicles, drones, the Internet of Things (IoT), and robotics—have transformed the global economy through widespread digitalization. This profound shift is reshaping the nature of work, impacting workers in uneven ways. While some workers benefit from enhanced productivity and new job opportunities, others face job displacement as their tasks become automated (Acemoglu and Restrepo 2018). Understanding which occupations and industries are most susceptible to these changes, and which demographic and skill groups are most vulnerable is crucial for policymakers to mitigate adverse effects and maximize the benefits of the ongoing technological change.

In this paper, we estimate the exposure of US industries and occupations to a diverse array of digital technologies introduced between 2012 and 2021, and evaluate their impact on employment across commuting zones (CZ). Using the natural language processing (NLP) methodology developed by Prytkova et al. (2024), we calculate semantic similarities between patents and the descriptions of industries and occupations from US standard classifications. This yields exposure scores that quantify the *relevance* of these 40 digital technologies to industries and occupations, extending the TechXposure database to the US context. We then combine this novel data source with US labor market data to estimate the overall effect of digital technologies on US employment and among different demographic and skill groups.

We start our analysis with a sample of patents identified as core digital inventions that emerged between 2012 and 2021 in Chaturvedi et al. (2023). These patents are grouped into 40 distinct technologies based on semantic similarities in their titles, following the methodology described by Prytkova et al. (2024). This set of technologies captures a diverse array of digital technologies beyond just AI and robotics, enabling a more precise and interpretable categorization.[1]

We preprocess the textual data from patents, industries, and occupations to impose a similar structure that brings out function-related content. Specifically, we utilize descriptions from 4-digit NAICS 2007 industries and 8-digit O*NET-SOC 2010 occupations. Then, we transform these preprocessed texts into contextual embeddings (i.e., dense vector representations) using the pre-trained MPNet v2 sentence transformer (Song et al. 2020). This method of encoding provides a superior capture of semantic meanings compared to traditional bag-of-word

[1]Specifically, the set includes technologies such as 3D printing, embedded systems like the Internet of Things (IoT) and remote monitoring, smart mobility solutions including intelligent logistics and electric vehicles, various payment systems including mobile payments and e-commerce, digital services such as cloud computing, workflow management systems, e-learning, and computer vision technologies like neural networks, along with digital health technologies. Detailed descriptions of these technologies are available in Tables 8 to 10 in the appendix.

approaches, which are commonly used in the literature (see Felten et al. 2018, Kogan et al. 2017, Kogan et al. 2019, Webb 2019, Felten et al. 2021, Kelly et al. 2021, Autor et al. 2024).[2] We compute the cosine similarity between each patent–industry and patent–occupation pair of embeddings, while filtering out any irrelevant or mistaken connections (false positives). Finally, we aggregate these exposure scores from the patent level up to the technology level, weighting each patent by the number of citations it received relative to the total number of citations all patents filed in the same year received, and adjusting for the right skewness of the exposure scores.

Our methodology generates estimates of exposure to 40 digital technologies introduced between 2012 and 2021 for each of the 1,110 8-digit O*NET-SOC 2010 occupations and 324 4-digit NAICS 2007 industries. These exposure scores gauge the *relevance* of a technology to an industry or occupation, particularly through its applicability to production processes or tasks. The identification of a link between a technology and an industry or occupation suggests potential for adoption, though actual adoption may not yet have occurred. Thus, our exposure score serves as a proxy for *actual adoption*.

We present descriptive results regarding exposure scores. For occupations, the highest exposures are observed in Computer and Mathematics, Office and Administrative Support (including clerks and salespersons), and Management occupations. We elucidate the reasons behind their higher exposure by analyzing the most impacted O*NET tasks and identifying the digital technologies most relevant to them. Consistently, the exposure scores of industries align with those of occupations. Notably, the Information and Cultural Industries; Professional, Scientific, and Technical Services; and Administrative and Support Services are the most exposed 2-digit industries. The first two sectors reflect the significant exposure of IT professions, while the latter corresponds to the exposure observed in clerical and sales positions.

Leveraging exposure scores, we estimate the net impact of digital technologies on US employment across various demographic and skill groups using a long-difference shift-share approach at the CZ level. We focus on the period from 2012 to 2019, intentionally concluding before the onset of the COVID-19 pandemic to prevent distortions in our estimates caused by the pandemic's economic effects. Accordingly, we recalibrate our exposure scores for this timeframe.

Our shift-share approach constructs a measure of the digital technology exposure at the CZ level by combining the share of 2-digit NAICS industries in total employment within each CZ in 2010, alongside the averaged exposure score of 2-digit NAICS industries to digital tech-

[2] Contextual embeddings consider the surrounding context of a word, allowing for distinct vector representations in different texts—a feature not present in bag-of-word methods where each word is consistently represented by a single vector.

nologies (i.e., the shock). We posit that our exposure scores are exogenous to local US labor market dynamics, as they predominantly reflect global innovations not confined to the US—evidenced by only 36% of our analyzed patents originating from the US. To underscore this point, we recalculated the shift-share CZ exposure omitting US-origin patents, resulting in a correlation of approximately 0.99 with the original measure, indicating that these innovations are globally pervasive and unaffected by US-specific local labor market dynamics. Our methodology aligns with the shift-share approach suggested by Borusyak et al. (2021), which places the exogeneity condition on the shock rather than the share. We present results using AKM0 standard errors, following the recommendations of Adão et al. (2019).

The geographical distribution of exposure to digital technologies reveals significant disparities across CZs. Coastal and highly urbanized areas exhibit the highest levels of exposure, exemplified by cities such as New York City and Boston in the Northeast, and San Francisco and Los Angeles on the West Coast. Moderate exposure levels are observed in the Mid-Atlantic, Southeastern, and Midwest regions, represented by cities like Atlanta, Charlotte, Chicago, Minneapolis, and Detroit. These areas balance traditional manufacturing bases with emerging digital industries. In contrast, the Central and Mountain West regions experience the lowest exposure, attributed to their rural characteristics and lower population densities.

Our paper presents three main results. First, we identify a positive impact of digital technologies on the employment-to-population ratio; specifically, a one-standard-deviation increase in CZ exposure results in a 0.67 percentage point (1.1%) increase in this ratio, which represents our most conservative estimate. This positive effect remains robust when controlling for demographic characteristics of the CZ and the inclusion of industry share in the baseline analysis.

Second, despite the overall positive effect, we find substantial heterogeneity among demographic groups. Specifically, we observe positive employment effects for young workers (ages 16–24) and senior workers (ages 45–64), while identifying a negative impact for core working-age individuals (ages 25–44). This suggests that workers in their core working years are most vulnerable to technological changes, as the obsolescence of existing skills or the automation of their tasks may diminish their employment opportunities. Moreover, our analysis shows no significant differences in impact across genders.

Third, we also find evidence of skill-biased technological change attributable to digital technologies. Examining the variation by educational attainment, we observe a negative impact on workers with a high school education or less, whereas a positive impact is noted for college graduates. This outcome implies that lower levels of educational attainment may hinder workers' ability to adapt to technological advancements, consequently leading to greater job displacement.

This work contributes to the literature on the changing nature of work due to technological change. This research advances the understanding of the changing nature of work due to technological change. A primary contribution of our work is the development of a new open-access database that quantifies the exposure of industries and occupations to 40 emerging digital technologies within the US context. These exposure scores are provided at a granular level, specifically for 4-digit NAICS 2007 industries and 8-digit O*NET-SOC 2010 occupations, enabling detailed analyses when combined with US labor market data. Previous research has largely concentrated on computerization (Frey and Osborne 2017), industrial robots (Acemoglu and Restrepo 2020), specific applications of AI (Webb 2019, Felten et al. 2021, Felten et al. 2018), or a broad spectrum of technologies (Autor et al. 2024, Kogan et al. 2017, Kogan et al. 2017). Our study expands on this existing literature by including a wider array of new and detailed digital technologies, and by providing precise measures of exposure using a state-of-the-art NLP approach.

Secondly, we contribute to the literature on the impact of digital technologies on employment, a topic with varied findings in the literature depending on the technology considered. Acemoglu and Restrepo (2020), focusing on industrial robots, along with Webb (2019) and Bonfiglioli et al. (2024), who examine AI, report negative impacts on US employment. In contrast, Mann and Püttmann (2023) and Autor et al. (2024), who apply broader definitions of automation technologies, observe overall positive employment effects. Our research, encompassing a broad spectrum of digital technologies beyond merely robots or AI, similarly identifies a positive overall impact on employment. Additionally, we note significant heterogeneity across different age groups and educational levels, indicating the occurrence of skill-biased technological change consistent with Kogan et al. (2021).

The remainder of the paper is structured as follows. Section 2 outlines the methodology of Prytkova et al. (2024) and addresses the particularities of the US context, followed by presenting estimates of industry and occupation exposure. Section 3 quantifies the employment effects of digital technologies across US CZs. Section 4 concludes.

2 Exposure

We replicate the approach developed in Prytkova et al. (2024) to derive the exposure scores for the US standard classifications of occupations and industries.[3]

We link technologies to industries and occupations primarily by matching technologies' functions to industrial processes and occupational tasks respectively using text as data. The

[3]The original approach was performed on international standard classifications, namely, NACE Rev.2 for industries and ISCO-08 for occupations.

underlying idea is to find similarity between useful functions of a technology and functions performed by capital and labor during production of products and services.

2.1 Data

Digital Technologies. The set of 40 prominent digital technologies is constructed in Prytkova et al. (2024). Each digital technology is represented by a semantically-coherent cluster of patent families. The original patent sample retrieved from the Derwent Innovation Index (DII) database comprises documents filed between 2012 and 2021 pertaining to digital automation of *i)* process and machine control in physical production sectors like manufacturing, agriculture, mining, and construction, and *ii)* process and workflow control in services.

In this paper, we estimate the employment impact of the same set of digital technologies in the United States. Table 1 lists all 40 digital technologies which are further grouped into 9 technology families.

This set of digital technologies covers the entire spectrum from purely intangible, process(ing) technologies such as Workflow Management, E-Learning, or Medical Imaging & Image Processing, through hybrid or embedded technologies that comprise both tangible and intangible components like Internet of Things or Intelligent Logistics, to technologies with a prominent tangible component, for instance, 3D Printer Hardware, Autonomous Vehicles, and Industrial Automation.

Alternatively, one can look at these technologies from a generality versus specialization perspective. Some technologies are rather transversal or application agnostic, for example, Cloud Computing, Machine-Learning & Neural Networks, and Information Processing. Some, on the contrary, are application- or domain-specific like Online Shopping Platforms, Smart Agriculture & Water Management, and Health Monitoring.

However, it is crucial to stress that, by construction, the novelty of these technologies resides in their *digital* automation nature, whether or not digital control is exercised over tangible capital or intangible processes.

We use patent titles for the calculation of the exposure scores. Each DII patent has a long title that describes the essence of the invention followed by a concise explanation of its use and function. We leverage this structure and mirror it in the industrial and occupational descriptions, as detailed in the next paragraphs.

NAICS 2007. We use the North American Industry Classification System (NAICS) 2007 to represent industries. NAICS underlying construction principle specifically groups industries based on *similarity of their production processes*. Given that our task is to measure technol-

Table 1: List of Digital Technologies

	Family		Emerging Digital 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 digital technologies obtained in Prytkova et al. (2024). For a short description of these technologies, refer to Tables 8 to 10 in Appendix 4.

ogy's relevance to production processes and outputs (products or services), this property of NAICS facilitates matching between digital technologies and industries.

We select the 4-digit (industry group) as the most disaggregated level for which we estimate exposure scores. The 5-digit level of NAICS represents product and market differentiation, while the 4-digit level focuses on commonalities of production processes across variants of products or services. For example, wheat, corn, or rice farming all belong to grain farming. Thus, the functional relevance of digital technologies for industries naturally favors the 4-digit level.

However, to avoid loosing useful information at the product level, we represent each 4-digit industry with a composite textual description by combining the 4-digit description with its nested 5-digit industrial descriptions. Thus, we prioritize functional matching of industries and technologies but account for product and service variety. In sum, we work with 324 4-digit industries represented with composite textual descriptions, replicating Prytkova et al. (2024) procedure.

O⋆NET-SOC 2010. We use O⋆NET 15.1 taxonomy to obtain detailed task descriptions for occupations defined in SOC 2010 classification. In general, the SOC system provides a standardized framework for classifying occupations down to the 6-digit level of disaggregation, i.e., Detailed SOC Occupations. In turn, O⋆NET expands SOC occupations by offering in-depth data on job tasks, skills, and work contexts, adding two more digits of disaggregation down to the 8-digit level, i.e., O⋆NET-SOC Occupations. In this taxonomy, there are 1110 8-digit occupations mapped to 840 6-digit Detailed SOC occupations.[4]

It is worth noting that, according to the stated classification principles, SOC occupations are "classified based on work performed and, in some cases, on the skills, education, and/or training needed to perform the work at a competent level." (U.S. Bureau of Labor Statistics, 2010, p. xii). This means that the primary grouping principle is the *nature of the work performed* hence a category can include occupations focused on similar tasks but with different levels of education and training required. Comparison between European occupational exposure scores obtained in Prytkova et al. (2024) based on ISCO-08 classification and those of the US based on SOC classification further demonstrates this difference, see Figure 7 in the appendix.

We estimate the exposure score of 1110 occupations at the 8-digit level. Different bundles of tasks result in variation between occupations in terms of the functions they perform. Similarly to industries, we measure the relevance of a digital technology to activities and functions performed by a worker at a given occupation. Given that SOC categories are homogeneous regarding the nature of work but not education or skill requirements, we calculate the exposure scores at the most disaggregated level as it provides greater consistency in education, skill lev-

[4]See O⋆NET-SOC 2010 in https://www.onetcenter.org/taxonomy.html for more details.

els, and wages within each occupation. Each occupation is represented by textual description of tasks performed at this occupation, with on average 19.3 tasks per 8-digit occupation.

2.2 Exposure as a Measure of Relevance of a Technology

Construction of Exposure Scores: A Brief Recap. We follow the same method proposed by Prytkova et al. (2024) to construct the exposure scores. First, we preprocess textual data of patents, industries, and occupations to impose a similar structure that brings out function-related content, as detailed in Section 2.1.

Then, we convert the preprocessed texts into *contextual embeddings*, i.e. dense vector representations, using the pre-trained MPNet v2 sentence transformer. The advantage of contextual or dynamic embeddings lies in encoding the semantic meaning of a word given its surrounding context; the same word in two different texts has two different vector representations. Thus, more semantic information is encoded in contextual embeddings than in static ones. In addition to rich semantic representation of individual documents, the MPNet v2 transformer produces embedding space where *distances* between document—vectors represent their semantic (dis)similarity. Thus, documents with similar content are positioned closely while unrelated documents are distant.

First, we calculate the cosine similarity between each patent–industry and patent–occupation pair of embeddings. It represents the cosine of the angle between two vectors, in our case two embeddings. Therefore, the more similar the description of technology and its function to the description of production processes or tasks, the higher the cosine similarity. Because each document—4-digit NAICS industry and 8-digit O*NET-SOC occupation—consists of several sentences, we leverage sentences and produce multiple cosine similarity scores per patent–industry and patent–occupation pair. This redundancy helps to filter out false links between invention and industry or occupation (false positives); see Prytkova et al. (2024) for more details.

Then, we aggregate exposure scores from patent to technology level. We weight each patent based on the number of citations it received in a given year to reflect variation in patent impact. Lastly, we add up weighted cosine similarity scores of patents that constitute a technology cluster and correct it for right skewness. Our final exposure score to technology $k \ \forall \ k \in [1, 40]$ is X_i^k for 4-digit NAICS industries and X_o^k for 8-digit O*NET-SOC occupations.

Interpretation. Our exposure scores are the measure of *relevance* of a technology to an industry or occupation, primarily via the relevance of its functions to production processes or tasks. Whenever the link between technology and industry or occupation is detected, it indicates the potential for adoption, but de facto adoption might not have happened yet. In sum, our exposure score is *a proxy for actual adoption*.

Moreover, the exposure score has simultaneously cumulative and dynamic nature. On the one hand, despite the evolution of the technology cluster over time, each patent's semantic content contributes to the aggregate exposure to the technology as a whole. On the other hand, any change in functions performed by a technology (expansion or shrinkage), if persists long enough, will lead to adjustments in technology's industry or occupation links.

Lastly, the exposure scores are neutral concerning the relationship between technology and labor, meaning they do not encode whether technology and labor act as complements or substitutes in the production of products and services.

We provide these data as an open–-access resource, the 'TechXposure' database. The database includes exposure measures starting at 8–digit O⋆NET-SOC occupations and 4–digit NAICS industries and then at every higher level of aggregation up to the top of its respective taxonomy. It is worth stressing that aggregation to increasingly broader SOC groups maintains homogeneity of task content but entails further dilution of education, skill, and training consistency for reasons discussed in Section 2.1.

2.3 Descriptives

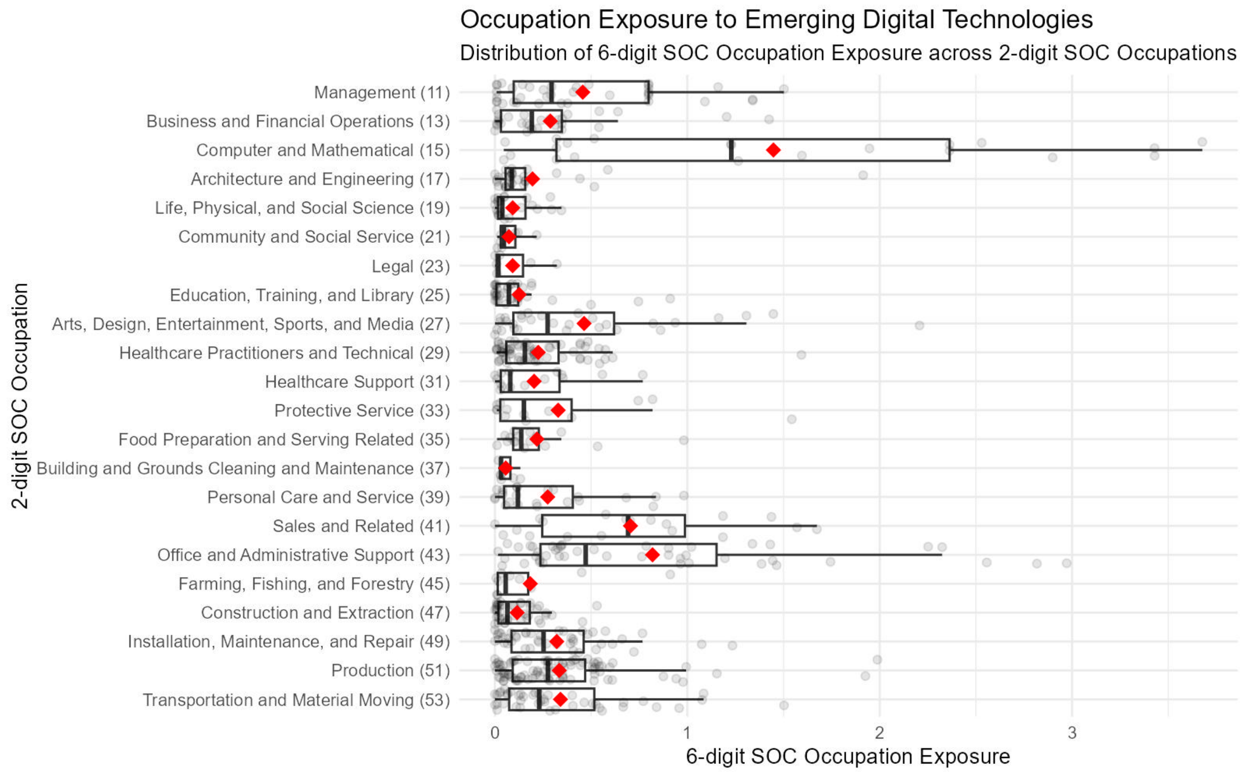
SOC/O⋆NET Occupations. We examine the exposure of occupations to all digital technologies by 2-digit SOC group. The overall exposure of an occupation o is defined as its average exposure across all 40 digital technologies: $X_o = \frac{1}{40}\sum_k X_o^k$.

Figure 1 shows the distribution of exposure of 8-digit occupations grouped by their respective 2-digit SOC major groups. Each boxplot corresponds to a 2-digit SOC group and comprises overall exposure scores of all 8-digit occupations in a given group. Because 2-digit SOC groups comprise mixed education and skill levels as discussed in Section 2.1, Figure 1 helps identify domains of work that are exposed to digital technologies rather than skill-based groups.

To gain additional insights, we combine results from Figure 1 with Table 2 that lists the top-10 most exposed tasks by 2-digit SOC group. We represent tasks as *verbs* used in O⋆NET description of an occupation. The lists of top-exposed tasks include tasks regardless of how frequently they are mentioned in the O⋆NET description; only the semantic similarity between a verb–task and a technology's function defines the list.

Not surprisingly, we observe that the Computer and Mathematics occupation group (15-0000) is the most exposed to digital technologies. This group contains occupations that create new digital technologies as well as themselves can be subject to digital automation due to the

Figure 1: Distribution of O*NET-SOC 2010 Occupation Exposure to Digital Technologies



Notes: This figure presents the distribution of exposure of 8-digit O*NET-SOC occupations to all digital technologies. Occupations with 8-digit codes are assigned to their corresponding 2-digit SOC groups, with each group displayed in its own separate boxplot. Vertical bars indicate the median exposure among 8-digit occupations within the same 2-digit group, and diamond points represent the average exposure for these 8-digit occupations.

highly-codified, algorithmic nature of work. The tasks performed by this occupation group can be routine and non-routine which might contribute to high heterogeneity of exposure within this group. Workers in these occupations are tasked to *develop*, *implement* and *analyze* an application, solution, technology, measures, policy, process, etc; *configure* systems, and equipment; *provide* technical support, assistance, and training; *analyze* information, data, or problem.

The second most exposed occupation groups are Office and Administrative Support occupations (43-0000) and Sales and Related occupations (41-0000). These groups comprise various clerks and salespersons working across all sectors. Predominantly, clerks process and handle information routinely within an organization supporting its operations, while salespersons do so while they survey markets and liaison with a buyer. The most exposed tasks of each category corroborate this fact. Clerical exposed tasks are to *enter* and *process* various data, records, and information; *receive* and *record* messages, documents, payments, calls, items, etc.; *operate* communication and recording equipment and software. Salesperson's exposed tasks are to *determine* sale or contract conditions (amounts, prices, customer's needs, prod-

Table 2: 10 Most Exposed Tasks by 2-digit SOC Occupation Groups

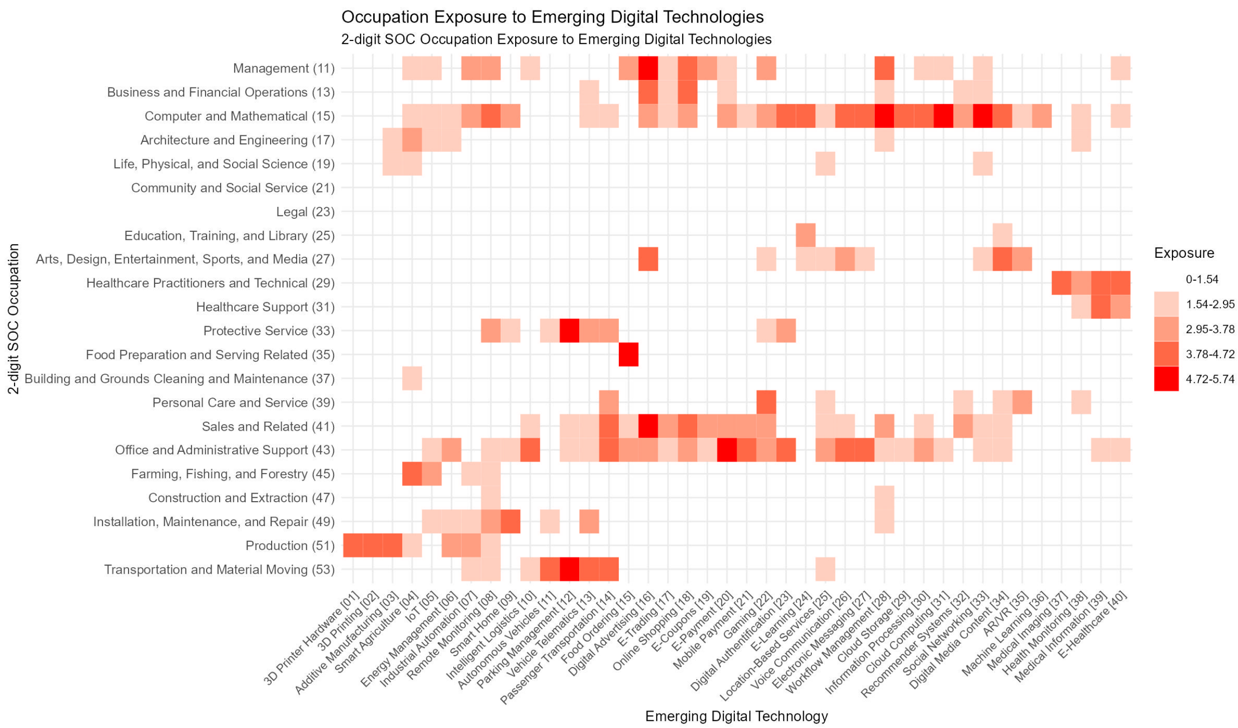
SOC code	Title	Tasks
11-0000	Management	monitor, administer, plan, direct, coordinate, train, develop, execute, position, regulate
13-0000	Business and Financial Operations	process, receive, design, arrange, analyze, measure, implement, develop, maintain, initiate
15-0000	Computer and Mathematical	provide, implement, configure, instruct, analyze, specify, process, develop, evaluate, identify
17-0000	Architecture and Engineering	design, document, perform, determine, select, calibrate, analyze, modify, maintain, integrate
19-0000	Life, Physical, and Social Science	collect, analyze, develop, interpret, program, direct, maintain, create, merge, adjust
21-0000	Community and Social Service	collect, evaluate, identify, provide, collaborate, develop, plan, oversee, instruct, monitor
23-0000	Legal	prepare, authorize, issue, use, read, summarize, examine, retrieve, recommend, enter
25-0000	Education, Training, and Library	provide, organize, design, develop, maintain, catalog, acquire, assess, teach, collaborate
27-0000	Arts, Design, Entertainment, Sports, and Media	communicate, edit, plan, set, regulate, control, obtain, adjust, operate, confer
29-0000	Healthcare Practitioners and Technical	operate, enter, position, maintain, develop, plan, monitor, protect, observe, process
31-0000	Healthcare Support	deliver, operate, prepare, check, perform, record, accept, provide, receive, inventory, store
33-0000	Protective Service	mark, provide, retrieve, enter, locate, record, check, identify, monitor, communicate
35-0000	Food Preparation and Serving Related	record, take, receive, pack, package, prepare, relay, cook, place, provide
37-0000	Building and Grounds Cleaning and Maintenance	study, cover, lift, push, swing, start, confer, connect, provide, carry
39-0000	Personal Care and Service	plan, operate, select, supervise, compute, start, control, observe, maintain, establish
41-0000	Sales and Related	determine, identify, relay, monitor, obtain, book, generate, provide, select, explain
43-0000	Office and Administrative Support	enter, receive, operate, record, arrange, page, process, compute, read, inform
45-0000	Farming, Fishing, and Forestry	weigh, advise, record, set, operate, monitor, manipulate, locate, direct, observe
47-0000	Construction and Extraction	install, perform, drive, coordinate, determine, monitor, layout, plan, demonstrate, purchase
49-0000	Installation, Maintenance, and Repair	adjust, install, repair, diagnose, maintain, test, instruct, set, schedule, mount
51-0000	Production	install, set, load, adjust, monitor, operate, select, transfer, implement, apply
53-0000	Transportation and Material Moving	communicate, operate, determine, drive, calculate, park, analyze, regulate, retrieve, report

Notes:

ucts, etc.), *identify* and *monitor* new clients, markets, opportunities; *relay* sale information back to the firm; *monitor* sales, performance, equipment, compliance; *obtain* information, documentation, permissions, agreements, authorization.

The third most exposed group is Management occupations (11-0000). Managerial roles involve non-routine cognitive tasks that pertain to direction (*administer, direct, develop*), oversight (*monitor, regulate*), implementation (*plan, execute*), and coordination (*coordinate, train*)

Figure 2: Occupation Exposure by Digital Technologies (2-digit SOC 2010)



Notes: Each cell shows the exposure of a 2-digit SOC occupation (row) to a given digital technology (column). Exposure scores below the 80th percentile (0-1.54) are transparent, whereas the four other groups represent respectively the 80th (1.54-2.95), 90th (2.95-3.78), 95th (3.78-4.72), and 99th (4.72-5.74) percentile of the distribution.

of business operation.

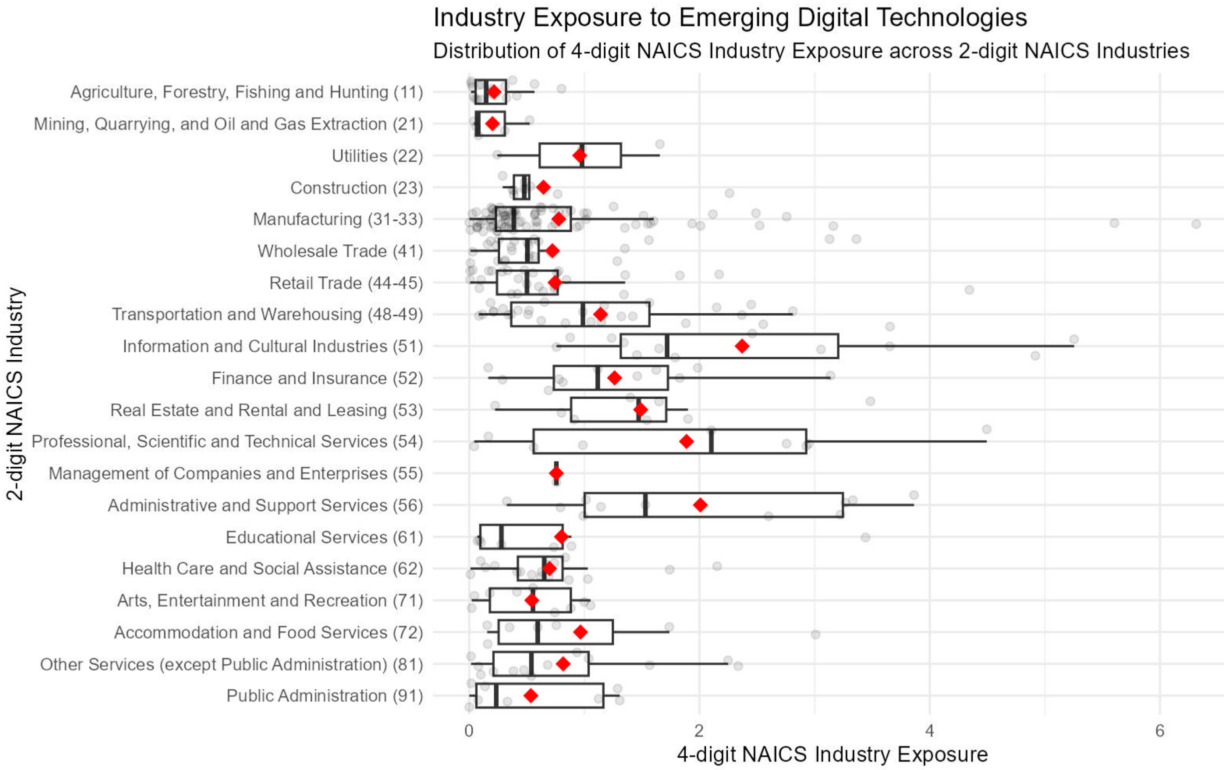
Lastly, we break the overall exposure of 2-digit SOC groups down to the exposure to individual digital technologies. Figure 2 plots exposure scores at the intersections of occupation (rows) and technology (columns) as a heatmap.

We observe that the most exposed occupations discussed above—IT workers, clerks, salespersons, and managers—display relevance to a wide range of digital technologies. All other 2-digit SOC groups show connections confined to a domain-specific subset of digital technologies, i.e. transportation occupations to mobility-related technologies, production and maintenance workers to tangible manufacturing technologies, healthcare occupations to medical technologies, art and design professions to social and digital media, etc. Figure 2 provides further evidence of the domain focus of SOC classification.

NAICS Industries. Analogously to occupations, we analyze the exposure of 2-digit NAICS industries first to all digital technologies and then proceed to individual technologies breakdown.

Figure 3 shows the distribution of 4-digit industries exposure scores grouped by 2-digit

Figure 3: Distribution of NAICS Industry Exposure to Digital Technologies



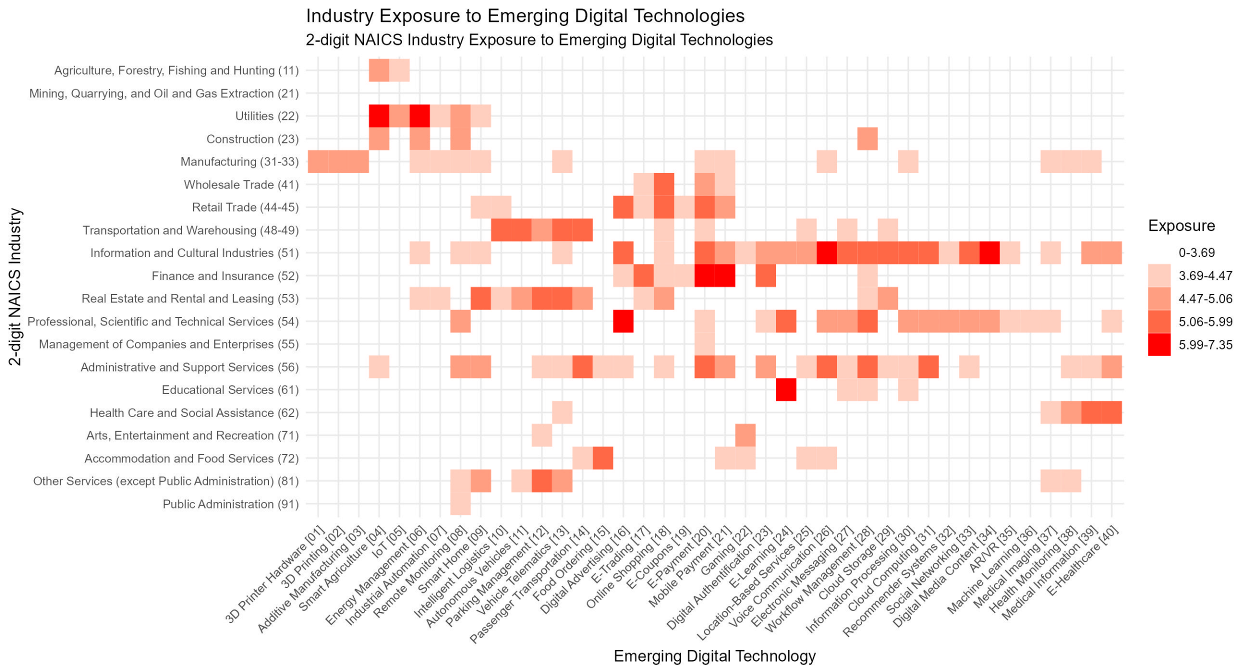
Notes: This figure presents the distribution of exposure to emerging digital technologies across 4-digit NAICS industries, with each 2-digit industry displayed separately in boxplots. Vertical bars indicate the median exposure for all 4-digit industries within the same 2-digit industry, and diamond points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

sector. On average, service sectors are more exposed to digital technologies than manufacturing but the most exposed 4-digit industries are located inside the Manufacturing (31-33) sector. These industries are Computer and Peripheral Equipment Manufacturing (3341) and Communications Equipment Manufacturing (3342) representing the tangible component of digital technologies. The software component is represented by the next two most exposed industries: Data Processing, Hosting, and Related Services (5182) and Computer Systems Design and Related Services (5415).

Unlike occupations, industries show a higher degree of heterogeneity both between and within 2-digit sectors. This is the result of a single sector encompassing various production processes along the value chain; digital technologies can be relevant for different processes and at different stages of the value chain. This contrasts with the SOC taxonomy, which categorizes occupations into homogeneous groups regarding the nature of work.

However, we can trace similarities with occupation exposure. Information and Cultural Industries (51) and Professional, Scientific, and Technical Services (54) combined mirror high exposure of IT professions. Administrative and Support Services (56) are highly exposed like

Figure 4: Industry Exposure by Digital Technologies (2-digit NAICS 2007)



Notes: Each cell shows the exposure of a 2-digit NAICS 2007 industry (row) to a given digital technology (column). Exposure scores below the 80th percentile (0-3.69) are transparent, whereas the four other groups represent respectively the 80th (3.69-4.47), 90th (4.47-5.06), 95th (5.06-5.99), and 99th (5.99-7.35) percentile of the distribution. Figure **??**, in the appendix, presents the same figure at the 2-digit level.

office clerks. Salespersons included in the 2-digit SOC group 41-0000 are scattered across several industries: Wholesale Trade (41), Retail Trade (44-45), Finance and Insurance (52), and Real Estate and Retail and Leasing (53). Management of Companies and Enterprises (55) displays a similar magnitude of exposure as managerial occupations but a near-zero variation.

Figure 4 shows the decomposition of the overall exposure scores into individual digital technologies for 2-digit NAICS sectors. We obtain very similar results to Prytkova et al. (2024) in terms of both pattern and magnitude of industrial exposure because, unlike ISCO and SOC occupational classifications, NACE and NAICS classifications have similar organization principles. Appendix 8 further corroborates this statement.

We observe that the exposure scores follow a diagonal pattern from the top-left to the bottom-right. Tangible technologies on the left are associated with Agriculture (11), Utilities (22), Construction (23), and Manufacturing (31-33) sectors. Intangible technologies are relevant for market services like Wholesale (41) and Retail (44-45) Trade, Finance and Insurance (52), Professional services (54), or Administrative and Support Services (56). Hybrid tangible and intangible technologies such as Embedded Systems and Smart Mobility are relevant for industries in between the extremes that operate physical infrastructures such as Transportation and Warehousing (48-49) and Real Estate, Rental, and Leasing (53). Lastly, public and social services at the end of the NAICS taxonomy are scarcely exposed to digital technologies.

3 Empirical Analysis

We estimate the impact of digital technologies on employment in the United States from 2012 to 2019 using a long-difference approach across 741 Commuting Zones (CZ). We start by presenting the CZ labor market data and our empirical strategy. Then, we estimate the baseline estimate and explore how these impacts vary among different demographic and skill groups.

3.1 Commuting Zone (CZ) Level Data

We use data from the American Community Survey (ACS) to obtain demographics and employment information at the county level (FIPS). We then aggregate these data to the CZ level using the crosswalk from FIPS areas to CZs from Dorn (2009).[5]

For demographic data used as control variables, we calculate the following characteristics for CZs in the baseline year of 2010: population, female population share, elderly population share (age 65 and above), high school graduate share, and college graduate share. It is important to note that the ACS data for educational attainment applies only to individuals aged 25 to 64.

Our main outcome variable is the change in the employment-to-population ratio, computed as the change in the ratio of employed individuals to the total population in each commuting zone. Additionally, to assess changes in employment among demographic and skill groups, we calculate this ratio for specific categories including gender, age groups, and educational levels ranging from less than high school to bachelor's degree or higher.

3.2 Empirical Strategy

We estimate the impact of digital technologies that emerged between 2012 and 2019 on the employment-to-population ratio over the same period using a shift-share approach in long-difference.

We construct a measure of digital technology exposure for each commuting zone c such that:

$$X_c = \sum_j l_{cj} X_j, \tag{1}$$

[5] Instead of using the Quarterly Census of Employment and Wages (QCEW), which provides more granular 3-digit NAICS level data but requires county-level employment imputation, we opt for ACS data. The ACS offers complete and consistent data across sources, whereas QCEW data often lacks values for key industries, with missing data distribution across counties (and consequently, CZs) correlating strongly with population size and digital technology exposure. This non-randomness of missing data is problematic for analysis.

Table 3: Average Employment Shares in 2010 and Exposure (2012–2019) by Industry

		Emp. Share		
	NAICS Industry	Mean	SD	Shock
11	Agriculture, Forestry, Fishing and Hunting	1.4	4.1	0.3
21	Mining, Quarrying, and Oil and Gas Extraction	0.5	2.1	0.2
22	Utilities	0.9	0.1	1.2
23	Construction	7.1	2.1	1.0
31-33	Manufacturing	11.0	23.1	2.8
41	Wholesale Trade	3.0	0.4	1.9
44-45	Retail Trade	11.5	1.4	2.2
48-49	Transportation and Warehousing	4.2	1.1	2.2
51	Information and Cultural Industries	2.3	0.6	3.7
52	Finance and Insurance	4.9	2.9	1.9
53	Real Estate and Rental and Leasing	2.0	0.4	2.5
54	Professional, Scientific and Technical Services	6.2	6.9	3.2
55	Management of Companies and Enterprises	0.1	0.0	0.7
56	Administrative and Support, Waste Management and Remediation Services	4.0	0.7	3.2
61	Educational Services	9.2	3.6	1.7
62	Health Care and Social Assistance	12.9	4.4	1.2
71	Arts, Entertainment and Recreation	2.0	0.9	0.9
72	Accommodation and Food Services	6.8	2.1	1.7
81	Other Services (except Public Administration)	4.9	0.3	1.9
91	Public Administration	4.9	4.6	0.6

Notes: This table presents the average employment shares in 2010 by 2-digit NAICS industry, which is averaged across all the CZs, and the average exposure to digital technologies which is the shock in the shift-share. CZs are weighted by population in 2010. The first column indicates the 2-digit NAICS codes, the second column is the name of the NAICS industry, the third column is the average employment share in 2010, the fourth column gives the standard deviation (SD), and the fifth column corresponds to the industry exposure to digital technologies.

where l_{cj} is the employment share of industry j in CZ c and $X_j = \frac{1}{40}\sum_k X_j^k$ is the average exposure score across all digital technologies for the 2-digit NAICS industry.

Table 3 shows the average employment shares by industry in 2010 across commuting zones, alongside the shocks, which reflect the average exposure scores. This table highlights the primary sources of variation in our shift-share instrument. Three industries have an average employment share above 10%: Manufacturing (31-33) with a notably high exposure score of 2.8 and a standard deviation (SD) of 23, indicating significant employment differences between CZs; Retail Trade (44-45), which also shows considerable exposure (score of 2.2) but with less variation across CZs (SD of 1.4) due to its proportional relation to population and employment size; and Health Care and Social Assistance (62), which, despite being more heterogeneous across CZs (SD of 4.4), has a lower digital technology exposure with an average shock of 1.2.

The three industries most exposed to digital technologies are Information and Cultural Industries (51) with an average shock of 3.7, Scientific and Technical Services (54) with an average shock of 3.2, and Manufacturing (31-33) with an average shock of 2.8. Although the first two industries experience the largest shocks, they have relatively small employment shares.

Notably, there is greater heterogeneity across CZs in the share of Scientific and Technical Services compared to Information and Cultural Industries.

The main sources of variation in the shift-share variable stem from industries that either have large—sometimes heterogeneous—employment shares across CZs or are highly exposed, thereby experiencing significant shocks.

We argue that the industry exposure to digital technologies, X_j, representing the shock in our shift-share design, is quasi-exogenous to employment changes at the US CZ level. Our exposure scores are based on the semantic similarity between patents and NAICS industry descriptions. Notably, only 36% of the patents in our sample are US-based, suggesting that the advancement of these technologies is predominantly a global phenomenon. Furthermore, although we use US standard classifications for industry descriptions, these are generally not US-specific, as similar industries exist globally. The uniformity of industry descriptions across countries implies that global technological trends are not uniquely driven by US local labor market characteristics.

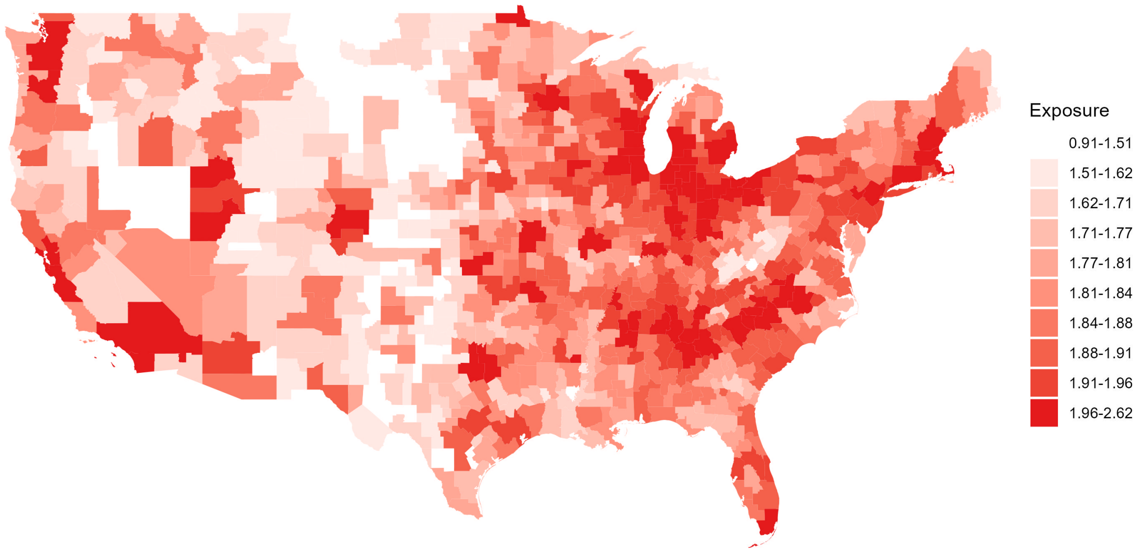
To reinforce our point, we recalculate the exposure scores after *excluding* US patents from the patent sample, denoted as $\widetilde{X}_j$. We then recalculate our shift-share CZ exposure using global (non-US) exposure, $\widetilde{X}_c$, according to Equation (1). In our empirical specification below, we use $\widetilde{X}_c$ to instrument the US exposure shift-share X_c.

Since our exogeneity condition is derived from the shock rather than the share, we rely on the equivalence result of Borusyak et al. (2021) to define our identifying assumptions. First, we assume that shocks are quasi-randomly assigned to industries, implying that local CZ employment dynamics in the US do not influence the relevance of technologies to industries as expressed by X_j.

Second, we assume that CZs more exposed to these technologies are not disproportionately affected by other labor market shocks or trends, and that the number of observed shocks is sufficiently large. Concerning this latter point, a common statistic reported is the Herfindahl–Hirschman Index (HHI) of average shock exposure, calculated as $\sum_j l_j^2 = 0.077$, where l_j is the average employment share in industry j in 2010 across all CZs, as reported in Table 3.

Figure 5 shows the geographic distribution of CZ exposure to digital technologies (2012–2019). The figure highlights significant disparities in digital technology exposure, with coastal and highly urbanized areas showing the highest levels. In the Northeast, extending from Washington D.C. to New York City and Boston, the darkest red shades indicate the highest exposure. This region is densely populated and hosts a concentration of finance and technology industries, a characteristic mirrored by similar high-exposure areas on the West Coast. Urban centers such as San Francisco, Seattle, and Los Angeles are also marked in dark red, underscoring their roles as technology hubs, particularly the Silicon Valley near San Francisco.

Figure 5: Geographic Distribution of CZ Exposure to Digital Technologies (2012–2019)



Notes: This figure illustrates the geographic distribution of exposure to digital technologies for Commuting Zones (CZ). CZ 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. CZ are categorized into deciles. CZ are shaded according to their exposure level, with the legend indicating the range of exposure.

Moderate exposure levels, depicted in mid-range red tones, appear in the Mid-Atlantic and Southeastern states, encompassing cities like Atlanta and Charlotte, along with parts of Florida. The Midwest, which includes cities such as Chicago, Minneapolis, and Detroit, also shows moderate exposure. This is attributed to a mix of traditionally high manufacturing shares and the increasing influence of digital-intensive industries.

Conversely, CZs in the Central and Mountain West regions, including Wyoming, Montana, and parts of Nebraska, display the lowest exposure levels due to their rural nature and lower population densities. Similarly, areas in the Deep South, such as Mississippi, Arkansas, and Alabama, also exhibit low levels of exposure to digital technologies.

Texas and California display a wide spectrum of CZ exposure, in which urban areas such as Dallas, Houston, and Los Angeles feature high levels of digital technology exposure, whereas more rural regions like West Texas and California's Central Valley show significantly lower levels. Transitional zones, where exposure shifts from high in urban centers to moderate or low in rural outskirts, are also observed around major inland cities like Denver and Phoenix.

These patterns highlight the importance of demographic characteristics, such as population density, in CZs as key confounding factors that should be controlled for in the regression framework outlined below. Additionally, they underscore the necessity of including state fixed effects when estimating the relationship with employment changes.

To account for these potential confounding factors related to demographic characteristics, we estimate the following empirical specification:

$$\Delta Y_c = \alpha + \beta X_c + Z\delta + \phi_{d(c)} + u_c, \quad (2)$$

where ΔY_c is the change in the employment to population ratio (in pp.) for CZ c between 2012 and 2019, X_c is the CZ exposure to digital technologies, as defined in Equation (1) and standardized, Z is a set of covariates which capture regional characteristics (including the logarithm of the population, the proportion of females, the proportion of the population aged over 65, the share of high school graduates, and the share of college graduates), $\phi_{d(c)}$ represents state fixed effects, and u_c is the error term.

Our coefficient of interest in this specification is $\hat{\beta}$. Since exposure is standardized across CZs, we interpret our coefficient as the effect of a one-standard-deviation increase in CZ exposure on the employment-to-population ratio, expressed in percentage points. Following recent literature on shift-share designs, we report AKM0 shift-share standard errors, which account for arbitrary cross-regional correlation in the regression residuals Adão et al. (2019).

3.3 Relationship between Exposure and Employment

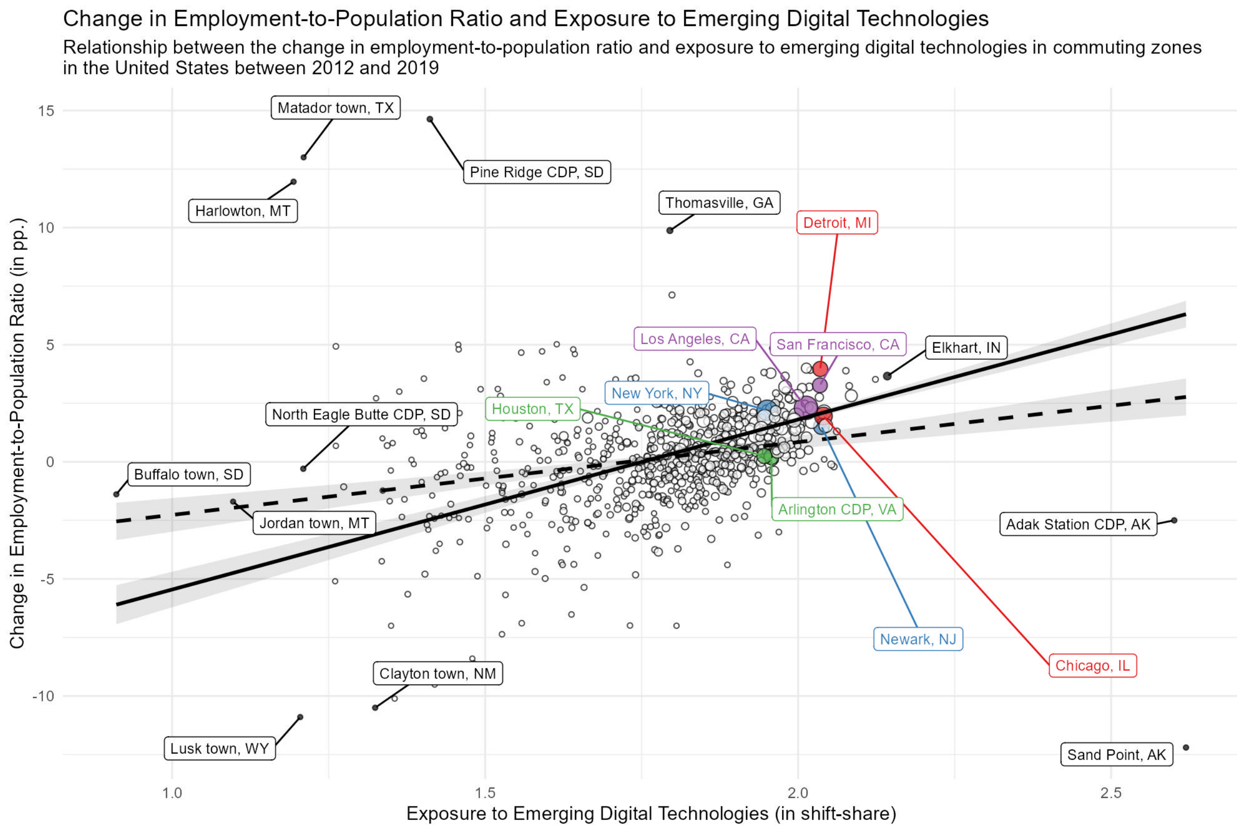
Figure 6 shows the relationship between the change in the employment-to-population ratio and exposure to emerging digital technologies across CZs between 2012 and 2019. Each point on the plot represents a CZ, weighted by its population size in 2010, with the horizontal axis measuring exposure to digital technologies and the vertical axis showing the change in the employment-to-population ratio in percentage points.

We observe a positive correlation, as depicted by the solid line, indicating that greater exposure to digital technologies is associated with an increase in the employment-to-population ratio over the period. This positive correlation persists, though at a smaller magnitude when considering unweighted observations, as shown by the dashed line. The largest US cities, such as San Francisco, New York, and Detroit, appear to be among the most exposed commuting zones, exhibiting, on average, positive changes in the employment-to-population ratio.

We estimate the instrumental variable shift-share approach described in Equation (2) in which we instrument the CZ exposure to digital technologies X_c, with its world counterpart which the CZ exposure calculated only with patents from other countries, that is, $\widetilde{X}_c$. Table 4 reports the results.

Column (1), which includes only state fixed effects, shows that exposure to digital technologies positively impacts the employment-to-population ratio, with an estimate significant at the 1% level. A one-standard-deviation increase in CZ exposure corresponds to a 0.94 per-

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 US Commuting Zones (CZ) between 2012 and 2019. Each point represents a CZ. 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 dashed line represents the same correlation when observations are unweighted. The grey shaded area indicates the 95% confidence interval.

centage point (pp.) change in the employment-to-population ratio. This finding indicates a robust positive relationship between digital technology and employment in a basic specification that includes only state fixed effects.

As demographic controls are introduced in column (2), the coefficient decreases to 0.67 yet remains positive and significant. This decrease suggests that part of the initially observed effect can be attributed to the demographic characteristics of the CZ. This is the most conservative estimate of the impact of digital technologies on the employment-to-population ratio, indicating that a one-standard-deviation increase in exposure results in a 0.67 pp. change. Given the average employment-to-population ratio of 0.58 in 2012, this change corresponds to a 1.1% increase.

In column (3), we introduce the industry share in 2010 as a control to more accurately isolate the impact of digital technologies on employment within CZ. This adjustment accounts

Table 4: Effect of Digital Technologies on US Employment

	IV – Dep. var: Δ Emp-to-pop. Ratio (2012-2019) × 100						
	Weighted				Unweighted		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exposure to Digital Technologies	0.94^{***}	0.67^{***}	0.92^{***}	0.81^{**}	0.08	0.55^{**}	0.76^{***}
	(0.12)	(0.11)	(0.16)	(0.21)	(0.16)	(0.19)	(0.09)
State FE	✓	✓	✓	✓	✓	✓	✓
Demographics		✓	✓	✓	✓	✓	✓
Industry share			✓	✓	✓	✓	✓
Exclude Top 10% Exp. CZ				✓			
Exclude Bottom 10% Pop. CZ						✓	
Exclude Bottom 20% Pop. CZ							✓
R^2	0.57	0.60	0.61	0.49	0.25	0.33	0.41
Adj. R^2	0.54	0.57	0.58	0.44	0.18	0.27	0.35
Num. obs.	741	741	741	666	741	666	592

Notes: This table presents the estimates of the effect of exposure to digital technologies on the employment-to-population ratio in US commuting zones (CZ) between 2012 and 2019. The exposure to digital technologies is constructed as a shift-share and standardized, while the change in employment-to-population ratio is expressed in percentage points. Regressions in columns (1) to (4) are weighted by population in 2010 and regressions in columns (5) to (7) are unweighted. Column (1) includes state fixed effects; Column (2) adds demographics controls in 2010, including the logarithm of the population, the proportion of females, the proportion of the population aged over 65, the share of high school graduates, and the share of college graduates; Column (3) adds the share of employment in the industry sector in 2010. Column (4) excludes the top 10% most exposed CZ. Column (5) is the unweighted regression; Column (6) excludes the bottom 10% most populated CZ in 2010; and Column (7) excludes the bottom 20%. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors between parentheses are derived following the AKM0 inference procedure from Adão et al. (2019).

for the employment structure and mitigates the influence of sector-specific trends such as routinization and international trade, which could confound the effect of digital technology exposure. By adding this control, we ensure that our findings capture the direct relationship between digital technologies and employment outcomes, rather than being artifacts of broader industrial trends. The coefficient increases slightly to 0.92, suggesting that our estimate remains robust against other labor market trends predominantly affecting CZs with a significant industry sector.

In column (4), we exclude the top 10% of commuting zones (CZ) with the highest digital exposure. This adjustment results in a reduced coefficient of 0.81, which remains positive and significant, although standard errors have increased. This specification highlights the substantial influence that highly exposed CZ may have on the overall positive relationship. Nonetheless, the persistent positive effect, despite the exclusion of these highly exposed areas including San Francisco, Boston, and Detroit, confirms that the observed relationship is not exclusively driven by the most digitally advanced CZ.

In column (5), we replicate the specification from column (3) but with unweighted observations. Thus, we assign equal weight to each CZ regardless of its population size in 2010. Here, the impact of digital technology exposure on employment becomes statistically insignificant. Columns (6) and (7), which also use unweighted observations but exclude the bottom

10% and 20% of the least populated CZ, suggest that the insignificance of the unweighted estimate arises from CZ with minimal populations. These are effectively outliers, as depicted in Figure 6. When these outliers are progressively excluded, the coefficients increase, aligning more closely with those observed in column (3).

3.4 Heterogeneity by Demographic and Skill Groups

Table 5 presents the estimation for the demographic groups, where each column gives the estimation for a group.

In columns (1) and (2), we find minimal gender differences in the impact of digital technologies on employment. Both female and male workers show a positive employment impact, with a one-standard-deviation increase in the employment-to-population ratio by 0.2 percentage points for females and 0.27 for males, corresponding to increases of 0.75% and 0.94%, respectively. These effects apply to individuals aged 16 to 64, as data on gender breakdown for older age groups is limited in the ACS data and they constitute a small fraction of employment

Examining age-specific effects, both youth (16–19 years) and young adults (20–24 years) experience significant positive impacts, with coefficients of 0.15 and 0.12 respectively, as shown in columns (3) and (4). The youth group shows a notable 7% increase in their employment ratio—the highest among all age groups—which may reflect a rise in part-time and entry-level job opportunities driven by digital technologies. Similarly, young adults benefit from a 2% increase in their employment ratio, suggesting that digital technologies likely facilitate labor market entry after higher education or vocational training by aligning skills with technological demands.

Conversely, column (5) reveals that the core working-age group (25–44 years) suffers a significant negative impact. A one-standard-deviation increase in CZ exposure results in a 0.45 percentage point decline (1.7%) in the employment ratio. This indicates that US workers in this age group are most affected by job displacement, where technological advances demand higher skill levels, leading to the obsolescence of existing skills or automation of their tasks.

In columns (6) and (7), for older workers aged 45-–54 and 55–-64 years, we see a reversal of the negative trend observed for the core working-age group. Both age groups show significant positive effects from digital technologies, with coefficients of 0.43 and 0.38, respectively. These results suggest that older workers effectively adapt to new technologies, possibly through skill accumulation and experience that boost employability in a shifting labor market. Additionally, their senior positions may shield them from the negative employment impacts more commonly experienced by non-managerial roles.

Table 5: Effect of Digital Technologies on US Employment by Demographic Groups

	IV – Dep. var: Δ Emp-to-pop. Ratio (2012–2019) $\times$ 100								
	Gender (Y16-64)		Age						
	Female	Male	Y16-19	Y20-24	Y25-44	Y45-54	Y55-64	Y65-74	Y75+
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Exposure to Digital Technologies	0.20^{**}	0.27^{***}	0.15^{***}	0.12^{**}	-0.45^{***}	0.43^{***}	0.38^{***}	0.03	0.01^{**}
	(0.08)	(0.07)	(0.02)	(0.06)	(0.11)	(0.11)	(0.07)	(0.03)	(0.01)
Emp-to-pop. Ratio in 2012	0.26	0.28	0.02	0.05	0.25	0.14	0.09	0.02	0.00
Change (in %)	0.75	0.94	7.02	2.02	-1.70	3.01	4.00	1.42	2.64
R^2	0.55	0.53	0.41	0.38	0.45	0.75	0.62	0.38	0.20
Adj. R^2	0.51	0.49	0.36	0.33	0.41	0.73	0.59	0.33	0.13
Num. obs.	741	741	741	741	741	741	741	741	741

Notes: This table presents the estimates of the effect of exposure to digital technologies on the employment-to-population ratio in US commuting zones (CZ) between 2012 and 2019 by demographic groups. The exposure to digital technologies is constructed as a shift-share and standardized, while the change in employment-to-population ratio is expressed in percentage points. Columns (1) and (2) correspond to the estimates for female and male workers. Columns (3) to (9) correspond to the estimates for workers by age group. Regressions are weighted by population in 2010 and include state fixed effects as well as demographic controls (including the logarithm of the population, the proportion of females, the proportion of the population aged over 65, the share of high school graduates, and the share of college graduates). $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors between parentheses are derived following the AKM0 inference procedure from Adão et al. (2019).

Table 6: Effect of Digital Technologies on US Employment by Education Groups

	IV – Dep. var: Δ Emp-to-pop. Ratio (2012–2019) $\times$ 100			
	Less HS	HS	Some Coll.	Bach.+
	(1)	(2)	(3)	(4)
Exposure to Digital Technologies	-0.05^{**}	-0.16^{**}	0.22^{**}	0.35^{**}
	(0.02)	(0.06)	(0.07)	(0.13)
Emp-to-pop. Ratio in 2012	0.04	0.12	0.15	0.17
Change (in %)	−1.08	−1.32	1.41	2.05
R^2	0.56	0.62	0.47	0.60
Adj. R^2	0.52	0.59	0.43	0.57
Num. obs.	741	741	741	741

Notes: This table presents the estimates of the effect of exposure to digital technologies on the employment-to-population ratio in US commuting zones (CZ) between 2012 and 2019 by education groups (for age group 25–64 years). The exposure to digital technologies is constructed as a shift-share and standardized, while the change in employment-to-population ratio is expressed in percentage points. Columns (1) to (4) correspond to the estimates for workers with different levels of education (in order): less than high school graduate, high school graduate (includes equivalency), some college or associate's degree, and bachelor's degree or higher. Regressions are weighted by population in 2010 and include state fixed effects as well as demographic controls (including the logarithm of the population, the proportion of females, the proportion of the population aged over 65, the share of high school graduates, and the share of college graduates). $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors between parentheses are derived following the AKM0 inference procedure from Adão et al. (2019).

Lastly, the oldest age groups (65–74 and 75+ years) experience minimal impact from digital technology on employment, with coefficients of 0.03 and 0.01, respectively, as shown in columns (8) and (9).

Table 6 presents the estimated impact of digital technologies on employment ratios across different educational groups, with each column representing a specific education level. Due to constraints in the ACS data structure, the employment changes reported pertain only to the core working-age and senior populations, aged 25–64 years.

Digital technologies negatively impact individuals with less than a high school education, as evidenced by a coefficient of -0.05 in column (1). This finding suggests that lower educational attainment may limit individuals' ability to adapt to technological advancements, potentially resulting in job displacement.

The negative impact is more pronounced for those with a high school diploma. A one-standard-deviation increase in digital technology exposure results in a -0.16 percentage point change in their employment ratio, corresponding to a -1.32% change. This significant negative effect suggests substantial displacement, likely due to the insufficiency of a high school education to leverage digital skills in the labor market.

Conversely, individuals with some college education experience positive employment outcomes from exposure to digital technologies. A one-standard-deviation increase in exposure results in a 0.22 coefficient, corresponding to a 1.41% increase in the employment-to-population ratio for this group. This indicates that individuals with some college education possess more adaptable skills or relevant qualifications that better align with the demands of a technologically integrated labor market.

Lastly, individuals holding bachelor's degrees or higher reap the greatest benefits from digital technologies. A one-standard-deviation increase in CZ exposure leads to a 0.35 percentage point rise in the employment-to-population ratio, translating to a 2.05% increase over the period. This positive impact likely arises from their advanced skill levels and specialized knowledge, which are highly complementary to digital technologies.

3.5 Timing of the Impact on Employment

A valid concern with using exposure to emerging digital technologies rather than adoption is the unclear timing of their impact on employment. Consider two sub-periods, 2012–2016 and 2016–2019, during which digital technologies evolve with different vintages. Our current long-difference specification does not clarify whether the observed employment impacts are due to the initial vintages in the first sub-period, the later vintages, or a combination of both, possibly with delays in adopting earlier vintages.

To address this timing issue, we implement an empirical specification that estimates the impact of digital technologies on the employment-to-population ratio over both sub-periods. Specifically, we recalibrate the shift-share exposure of CZs for each period $t = \{1, 2\}$ as follows:

$$X_c^t = \sum_j l_{cj} X_j^t,$$

where X_j^t is the average exposure score of each 2-digit NAICS industry to all digital technologies during the period t.

As expected, X_c^1 and X_c^2 are highly correlated since CZs most exposed between 2012 and 2016 tend to remain highly exposed between 2016 and 2019, reflecting the cumulative development of digital technologies. Including both exposure measures in the same specification would lead to multicollinearity issues. To address this, we calculate the change in exposure across the periods as $\Delta X_c \equiv X_c^2 / X_c^1$, which reflects the intensification of digital technology exposure in the second period relative to the first.[6]

Table 7 presents the estimates by sub-periods. Column (1) shows that exposure to digital technology positively affects the employment-to-population ratio during the first sub-period from 2012 to 2016. In column (2), we introduce the relative change in exposure between the

[6]Figure 6 in the Appendix shows the correlation between CZ exposures in both periods and the change in exposure between them.

Table 7: Effect of Digital Technologies on US Employment by Periods

	IV – Dep. var: Δ Emp-to-pop. Ratio $\times$ 100				
	2012–2016		2016–2019		2012–2019
	(1)	(2)	(3)	(4)	(5)
Exposure P1 (2012–2016)	0.36^{**}	0.37^{**}		0.45^{***}	0.82^{***}
	(0.09)	(0.13)		(0.19)	(0.21)
Exposure P2 (2016–2019)			0.30^{***}		
			(0.08)		
Exposure P2/P1		0.01		0.23^{*}	0.24
		(0.21)		(0.16)	(0.24)
R^2	0.44	0.44	0.56	0.56	0.61
Adj. R^2	0.40	0.40	0.52	0.52	0.57
Num. obs.	741	741	741	741	741

Notes: This table presents the estimates of the effect of exposure to digital technologies on the employment-to-population ratio in US commuting zones (CZ) between 2012 and 2019 by periods. The exposure to digital technologies is constructed as a shift-share and then standardized, the ratio between exposure over both sub-periods is also standardized, while the change in employment-to-population ratio is expressed in percentage points. All regressions are weighted by population in 2010 and include state fixed effects and demographics controls in 2010 (including the logarithm of the population, the proportion of females, the proportion of the population aged over 65, the share of high school graduates, and the share of college graduates). Columns (1) and (2) focus on the change in the employment-to-population ratio between 2012 and 2016; Columns (3) and (4) between 2016 and 2019; and Column (5) between 2012 and 2019. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors between parentheses are derived following the AKM0 inference procedure from Adão et al. (2019).

first and second sub-periods as a control variable. The associated coefficient is negligible and statistically insignificant, indicating that the relative changes in exposure during the second sub-period do not influence employment changes in the first sub-period.

In columns (3) and (4), we examine changes in the employment-to-population ratio over the second sub-period (2016–2019). Column (3) shows a positive and significant impact of digital technologies on employment during this period. In column (4), when we include the lagged exposure from the first sub-period and the relative change in exposure, the lag coefficient remains positive and significant. Additionally, the coefficient for the intensification of exposure in the second period (relative to the first) is positive and significant at the 10% level. This suggests that, on the one hand, CZs with greater relative exposure in the second period experience employment gains, which indicates CZs catching up; on the other, the impact of digital technologies persists across periods, as new vintages build upon earlier ones, reinforcing the technological advances of initially highly exposed CZs.

Lastly, column (5) examines the change in the employment-to-population ratio over the entire period. The coefficient for the change in exposure is positive but not statistically significant, suggesting that CZs that became more exposed in the second period did not necessarily benefit from it. Consequently, the primary impact appears attributable to the first-period exposure. However, it remains unclear whether the employment effects are due to the initial period alone or if the second period's developments in digital technologies also contribute, building upon the advancements from the earlier period.

4 Conclusion

This paper shows that digital technologies have an overall positive impact on US employment, with a 1.1% increase in the employment-to-population ratio following a one-standard-deviation rise in digital technology exposure across commuting zones. However, this overall benefit masks significant heterogeneity among demographic groups. Younger workers (ages 16–24) and older workers (ages 45–64) experience positive employment outcomes, whereas core working-age individuals (ages 25–44) face adverse effects, likely due to the obsolescence of skills and task automation. Additionally, our results reveal a skill-biased technological change, negatively impacting workers with lower educational levels while benefiting those with college degrees. This research highlights the importance of reskilling and upskilling policies for vulnerable workers, particularly those in core working-age groups and with lower education levels, who are negatively impacted by digital technologies.

Another significant contribution of our paper is the expansion of the TechXposure database to incorporate U.S. standard classifications, thereby adapting it to the US context. Using the advanced NLP approach developed in Prytkova et al. (2024), we estimate the exposure of industries and occupations to digital technologies that have emerged over the past decade at a very granular level (i.e., up to 4-digit level for industries and 8-digit level for occupations). These exposure scores capture the *relevance* of a digital technology to specific occupations or industries. We believe this enriched dataset offers valuable new avenues for future research, particularly in exploring the impact of specific digital technologies on US employment and wages, extending beyond the traditional focus areas like AI and robotics.

References

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. (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.

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

Bonfiglioli, A., Crinò, R., Gancia, G., and Papadakis, I. (2024). Artificial Intelligence and Jobs: Evidence from US Commuting Zones. *Economic Policy*, 40(121):145–194.

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

Dorn, D. (2009). *Essays on Inequality, Spatial Interaction, and the Demand for Skills*. PhD thesis, University of St. Gallen.

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.

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. (2019). Technological Change and Occupations over the Long Run. Working paper.

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. Working paper.

Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. (2017). Technological Innovation, Resource Allocation, and Growth. *The Quarterly Journal of Economics*, 132:665–712.

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

Prytkova, E., Petit, F., Li, D., Chaturvedi, S., and Ciarli, T. (2024). The Employment Impact of Emerging Digital Technologies. Working Paper.

Song, K., Tan, X., Qin, T., Lu, J., and Liu, T.-Y. (2020). Mpnet: Masked and permuted pre-training for language understanding. *Advances in Neural Information Processing Systems*, 33:16857–16867.

U.S. Bureau of Labor Statistics (2010). 2010 SOC user guide. Accessed: 2025-02-17.

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

Appendix

Table 8: Description of the 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 digital technologies ranging from 1 to 14.

Table 9: Description of the 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.

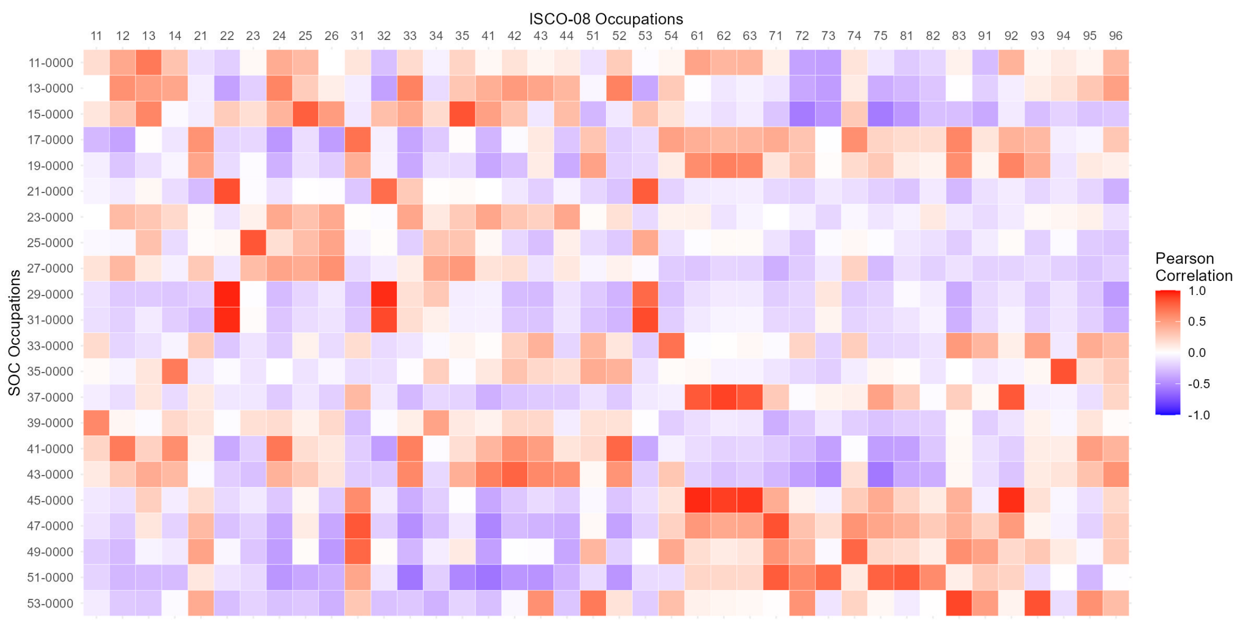
Notes: This table provides descriptions of digital technologies ranging from 15 to 28.

Table 10: Description of the 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.

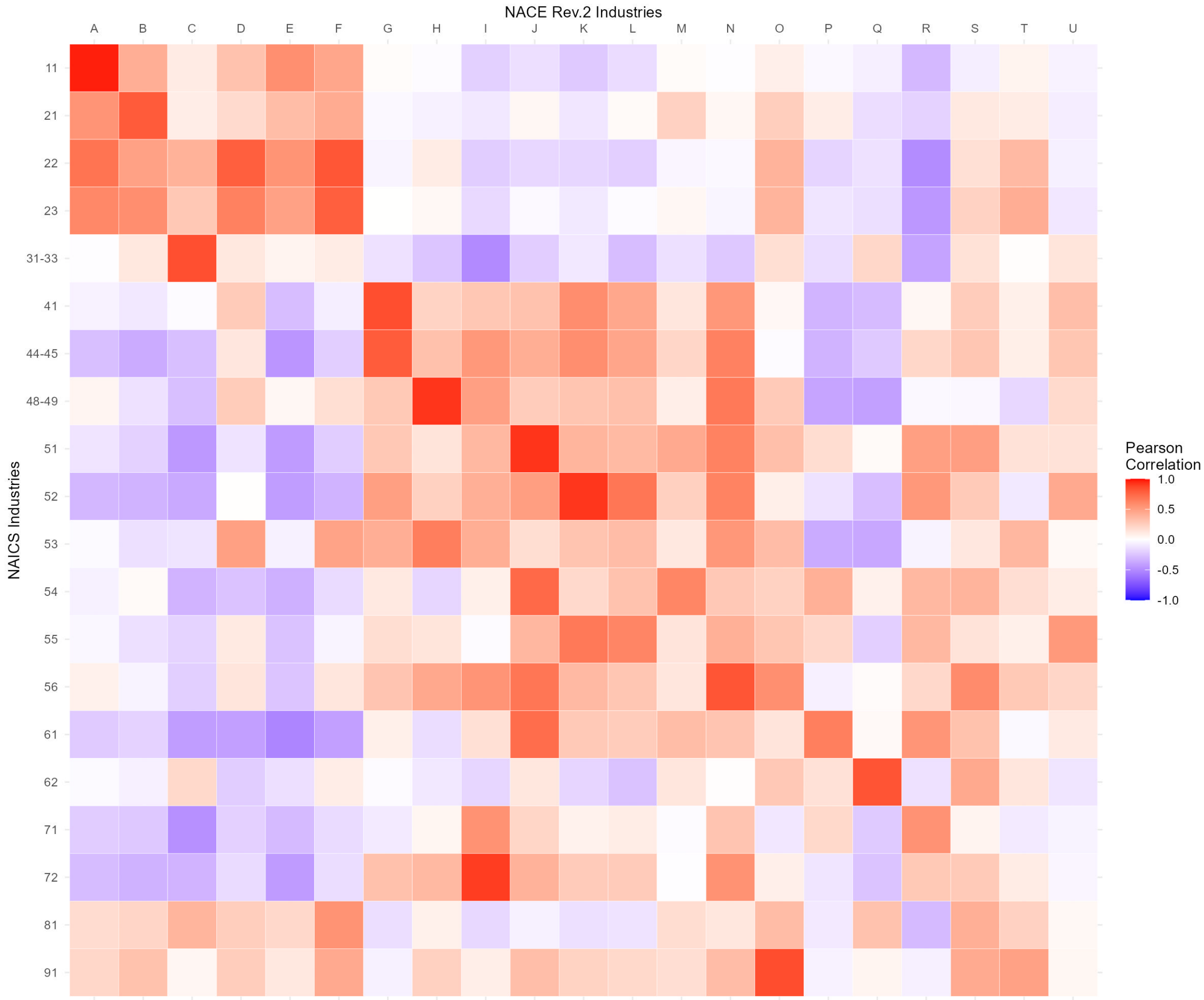
Notes: This table provides descriptions of digital technologies ranging from 29 to 40.

Figure 7: Correlation of Exposure Scores between 2-digit SOC and ISCO Occupational Classifications



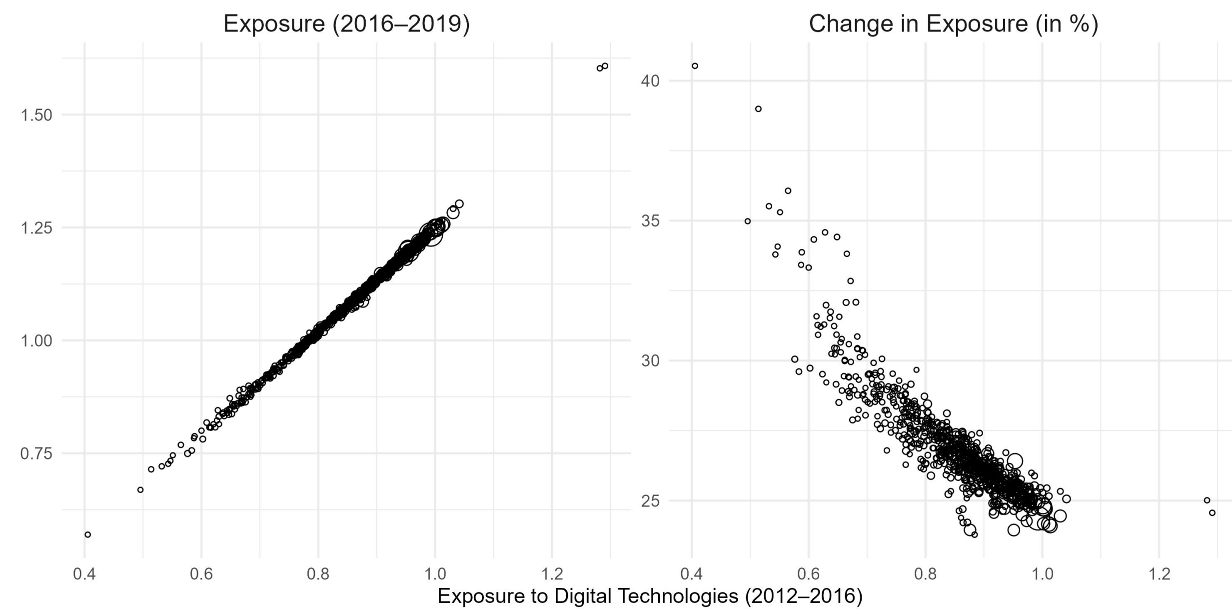
Notes: This figure presents the correlation of exposure scores derived using the 2-digit SOC 2010 occupation classification and derived using the 2-digit ISCO-08 classification.

Figure 8: Correlation of Exposure Scores between 1-digit NAICS and NACE Industrial Classifications



Notes: This figure presents the correlation of exposure scores derived using the 2-digit NAICS 2007 industry classification and derived using the 1-digit NACE Rev.2 classification.

Figure 9: Correlation of Commuting Zone Exposure Scores Between Periods



Notes: The left-hand side figure illustrates the correlation between the exposure to digital technologies for Commuting Zones (CZ) over the period 2012–2016 and 2016–2019, and the right-hand side figure shows the correlation between the CZ exposure over the period 2012–2016 and the change in CZ exposure between both periods, expressed in percent. CZ 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 2016 and 2016 to 2019. Observations are weighted by their population in 2010.