Technological change and labor outcomes: measuring the role of automation and labor augmentation

J. Daniel Aromí[1,2], María Paula Bonel[1,2], Martín Llada[1]

[1]IIEP UBA-Conicet, Facultad de Ciencias Económicas, Univ. de Buenos Aires.

[2]Facultad de Ciencias Económicas, UCA, Buenos Aires, Argentina.

Contributing authors: josedanielaromi@uca.edu.ar; mpaulabonel@uca.edu.ar; Martin.llada@uba.ar;

Abstract

This work studies the link between new technologies and labor outcomes. Using a comprehensive collection of business narratives for the period 2005-2024, we construct a detailed characterization of technology covering 151 industries and 545 occupations. We document large asymmetries in the importance of augmentation and automation across sectors and occupations. In most industries and occupations, labor automation is more prominent than labor augmentation. Estimated augmentation is positively associated with employment and wage growth. This link is more intense in the case of white-collar workers. In contrast, we find no significant association in the case of automation.

Keywords: Technological change, labor demand, automation, labor augmentation

JEL Classification: E24 , J23 , J24

Note:: Authors declare no conflict of interest.

1 Introduction

The impact of technological change on labor is a function of the extent to which the implemented innovations feature automation of tasks previously performed by workers versus complementarities that augment human capabilities. This link between technology and labor is of central importance and has been the focus of numerous analyses (Acemoglu et al., 2022; Aghion, Antonin, Bunel, & Jaravel, 2023; A.K. Agrawal, Gans, & Goldfarb, 2023; D. Autor, Chin, Salomons, & Seegmiller, 2024; Brynjolfsson, Li, & Raymond, 2025; Johnson & Acemoglu, 2023; Leontief & Duchin, 1986).Despite ongoing efforts, the lack of precise, systematic information on how new technologies are adopted remains a significant obstacle to research in this field. As a result, much of the existing analysis has concentrated on automation in manufacturing or assessed the effects of new technologies within specific industries. Developing a granular, cross-sectoral framework that captures the relationship between technology and labor continues to be a complex and unresolved challenge.

In this work, we propose a novel strategy in which business narratives are used to estimate technology-labor complementarities and automation over the period 2005-2024. We believe that this type of data can prove particularly valuable given the ambiguous nature of technological innovations.The extent to which innovations complement or substitute labor depends largely on how they are adopted (A. Agrawal, Gans, & Goldfarb, 2019; A.K. Agrawal et al., 2023). In many cases, a single technological breakthrough—even within the same organization—can lead to labor augmentation for some workers while resulting in automation for others. Business narratives can be used to extract information regarding these details and, in this way, resolve ambiguities.

We collect a broad set of business discussions from earnings conference calls transcripts and classify relevant excerpts in terms of the presence of references to automation or augmentation. After implementing human validation of automatic classifications, we use these labels to compute industry-level automation and augmentation metrics. Finally, we combine these metrics with BLS industry-occupation labor data. In this way, we are able to describe technological developments and assess their labor implications based on an industry/occupation map with a high level of specificity (151 industries, 545 occupations).

Our metrics demonstrate large asymmetries in the importance of augmentation and automation across sectors and occupations. Broad sectors such as Information and Professional, Scientific, and Technical Services are characterized by high levels for both metrics. On the other hand, we find that for important sectors, such as Hotels and Restaurants, Construction, and Utilities, these indicators are very low. These contrasting measures seem plausible and suggests that business discussions provide valuable information regarding important heterogeneities in the relationship between labor and new technologies across economic sectors.

In most industries and occupations, labor automation is found to be more prominent than labor augmentation. This is particularly evident in the case of Manufacturing, Transport and logistics industries. For these broad sectors, the frequency of discussions that are classified as instances of automation is up to four times the frequency of discussions classified as instances of augmentation. In contrast, there are no broad sectors showing a high level of augmentation and a low level of automation.

We estimate the relationship between labor outcomes and the indicators characterizing technologies using parsimonious panel data models. According to these estimates, augmentation is positively associated with employment and wage growth. This association is more intense in the case of white-collar occupations. In contrast, we find no association in the case of automation. These results are consistent with expansions in labor demand as a result of the implementation of new technologies that increase labor productivity and create new tasks for human labor. The null results for automation could be explained by indirect channels that compensate the direct negative effect of automation on labor demand (Aghion et al., 2023; A. Agrawal et al., 2019; A.K. Agrawal et al., 2023; D. Autor, 2022). These indirect effects could include the expansion of sectors due to increased productivity and the emergence of complementary tasks that are linked to automation innovations and are not precisely captured by the labor augmentation metric. Beyond the sometimes challenging classification of innovations into labor augmenting and automation, a general message that emerges from our estimates is that, during the period of our analysis, new technologies are not typically associated to a negative effect on labor demand.

Our study follows a considerable set of contributions that analyzed the impact of technological change on labor outcomes. For example, four decades ago, Leontief and Duchin (1986) used a general equilibrium framework to evaluate the impact of automation with a focus on manufacturing, administrative jobs, healthcare, and education. Their analysis suggests that there are no reasons to expect a large negative impact on labor demand. More recently, there has been a surge in the number of studies with this focus (Acemoglu & Autor, 2011; Acemoglu, Kong, & Restrepo, 2024; Acemoglu & Restrepo, 2018, 2019; Aghion, Antonin, Bunel, & Jaravel, 2020; Aghion et al., 2023; D. Autor & Salomons, 2018). Some of these analyses have led to somber outlooks. For instance, Johnson and Acemoglu (2023) provide a cautionary message: "Current tech sectors efforts are excessively focused on ever more automation... Automation needs to be accompanied by creating new tasks that make workers more productive, at least if you want to increase the odds that we will have any meaningful form of shared prosperity in the future.". But there are other views that communicate a more nuanced assessment in which efficiency gains from automation generate growth and, in this way, the impact of new technologies turns ambiguous. These views are in part supported by contributions that have measured productivity gains associated to the adoption of AI tools (Brynjolfsson & Hitt, 2003; Brynjolfsson et al., 2025; Dell'Acqua et al., 2023). Reflecting this perspective, A. Agrawal et al. (2019) recommend "... caution on drawing broad inferences from the research on factory automation... ...in forecasting the net near-term consequences of artificial intelligence for labor markets.". These contrasting views demonstrate the challenges associated with providing a precise characterization of new technologies and their implications for labor demand.

.

In terms of the general framework, our study follows the task approach to labor economics (Acemoglu & Autor, 2011; Acemoglu et al., 2024; Acemoglu & Restrepo, 2018, 2019). Under this perspective, the economy is characterized by the set of tasks required to produce goods and services. Tasks can be performed by workers or machines. Technological change alters the allocation of tasks through displacement and reinstatement effects.

Our paper is closely related to the wide-ranging analysis presented in D. Autor et al. (2024). Through novel methodologies, this study provides detailed evidence on how technological innovations lead to new occupations. The authors use text data from patent records to generate indicators of labor augmentation and automation for the period 1940-2018. The metrics are produced identifying patents with text similar to occupations, which are classified as instances of augmentation, and patents with text similar to tasks, which are classified as instances of automation. The study finds that augmentation causes new occupations and higher employment and automation produces reductions in labor demand. Departing from this precedent, our study uses business narratives instead of patent data. We conjecture that this is a valuable alternative, as it might provide detailed descriptions of the various ways in which new technologies are adopted by business organizations.

Another closely related paper provides a comprehensive characterization of the relationship between new technologies and labor is Tomlinson, Jaffe, Wang, Counts, and Suri (2025). With a focus on AI, this work measures its applicability to different occupations. Their primary data are conversations in which users seek assistance from a large language model. Informing about complementarities and instances of labor substitution, the authors identify activities which are typically performed by AI and instances in which workers perform the task assisted by AI. Compared to this study, our analysis is focused on an ample set of technologies that goes beyond large language models, covers a longer time period and estimates the relationship between technology and labor market outcomes.

In terms of data and methodology, our work is connected to previous studies that have extracted economic information from earnings conference calls (Aromi, Bonel, Llada, et al., 2024; Hassan et al., 2025). At the time of writing, we know of no work that has analyzed the impact of new technologies on labor with this data source. The two studies that have analyzed new technologies focused on the impact on the value of the firm and the process of technology difussion (Bloom, Hassan, Kalyani, Lerner, & Tahoun, 2021; Chava, Du, & Paradkar, 2020).

This paper is organized as follows. In the next section, we detail the data and methodology. In section 3, we inspect the resulting metrics of augmentation and automation. The following section focuses on the association between labor outcomes and the indicators that characterize new technologies. We provide concluding remarks in the last section.

2 Data and Methodology

The main data in our analyses are earnings conference calls transcripts. Earnings conference calls are meetings in which company executives provide a description of recent results and evaluate future scenarios. This exposition is followed by a Q&A section in which investors, analysts and journalist are able to place questions. We collected all available transcripts of companies whose shares are traded in US markets for the period 2005-2024. Each of the documents reflects these discussions in a very accurate manner. Our data source is Refinitiv, a commercial database.[1] The corpus contains over 250,000 documents from 7,449 companies, with each transcript averaging about 400 sentences. The number of documents by broad areas is provided in table 1. Firms in this dataset are classified by SIC industry.

Sector	**Companies**	**Transcripts**	**Sentences**
Manufacturing	2572	90848	36567398
Services	1377	43398	17426019
Finance, Insur. and Real estate	1077	42903	17059658
Other (not classified)	959	19831	7885318
Transp., Commun., Electric, Gas...	521	20243	8328570
Retail Trade	355	14876	6332239
Mining	346	11565	4805239
Wholesale Trade	147	6194	2360357
Construction	74	3248	1428184
Agriculture, Forestry and Fishing	21	496	172289
Sum	**7449**	**253602**	**102365271**

Table 1: Distribution of companies, transcripts, and sentences by sector

The second type of data used in this study are labor outcomes data. More specifically, we analyze employment and annual wage data from the May 2004 and May 2024 vintages of the Occupational Employment and Wage Statistics produced by the Bureau of Labor Statistics. Employment and annual wages data are expressed at 4 digit NAICS industry level and SOC six-digit occupation.[2] To implement our analysis we need to merge the datasets that are expressed using different classifications. In the case of earning conference call transcripts we use a crosswalk transform data classified by SIC into data classified according to NAICS. We also need to use crosswalk to combine the two vintages of employment and wage data since they are expressed using different versions of NAICS. The resulting dataset covers 151 industries and 545 occupations. This is an unbalanced panel with 13278 industry-occupation observations.

[1] https://eikon.refinitiv.com/.

[2] For some cases in which the four digit NAICS data was unavailable we selected data available at a higher level of disaggregation.

2.1 Classification of business discussions

Business discussions constitute an unstructured type of data from which information can be extracted after preprocessing the documents and implementing a sequence of classification steps. Our methodology involves a combination of traditional dictionary methods and the use of modern large language models. First, we split the text into sentences. Then, we use a dictionary of phrases to identify sentences that refer to new technologies and extract an excerpt of the text that includes that sentence. Next, we use a large language model to classify the selected excerpts in terms of two aspects: the presence of references to labor issues and, for each those excerpts that refer to labor issues, we generate a second classification that identifies references to automation or labor augmentation. In the final step, these classifications are aggregated to produce industry indicators that characterize the relationship between new technologies and labor.

More in detail, we apply four steps to preprocess and classify business discussions as described below:

1. Selection of phrases that refer to new technologies: each document is split into sentences and, for each sentence, we evaluate if it includes at least one phrase that belongs to a dictionary of new technology terms. The dictionary was constructed by the authors and is reported in figure A of the appendix. For example, the dictionary includes terms such as robots, automation, artificial intelligence and algorithm. The dictionary is intended to capture a broad set of applications that rely on electronic devices and computer software. For each sentence that includes at least one term related to new technologies, we extracted an excerpt that includes the immediately preceding and succeeding sentences. This step results in a collection of 210,728 excerpts.
2. Identify selected excerpts that refer to labor issues: the task is performed by a large language model. We use OpenAI's GPT4-o-mini and binary classification. Out of the 210,728 excerpts related to new technologies only 99,349 are classified as related to labor issues.
3. Classify excerpts in terms of augmentation vs. automation: The subset of excerpts labeled as labor related in step 3 are classified using a 0-10 scale where 0 indicates the content is strongly associated with labor automation and 10 indicates a strong association with labor augmentation. This task is also performed by OpenAI's GPT4-o-mini.

To generate the classifications we access OpenAI's API and submit a large sequence of prompts to request the classification tasks. As shown in figure 1, we composed a single prompt template to request the labels corresponding to the classifications of steps 3 and 4.

Figure 2 shows the distribution of labels that classify excerpts in terms of augmentation vs. automation. The first notable feature is the high frequency of the label 0. This is consistent with predominant views regarding the strong focus on automation that has dominated technological change Johnson and Acemoglu (2023). On the other

Task: You will analyze 5 excerpts from earnings conference call transcripts and classify it based on the expressed views regarding the relationship between technology and the workforce (substitution vs. complementarity).
Classify the excerpt based on whether it aligns more with technology as **substitution/automation** (replacing human labor) or technology as **labor augmentation/complementary** (enhancing human skills and capabilities).
Use a 0 to 10 scale, where:

- **0** = Fully aligned with technology as automation/labor-saving/worker substitution
- **10** = Fully aligned with technology as labor-augmenting/skill-enhancing/worker empowerment

If the text is not related to the labor force, do not provide a number, just choose the label `'na'`. That is, if the role of workers is not connected to the excerpt use the `'na'` label.
Provide the output as a comma separated sequence of labels, each corresponding to the sequence of excerpts. Do not provide any text beyond that comma separated sequence of labels.
Here are the 5 excerpts to classify:
`{Text.to.classify}`

Fig. 1: Prompt template for excerpt classification

hand, there is also an important fraction of excerpts that are assigned a label of 10 confirming labor augmentation is another relevant theme that is present in business discussions.

Table 2 shows sample excerpts and the corresponding labels assigned by the large language model. We selected a diverse set of excerpts to illustrate the depth of the corpus. Selected excerpts correspond to early and late years of our sample period. Also, the excerpts discuss diverse sectors of the economy: manufacturing, security services, information services, financial services and healthcare. This example shows two instances in which there are very strong contrasts in the nature of the labor implication of new technologies. This contrast is correctly captured by the labels assigned by the large language model.

2.2 Validating large language model's classifications

To assess the performance of large language model's classifications, we perform manual classifications of excerpts. We drew a random sample of 500 excerpts that contain new technology terms and use the prompt template to guide human classifications that were carried out by the authors.

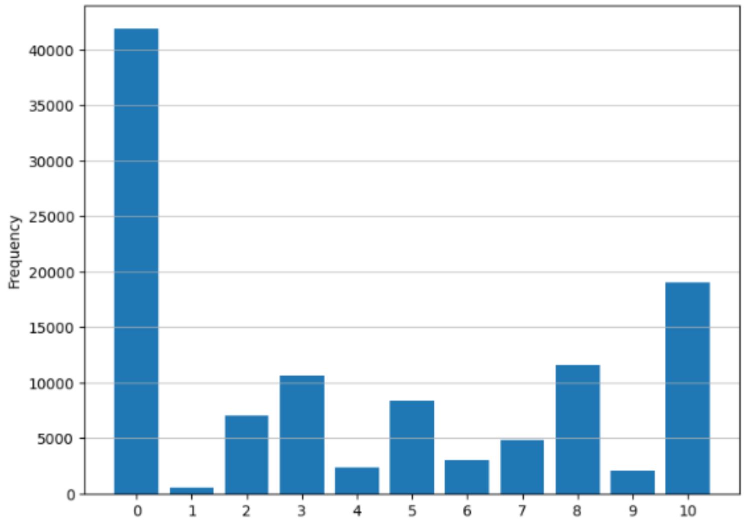

Fig. 2: Histogram of labels classifying excerpts in terms of to augmentation vs. automation.

First, we consider the classification of excerpts according to work relatedness. For this binary classification task, we computed the confusion matrix shown in table 3. We find that under the current specification of the automatic classification exercise, we are able to identify most positive instances of labor related discussions but, at the same time, we observe a high number of false positives. According to human experts, 175 excerpts are related to labor issues out of which 151 are correctly flagged by automatic classification. This desirable outcome contrasts with 141 false negatives. This outcome suggests the value of further refinements in the specification of the automatic classification exercise.

We also evaluate the performance of automatic classifications for the case of automation and augmentation. Table 4 compares the labels assigned by human authors and the large language model for four sample excerpts. For these cases, we find a strong correspondence between human and automatic classifications. This result is extended considering the 500 excerpts of the validation exercise. Indicating a high level of agreement between human and large language model classifications, the Pearson correlation coefficient is 0.73. In conclusion, the validation exercises suggest that the large language model produces satisfactory augmentation vs. automation labels, detects most of the labor-related excerpts, but displays a high fraction of false negatives in this second task.

2.3 Indicators of augmentation and automation

We construct industry level indicators of automation and augmentation measuring the frequency of discussions that refer to strong complementarities and intense substitution respectively. For augmentation, given an specific industry, we count the number

Table 2: Sample classified excerpts

Year	Company	Excerpt	Label
2007	iRobot	... we developed a proof of concept robot by integrating a TASER electronic control device on to our PackBot platform. Using a PackBot equipped with an integrated TASER, police officers will be able to remotely engage, neutralize and control dangerous suspects.	10
2024	Infosys Tech. Ltd.	We have 100,000 employees trained in generative AI areas. We have developed a range of use cases and benefit scenarios across different industries for our clients.	10
2005	Mitek Systems, Inc.	... technology has not been able to automatically screen these checks for authentication... ...with the increased processing speed, the lower cost of memory, the advances in algorithms and pattern recognition and neural networks that we deploy, we can now provide that kind of 100% screening. Automate processes which were formerly labor-intensive and not nearly as accurate.	0
2024	Nutex Health Inc	... we remain committed to embracing advanced technologies that can improve patient care and operational efficiencies whether it's through electronic medical records, telemedicine capabilities artificial intelligence to automate our coding billing collection services or other digital tools.	0

Note: the labels indicate a value between 0 and 10 where 0 indicates the content is strongly associated with labor automation and 10 indicates a strong association with labor augmentation.

		GPT4-o-mini	
		Yes	No
Human	Yes	151	24
	No	141	184

Table 3: Confusion matrix for work relatedness excerpt classification

of new technology excerpts that are labor related and are assigned the label 10 when classified in terms of augmentation vs. automation. This quantity is divided by the total number of words in the transcripts corresponding to that industry and rescaled multiplying by 1000. Similarly, for automation, we compute the frequency of excerpts that are labeled 0 per thousand words in the corresponding industry transcripts.

More formally, to define the indices characterizing technology by industry i, we start by identifying a set of transcripts Tr_i and a collection of associated excerpts E_i.

Table 4: Sample classified excerpts

Year	Excerpt	Human	GPT4
2024	... a new chat application with its own learning curve. We're just starting to scratch the surface of what's possible when thinking *AI less as an assistant and more as a teammate.* Today, AI can take on individual tasks and *assist in completing workflow steps.*	8	8
2016	... collaborations are both progressing through *joint robotic-assisted PCI* research and clinical programs. *The physicians are using our robotic system to treat patients*, and we are initiating collaborative research projects to further improve patient care as well as physician and staff safety. Initial work with physicians has been productive,...	9	10
2022	... there's about probably a \$15 million to \$30 million opportunity just around *automation and centralization of functions that are being done manually today* in a very decentralized fashion.	0	2
2021	... we offer our customers competitive advantage with robots, *automating functions that previously had to be done manually.* Our technology crosses the gap and makes functions like robotic picking, packing and sortation useful in a commercial setting.	1	0

Note: human and GPT4-o-mini labels indicate a value between 0 and 10 where 0 indicates the content is strongly associated with labor automation and 10 indicates a strong association with labor augmentation.

Then, the indicator of augmentation is given by:

$$Aug_i = \frac{\sum_{e \in E_i} I(label_e = 10)}{\sum_{tr \in Tr_i} wordcount_{tr}} * 1000 \tag{1}$$

For automation we apply a similar formula that measures the frequency of 0 labeled excerpts per thousand words:

$$Aut_i = \frac{\sum_{e \in E_i} I(label_e = 0)}{\sum_{tr \in Tr_i} wordcount_{tr}} * 1000 \tag{2}$$

In the empirical analysis described below, both indices are expressed in terms of log counts.

In addition to industry technology characterization, we are interested in characterizing occupations in terms of augmentation vs. automation. To construct this metric, we exploit employment data by industry and occupation to compute weighted average that take the form of mappings from industry indicators to occupation scores. More specifically, we use the May 2004 vintage of BLS's Occupational

Employment and Wage Statistics to set the employment of industry i in occupation o as the an industry-occupation weight: $Empl_{i,o}^{2005}$. Then, the metric of augmentation for occupation o is given by the weighted average of industry scores. We follow the same procedure to generate the corresponding indicator of automation. Formally, the occupation level indicators are given by:

$$Aug_o = \frac{\sum_{i \in I} Aug_i * Empl_{i,o}^{2005}}{\sum_{i \in I} Empl_{i,o}^{2005}} \tag{3}$$

$$Aut_o = \frac{\sum_{i \in I} Aut_i * Empl_{i,o}^{2005}}{\sum_{i \in I} Empl_{i,o}^{2005}} \tag{4}$$

2.4 Empirical models

We estimate the association between technological change and new technologies using BLS's employment and wage data. We evaluate three outocomes: employment growth, annual wages growth and wagebill growth. These data are at the industry-occupation level. Our main specification is given by:

$$Growth_{i,o}^{05-24} = \alpha + \beta_{Aug} Aug_i + \beta_{Aut} Aut_i + \lambda_I + \lambda_O + u_{i,o} \tag{5}$$

Where $Growth_{i,o}^{05-24}$ is cumulative growth of the respective labor outcome for industry i and occupation o and λ_I and λ_O are 2-Digit industry and occupation fixed effects. We cluster standard errors by first digit of NAICS industry code and first digit of SOC occupation code and compute weighted least squres using beginning of sample employment levels. We winsorized the labor outcomes indicators using the top and bottom percentiles. In the regressions, the indicators of augmentation and automation are expressed as log counts.

3 Estimates of labor augmentation and automation

Figure 3 shows the estimates that characterize the link between new technologies and labor at the 2 digit level of the NAICS code. The figure shows important heterogeneities across broad industry sectors. We apply simple k-means clustering to identify sectors with similar metrics. There is a group that displays a high level of automation and augmentation. This cluster includes knowledge and information activities, personal services and administrative support. The other three groups display a low level of augmentation but show large differences in terms of automation. A high automation cluster covers logistics (transport, storage,...) and a couple of manufacturing groupings (including machinery, food and textile manufacturing). There is also a second cluster with low augmentation and intermediate automation that covers retail, wholesale trade, mining and a third manufacturing grouping (including chemical, paper and plastic manufacturing). Finally, the cluster with low augmentation and low automation covers construction, utilities and important service sectors such as hotels and restaurants, finance and insurance, healthcare and arts, entertainment and recreation.

The clusters we computed provide one engaging way to structure a high level representation of new technologies and their impact on labor. Instead of analyzing the two indicators independently, we could express the information as two dimensions where the first one measures the relative intensity of automation vs. augmentation and a second dimension expresses the total level references to new technologies. For each industry, this would represent the angle and the hypotenuse projected in the plane. We show this alternative groupings in figure B1 of the appendix. This alternative groupings show some interesting differences. For example, the two large services sector corresponding to healthcare and hotel and restaurants are now part of different clusters since, in the case of healthcare, in relative terms, augmentation is notably stronger than automation.

It is important to note that these 2-digit NAICS groupings hide important differences that are observed inside of each broad industry group. In the appendix, figures B2 and B3 show augmentation and automation indicators for selected 4 digit NAICS sectors. This disaggregated metrics show that while a similar level of automation is observed for Truck Transportation and for Taxi and Limousine Services, augmentation is substantially more intense in the case of Taxi and Limousine Services. Turning to healthcare industry, we observe that Office of Physicians is a NAICS industry characterized by significantly more augmentation than Nursing Care Facilities. Finally, we can also verify that broad industry 54 (Professional, Scientific, and Technical Services) covers industries with moderate levels of automation and augmentation, such as Architecture and Engineering Services, industries with high levels for both metrics, such as Advertising and Public Relations, and industries with even higher levels, such as Computer System Design and Related Services.

Table 5 shows the industries with the highest and lowest levels of augmentation and automation. Among the five industries with the highest levels of augmentation, three correspond to NAICS sector 51, Information, one is a manufacturing industry (Magnetic and Optimal Media) and the remaining corresponds to Electronic and Precision Equipment Repair and Maintenance. In the case of automation, manufacturing industries take 4 out of 5 top positions. The remaining industry with notable high automation is a service: Couriers and Express Delivery.

3.1 Characterization of occupations

As described above, we are also able to characterize occupations in terms of augmentation and automation. This results from combining industry classifications with employment data at the industry-occupation level. The resulting classifications are shown, at the 2 digit SOC code level, in figure 4. Computer and Mathematical Occupations and Education, Training and Library Occupation show the highest levels of augmentation. With differences in terms of automation, we find three large categories display some of the lowest levels of augmentation. These categories are Production Occupations, with very high automation, Construction and Extraction Occupations, with moderate levels of automation, and Food Preparation and Serving Related Occupations, with low automation. The lowest level of automation corresponds to Personal

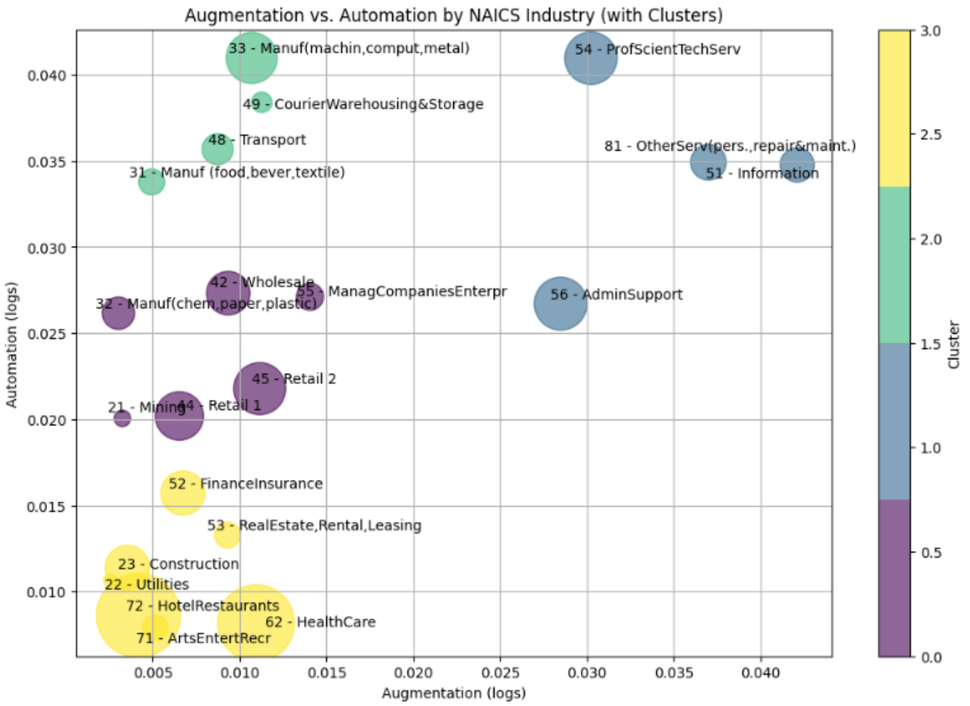

Fig. 3: Estimates of labor augmentation and automation - 2 Digit NAICS code

Note: The size of each circle is proportional to broad industry 2005 employment. Collors indicate k-means clusters.

NAICS code	Industry Name	Indicator
Augmentation - Top 5:		
8112	Electronic and Precision Equipment Repair and Maintenance	0.093
519	Web Search Portals, Libraries, Archives, and Other Information Services	0.093
5162	Media Streaming Distr. Serv., Social Networks, and Other Media Networks and Content Prov.	0.092
5132	Software Publishers	0.085
3346	Manufacturing and Reproducing Magnetic and Optical Media	0.081
Automation - Top 5:		
3335	Metalworking Machinery Manufacturing	0.258
3149	Other Textile Product Mills	0.255
3352	Household Appliance Manufacturing	0.103
4921	Couriers and Express Delivery Services	0.087
3341	Computer and Peripheral Equipment Manuf.	0.081

Table 5: Ranking of NAICS sectors by augmentation and automation

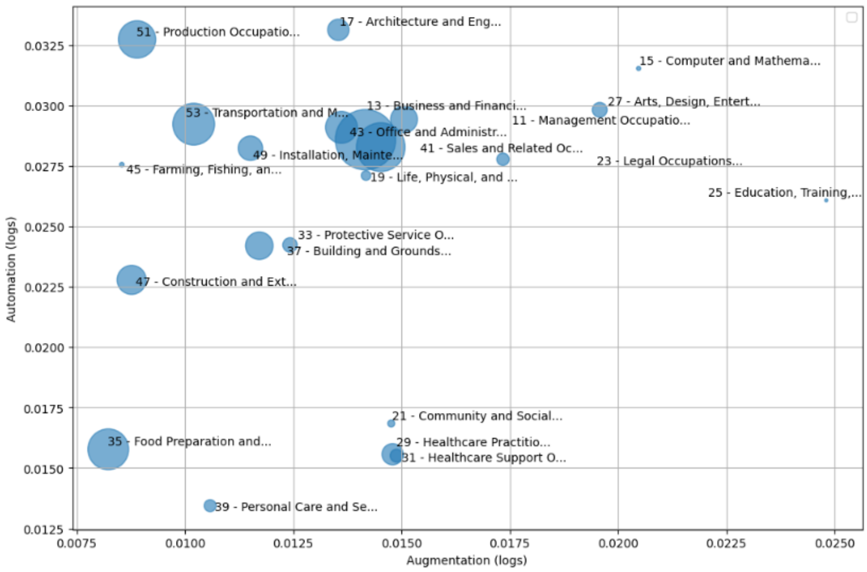

Fig. 4: Augmentation and automation indices: 2-digit SOC occupations

Care and Service Occupations. In another interesting example, we find that while various healthcare occupations display a similar level of automation than Food Preparations and Serving Related Occupations, their level of augmentation is significantly higher.

As in the case of industry classification, we note that broad categories hide important heterogeneities. In figure B4 of the appendix, we report the classifications that correspond to occupations covered in the 2 digit broad category Sales and Related Occupations. We observe important differences inside this broad grouping. For example, we can construct a sequence of categories with increasing levels of augmentation that start with Cashiers and Retail Sales Persons, follows with Insurance Sales and Telemarketers and ends with Sales Engineers and Advertising Sales Agents, the two categories with the highest level of augmentation.

4 New technologies and labor outcomes

We estimate the association between labor augmentation and automation metrics and three alternative labor outcomes. Labor outcomes correspond to the cumulative growth of employment, annual wages, and wagebill as reflected in the May 2005 and May 2024 of BLS's Occupational Employment and Wage Statistics. In the first specification, augmentation and automation metrics are expressed as log-counts. As a robustness check, we also estimate a model in which the regressors are indicator functions for the top deciles of the automation and augmentation metrics. This alternative specification is convenient given its straightforward interpretability.

	Wagebill	Employment	Wages
Log of frequency			
Augmentation	10.55***	3.24**	7.92***
	(1.4)	(0.9)	(1.4)
Automation	0.30	0.56	-0.45
	(1.1)	(0.5)	(1.9)
Top decile dummies			
$I[Aug_i > Perc.90]$	0.41**	0.10*	0.31**
	(0.1)	(0.5)	(0.1)
$I[Aut_i > Perc.90]$	0.00	0.13	-0.12
	(0.2)	(0.1)	(0.1)

Table 6: Estimated labor outcomes models

Notes: WLS weighted by industry-occupation employment in 2005. The specification includes fixed effects by 2-digit industry and occupation code. Standard errors clustered by first-digit industry and occupation codes.

Table 6 shows the estimated coefficients for alternative model specifications. We find a consistent emerging message. In the case of labor augmentation, we find a positive association with labor outcomes. For example, the second panel indicates that mean employment growth is 10% higher when the augmentation metric is in the top decile. This association is more intense in the case of wages and wagebills. In the last case, estimates indicate that mean wagebills growth is 40% when the augmentation metric is in the top decile. On the other hand, in the case of automation, we do not observe an association between any labor outcome and the text-based metric.

These results are consistent expanding labor demand in response to technological developments that increase labor productivity and create new tasks for human labor. The null results for the case of the automation can be explained by countervailing channels that are linked to automation. While the direct effect of automation is to replace human labor, there are indirect effects such as the expansion of sectors due to increased productivity and the emergence of complementary tasks that are not precisely captured by the labor augmentation metric (Aghion et al., 2023; A. Agrawal et al., 2019; A.K. Agrawal et al., 2023; D. Autor, 2022).

Previous studies have discussed the differential impact of technological change as a function of the type of work and the implications of this patterns for inequality (Acemoglu et al., 2022; D.H. Autor & Dorn, 2009; Prettner & Strulik, 2020). To advance an analysis of this question with our novel data, we divide occupations into blue-collar and white-collar. We apply a straightforward split in which white-collar occupations correspond to management, knowledge, and professional occupations, that are occupations with SOC codes that start with the first digit equal to 1 or 2. The remaining categories are labeled blue-collar.

	Wagebill	Employment	Wages
$I[Aug_i > Perc.90] * BlueCollar$	0.35***	0.01	0.34***
	(0.1)	(0.03)	(0.1)
$I[Aug_i > Perc.90] * WhiteCollar$	0.56**	0.32***	0.23
	(0.2)	(0.03)	(0.2)
$I[Aut_i > Perc.90] * BlueCollar$	0.00	0.14	-0.13
	(0.2)	(0.1)	(0.1)
$I[Aut_i > Perc.90] * WhiteCollar$	-0.03	0.05	-0.07
	(0.2)	(0.2)	(0.1)

Table 7: Caption

Notes: WLS weighted by industry-occupation employment in 2005. The specification includes fixed effects by 2-digit industry and occupation code. Standard errors clustered by first-digit industry and occupation codes. Occupation classification according to the first digit of the SOC CODE.

The estimated parameters for these more flexible models are shown in table 7. The estimates indicate that the positive association between labor outcomes and augmentation is verified for the two categories of occupations. The link is stronger in the case of white collar occupations. Also, confirming previous findings, automation metrics are not associated with any indicator of labor outcomes.

5 Concluding remarks

This work proposes a novel metric that characterizes new technologies in terms of the connection with labor. More in detail, we computed augmentation and automation metric extracting information from business narratives. We document large asymmetries in the importance of augmentation and automation across different sectors of the economy. Also, our estimates confirm that labor automation is more prominent than labor augmentation. Finally, estimating models in which these metrics explain three labor outcomes, we observe that augmentation is positively associated with employment and wage growth. On the other hand, there is no significant association between any of the three labor outcomes and the level of automation.

There are multiple directions in which the current exercise can be extended. One aspect that can be evaluated in more detail is given by the information extraction techniques. Text data is a rich and complex source of information. We can consider variations in the specifications of the classification task and the aggregation methods. These variations can result in gains in the precision of the metrics that characterize the implication of new technologies for labor.

A second direction to consider is the expansion of the data under analysis. We could incorporate complementary text data such as narratives from financial statements, trade publications, and patents' documentation. Labor outcome from additional waves of BLS's Occupational Employment and Wage Statistics could allow for some evidence on dynamics during the period under analysis. Finally, another potentially fruitful extended exercise is the implementation of comparative analysis of the metrics presented in this study and other indicators of augmentation and automation.

Acknowledgements

We would like to thank Jeremías Schneider, Sebastian Manzi and Agustín Río, for very valuable research assistance. We benefited from useful comments received at CAF and UCA seminars. We also thank Virginia Robano, Mariano Rabassa and Agustín Coll for valuable discussions.

References

Acemoglu, D., Anderson, G., Beede, D.N., Buffington, C., Childress, E.E., Dinlersoz, E.M., ... others (2022). *Automation and the workforce: A firm-level view from the 2019 annual business survey* (Vol. 30659). National Bureau of Economic Research.

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. *Handbook of labor economics* (Vol. 4, pp. 1043–1171). Elsevier.

Acemoglu, D., Kong, F., Restrepo, P. (2024). Tasks at work: comparative advantage, technology and labor demand.

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

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

Aghion, P., Antonin, C., Bunel, S., Jaravel, X. (2020). What are the labor and product market effects of automation? new evidence from france.

Aghion, P., Antonin, C., Bunel, S., Jaravel, X. (2023). The effects of automation on labor demand. *Robots and AI*, 15–39, Agrawal, A., Gans, J.S., Goldfarb, A. (2019). Artificial intelligence: the ambiguous labor market impact of automating prediction. *Journal of Economic Perspectives*, *33*(2), 31–50,

Agrawal, A.K., Gans, J.S., Goldfarb, A. (2023). *The turing transformation: Artificial intelligence, intelligence augmentation, and skill premiums* (Tech. Rep.). National Bureau of Economic Research.

Aromi, J.D., Bonel, M., Llada, M., et al. (2024). Listening to the price-setters: Inferring inflation expectations from synthetic surveys. *Available at SSRN 4955056*, ,

Autor, D. (2022). *The labor market impacts of technological change: From unbridled enthusiasm to qualified optimism to vast uncertainty* (Tech. Rep.). National Bureau of Economic Research.

Autor, D., Chin, C., Salomons, A., Seegmiller, B. (2024). New frontiers: The origins and content of new work, 1940–2018. *The Quarterly Journal of Economics*, *139*(3), 1399–1465,

Autor, D., & Salomons, A. (2018). *Is automation labor-displacing? productivity growth, employment, and the labor share* (Tech. Rep.). National Bureau of Economic Research.

Autor, D.H., & Dorn, D. (2009). Inequality and specialization: the growth of low-skill service jobs in the united states. *NBER Working Paper Series*, *15150*, ,

Bloom, N., Hassan, T.A., Kalyani, A., Lerner, J., Tahoun, A. (2021). The diffusion of disruptive technologies.

Brynjolfsson, E., & Hitt, L.M. (2003). Computing productivity: Firm-level evidence. *Review of economics and statistics*, *85*(4), 793–808,

Brynjolfsson, E., Li, D., Raymond, L. (2025). Generative ai at work. *The Quarterly Journal of Economics*, *140*(2), 889–942,

Chava, S., Du, W., Paradkar, N. (2020). More than buzzwords? firms' discussions of emerging technologies in earnings conference calls. *Firms' discussions of emerging technologies in earnings conference calls.*

Dell'Acqua, F., McFowland III, E., Mollick, E.R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., … Lakhani, K.R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of ai on knowledge worker productivity and quality. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*(24-013), ,

Hassan, T.A., Hollander, S., Kalyani, A., van Lent, L., Schwedeler, M., Tahoun, A. (2025). Text as data in economic analysis. *Journal of Economic Perspectives*, *39*(3), 193–220,

Johnson, S., & Acemoglu, D. (2023). *Power and progress: Our thousand-year struggle over technology and prosperity— winners of the 2024 nobel prize for economics.* PublicAffairs.

Leontief, W., & Duchin, F. (1986). *The future impact of automation on workers.* New York: Oxford University Press.

Prettner, K., & Strulik, H. (2020). Innovation, automation, and inequality: Policy challenges in the race against the machine. *Journal of Monetary Economics*, *116*, 249–265,

Tomlinson, K., Jaffe, S., Wang, W., Counts, S., Suri, S. (2025). Working with ai: Measuring the occupational implications of generative ai. *arXiv preprint arXiv:2507.07935*, ,

Appendix A New technology phrases

*-ai, **ai**, ai-*, algorithm*, artificial intelligence, automata*, automated, automating, **automation**, automatiz*, autonomous system*, big data, bot, bots, cashierless, chatbot*, cobot*, cognitive computing, computer vision, data science, deep learning, digital assistant, digital transformation, fourth industrial revolution, gpu*, human-machine, humanless, industry 4.0, **information technology**, machine intelligence, machine learning, recommendation system, **robot***, rpa, self-checkout, self-service, self-serve, smart machine, technological disruption, unmanned, virtual assistant

Appendix B Alternative cluster of broad sectors

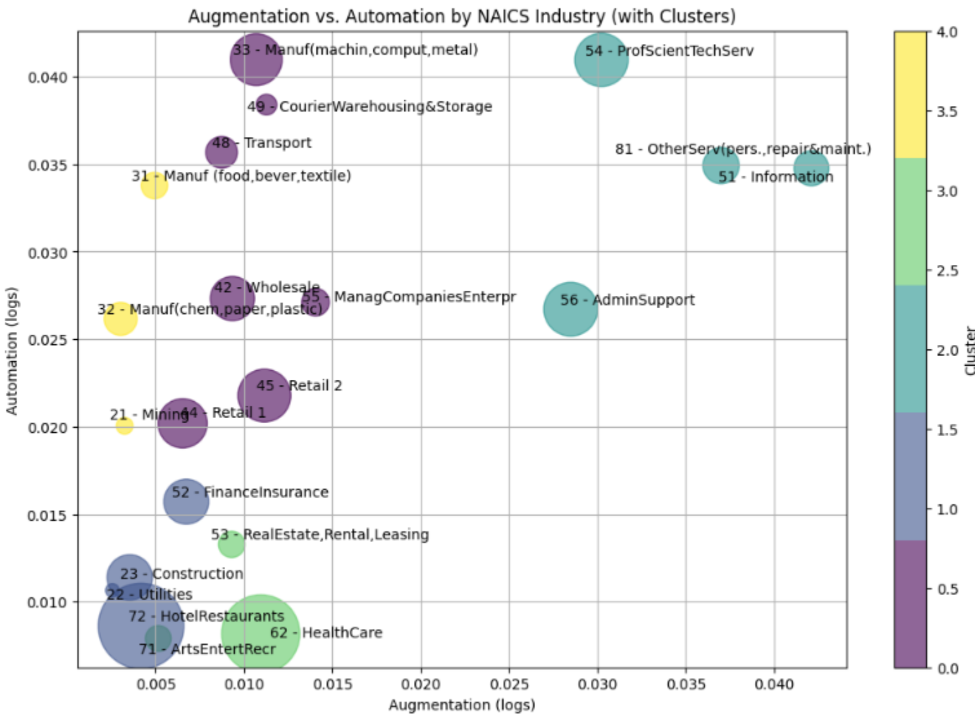

Fig. B1: Augmentation and automation: 2-digit NAICS industries - Clusters by angle and hypothenuse

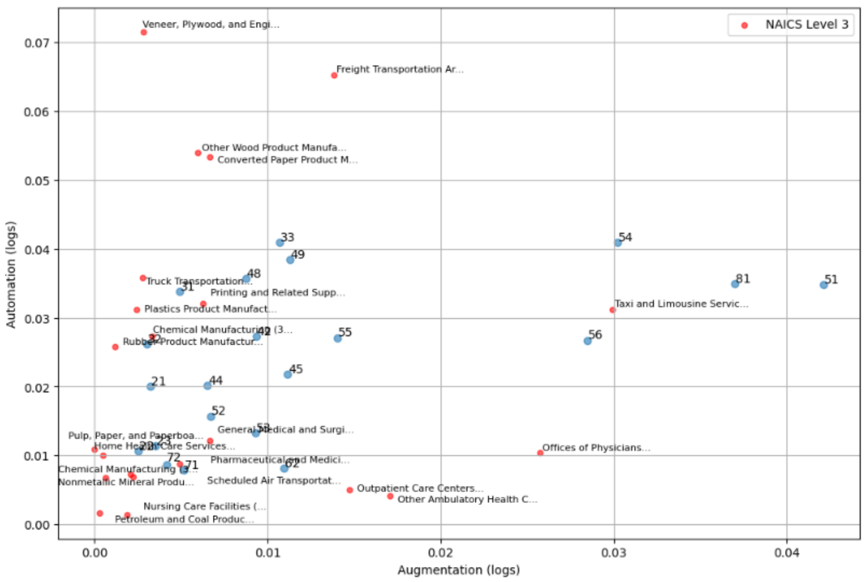

Fig. B2: Industry augmentation and automation: 48 (Transport), 32 (Industry) and 62 (Healthcare)

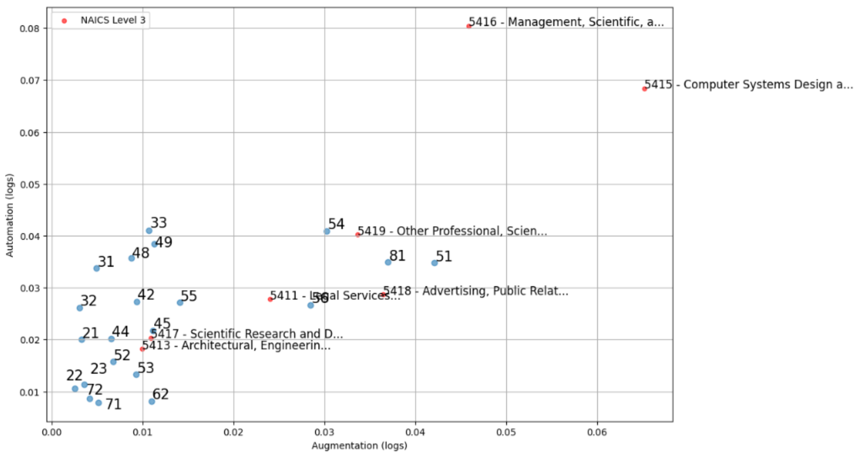

Fig. B3: Industry augmentation and automation: 54 (Professional, Scientific and Technical Services)

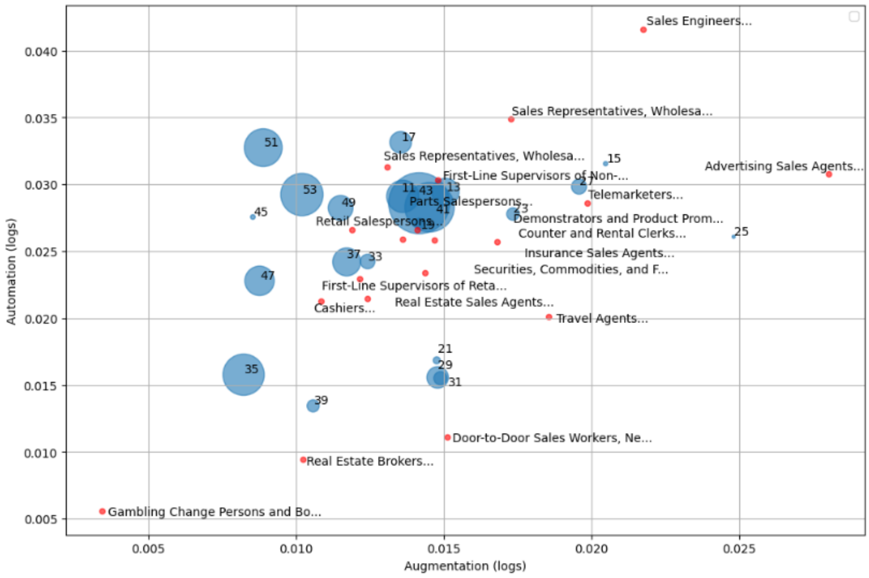

Fig. B4: Augmentation and automation: Sales