socioeconomic disruption by artificial intelligence

A comparative analysis between industries in the European Union

Table of contents

1 Abstract **1**

2 Introduction **2**

2.1 Effects of Automation on labor . . . 2

2.2 Effects of computerization . . . 3

2.3 Effects of AI . . . 3

2.4 Changes of Occupational composition (no net displacement but need to re-skill) . . . 4

2.5 Changes in labor share . . . 4

2.6 Summary of the different findings -> why findings differ . . . 5

2.7 Definitions of AI . . . 5

3 Methodology **6**

3.1 Data Sources . . . 6

3.2 Data Acquisition . . . 6

3.3 Preprocessing . . . 7

3.4 Hypothesis . . . 11

4 Results **13**

5 Discussion **18**

5.1 Implications . . . 18

5.2 Limitations . . . 18

5.2.1 Small data set . . . 18

5.2.2 Patent weighting . . . 18

5.2.3 Additional CPC classes . . . 18

5.2.4 Further control variables . . . 18

5.3 Further research . . . 18

6 Conclusion **19**

References **20**

Appendix **23**

List of Figures

1 Example of untransformed data for all Industries and NACE Code ‘Number of Employees’ plotted over years 10
2 Example of linear detrended data for all Industries and NACE Code ‘Number of Employees’ plotted over years 11

List of Tables

1 Selected CPC Codes 6
2 Regression results - Mining and Quarrying (B) 13
3 Regression results - Manufacturing (C) 13
4 Regression results - Electricity, gas, steam and air conditioning supply (D) 14
5 Regression results - Construction (F) 14
6 Regression results - Transportation and storage (H) 15
7 Regression results - Information and communication (J) 15
8 Retrieved significant p values for coefficients of number of patents by industry and indicator 16
9 Significant p values with coefficient sign, sample size and total number of patents 18
10 Selected main keywords for NACE industries 23
11 Example queries posted to the OPS API 24

1 Abstract

2 Introduction

Mokyr et al. (2015, p. 32) identifies two forms of technological anxiety, the fear of labor displacement through technology and and the fear of morally negative applications resulting in declining welfare. The majority of the US population has been found to assess the potential impact of automation as unfavorable rather than beneficial (Anderson, 2017).

Since AI is a still fairly new topic in the literature and has only seen real increase in dominance and interest in recent years (Acemoglu et al., 2020a, p. 23f.), is is worth noting the effects of previous technologies as the adoption of machines (specifically often industrial robots (Acemoglu and Restrepo, 2020a; see for example **?**)) and software (also referred to as computerization(Autor and Dorn, 2013; Frey and Osborne, 2017; see for example Pajarinen et al., 2015)) have been seen as previous stages in the evolution of automation with AI composing the next stage (Acemoglu, 2021, p. 19). Furthermore, these technologies have been summarized under the umbrella term "automation" (Mann and Püttmann, 2018, p. 40) indicating common characteristics and thereby - possibly - common effects.

2.1 Effects of Automation on labor

In a 2018 study, the introduction of automation technology was found to have positive effects on employment gains, but only within the same commuting zone (Mann and Püttmann, 2018, p. 26). These findings contradict the results from Autor et al. (2015) [p. 632], that found no relation between exposure to automation and employment as a whole but found a significant decline in employment related to routine tasks in the non-manufacturing sector (p. 641). (**?**) found no relationship between the usage of industrial robots and net employment. However, usage of industrial robots was found to lower employment of low-skilled workers. However, a later study also looking at employment effects induced by usage of industrial robots found a significant decline of employment as well as a reduction in wages related to robot exposure within a commuting zone (Acemoglu and Restrepo, 2020a, pp. 2215f, 2218). Dauth et al. (2017) [p. 25] found no relation between robot exposure and employment in the German market. A few years later, Dauth et al. (2021, p. 3126ff) found robot exposure to lead to within-firm and between-firm job displacement, with displaced workers having difficulties reallocating their jobs within the same industry, leading to a migration of workers from manufacturing (where robot exposure is most present) to the service sector. They also exhibited that a lack of worker protections (for example unionization or tenure) is related to greater displacement. These results were also confirmed by Boustan et al. (2022) [p. 21, 23] who observed that displaced workers acquire new skills and concluded job displacement by automation to be less discernible amongst unionized and high-skilled workers. Similarily, Acemoglu and Restrepo (2020a, p. 2215f., 2218) provided evidence showing automation (adoption of industrial robots) within a commuting zone (local labor market) relating to significant declines in employment as well as wages. By studying 53 developing countries, Cirera and Sabetti (2019, p. 172) did not find a relationship between exposure to automation and firm level employment. Hoewever, while a net effect on employment was absent, in line with the aforementioned literature, they did find automation to alter the composition of tasks and skills within firms (p. 172).

In a purely theoretical approach to the effects of automation on labor, Acemoglu and Restrepo (2017) [p. 12, 15] concluded that automation leads to labor displacement and the displacement of low skilled-labor leading to an increase in the wage gap (pay gap between low-skilled and high-skilled workers) while the displacement of high-skilled labor is followed by a reduction in the wage gap as high-skill labor reallocates into medium- and low-skilled occupations. This reallocation from displaced high-skill labor into lower skilled occupations has also been shown by Beaudry et al. (2016, p. 21) who studied the effects on labor when prices for specific types of labor fall - as is induced when substitution (through technology) becomes economically viable. While labor displacement induced by the introduction of automation is followed by increased inequality between low-skill and high-skilled labor in the short run (Acemoglu and Restrepo, 2018a, p. 1519), the creation of new tasks - that is followed by increased productivity gains from automation - is seen to reduce this gap in the long run (p. 1521). However, this positive outlook of a net positive on employment only holds true as long as the productivity effects which accompany the adoption of automation technologies offset the displacement effects incurred in the first place - and should the offset be insufficient, automation is found to negatively impact the demand for labor and its wages (Acemoglu and Restrepo, 2018b, p. 227). There is also growing evidence suggesting automation to cause a decline in real wages of low-skilled workers, for example Acemoglu and Restrepo (2020b) [p. 360f.] found strong relationships between the adoption of automation technology and wages. Acemoglu and Restrepo (2022, p. 1993) found a relationship between labor displacement and a decrease in relative wages, concluding automation to cause an increase in wage inequality (p. 1998). Automation is also attributed to the decline in the demand for labor in the US over recent decades (Acemoglu and Restrepo, 2019, p. 21).

Furthermore, Arntz et al. (2016) [p. 14f.] studying 21 OECD countries found 9% in the US, and over all countries studies a 6-12% high risk of employment to be substitutable for automation, while (**?**) came to the conclusion that labor displacement by machines mostly affects routine tasks.

2.2 Effects of computerization

In a study from Finland, Pajarinen et al. (2015) found that computerization is likely to place high risk of displacement on 35% of the Finish labor market [p. 5], 33% of Norwegian labor (p. 5) as well as 49% in the US [p. 5]. Frey and Osborne (2017, p. 41) found 47% of US employment to have a a high risk suitability for substitution by computerization. They further classify the process of automation into two “waves” with the first wave affecting routine tasks (transportation, logistics, office, and administration) [p. 41] followed by a second wave that, once technological obstacles are overcome, will effect the jobs involving creative or abstract tasks [p. 43]. Evidence also suggests computerization to significantly induce labor displacement from occupations relying on routine tasks into higher-skilled occupations as well as low-skilled service occupations (Autor and Dorn, 2013, p. 1573)

2.3 Effects of AI

Brynjolfsson et al. (2018) [p. 46] found that machine learning affects different types of tasks than earlier forms of automation. A year later, in a study comparing the impact of AI on the job market between industries, Webb (2019, p. 46) shows that AI affects mostly the highly educated workforce and that this group is affected significantly more by AI than the presence of software or robots. Under the assumption that the current trend in technological evolution is set to continue, the speed of labor displacement through technological innovation is found likely to outpace the speed at which labor can be relocated (Mokyr et al., 2015, p. 43f.). By constructing impact scores of Artificial Intelligence on occupations, Felten et al. (2019) found low-income occupations to experience a decline in wage growth that is attributed to the increased presence of AI and middle and high-income occupations to experience an increase in wage growth [p. 6]. Furthermore, the authors found found that occupations with a medium and high degree of automation (degree of automation being the presence of automation technologies - not just AI) positively correlate with employment when exposed to Artificial Intelligence, while they did not find any relationship for occupations already exhibiting a low degree of automation [p. 5].

It has also been noted that the presence of Artificial Intelligence does not have a linear impact on labor but depends on influencing factors, such as price elasticity, complementarities, or elasticity of labor that govern the implementation of these technologies (Brynjolfsson and Mitchell, 2017, p. 1533f.). Additionally, the adoption of AI technology is found to significantly alter the skill-demand distribution of firms, with the number of previously highly demanded skills declining while simultaneously creating demand for new skills (Acemoglu et al., 2020a, p. 19).

2.4 Changes of Occupational composition (no net displacement but need to re-skill)

Furthermore, it is important to note that previous research on the effects of robots, software and AI - that have been summarized under the umbrella term “automation” (Mann and Püttmann, 2018, p. 40) - in general may not have found net negative effects on employment but a restructuring of composition of occupations. The aforementioned study from Autor et al. (2015) [p. 644] found automation, while having no aggregate effects on employment, lead to a decline in occupations involving routine tasks and and an increase in non-routine (abstract) tasks. The same effect was found in (**?**) studying the introduction of industrial robots.
These effects remain only harmless as long as the assumption holds true that displaced labor can in fact always reallocate itself to new tasks. Should this assumption be contradicted, and the the negative effects of automation on employment are no longer offset by the positive effects of reallocation, the phenomenon of occupational migration would turn into an observation of job destruction.

2.5 Changes in labor share

The introduction of capital, whether to complement or substitute labor, intuitively leads to a decline of a firms profits paid to labor as the share of labors input relative to the output value decreases. And in fact Karabarbounis and Neiman (2014) [p. 99] show that the observed decline in capital prices explains almost half the decline in global labor share, that has been observed in recent decades. This might seem problematic as an increasing portion of a firms revenue remains as corporate profits and savings (given that the capital invested leads to a decrease in marginal costs - through substitution of labor and/ or increased production) rather than being redistributed to labor. Karabarbounis and Neiman (2014, p. 102) further show that the observed decline in labor share is accompanied by an increase in corporate revenue and savings. This is also brought forward from Acemoglu and Restrepo (2019) [p. 27] who conclude that "[...] automation always reduces the labor share and may reduce labor demand [...]" but also mention that the creation of new tasks necessarily increases the labor share. These results where further solidified by Acemoglu et al. (2020b, p. 387) who investigated the French manufacturing market and found firms exposed to automation (in this study measured by the introduction of robots) to experience significant declines in their labor share.

2.6 Summary of the different findings -> why findings differ

The net impact assessment of automation on socioeconomic factors widely differs in the aforementioned literature (see also Frank et al., 2019, p. 6532). Some research has focused on local labor markets (commuting zones) (see Acemoglu and Restrepo, 2020a; Autor et al., 2015; Autor and Dorn, 2013), while other research has researched national effects [see xxx] and international effects (see Graetz and Michaels, 2018). While one would expect to see the same relationship between the chosen variables on all levels and apart from differences in research design, it may be difficult to assess effects on a greater aggregate level as the number of variables that would need to be included to account for differences between and within groups becomes unfeasible.

2.7 Definitions of AI

The classification if Artificial Intelligence remains also difficult due to the fact that there is yet no widespread agreement on the definition of intelligence itself (Legg and Hutter, 2007).

Given the various contradicting results on the relationship between automation and labor effects and the increasing presence of AI, this research aims to add to the current corpus of literature by assessing the relationship between AI innovation and socioeconomic factors. Specifically, the research question is as follows: How does AI innovation across industries impact labor?

3 Methodology

The methodological approach has similarities to Mann and Püttmann (2018, p. 13) who used patent counts as a proxy for estimating the level of automation present within a US commuting zone. However, the method of selecting patents differs. While Mann and Püttmann (2018) classified texts based on the tasks they may effect within occupations, the presented approach here uses API query composition to preselect patents whose title or abstract match keywords reserved to an industry.

3.1 Data Sources

Data about patent publications is obtained from the European Patent Office's Open Patent Services (OPS) API (European Patent Office, 2023) as well as the Annual Structural Business Statistics (SBS) by Eurostat (Eurostat, 2023a). Furthermore, Eurostats code lists of Statistical classification of economic activities in the European Community (NACE Revision 2) (Eurostat, 2023b) (henceforth "NACE") and Economic Indicators for Eurostat's SBS (Eurostat, 2023c). While there are a variety of possible technologies that may fall under the umbrella term "Artificial Intelligence", as this research aims to assess AI's socioeconomic impact - which, if negative, falls into the governmental realm - a legal definition of AI is preferable as a classifier. Furthermore, it is arguable that the political definition is likely to have the greatest (socio)economic impact in the near future due to possible (and probable) regulation. As there is no legal definition yet - at least in the EU - technologies listed in the European Commisions latest proposal for the "Artificial Intelligence Act['s]" (European Commission, 2021) annex (European Commission, 202AD) will be used.[1]

Additionally Cooperative Patent Classification (CPC) codes are used to retrieve patents that utilize artificial intelligence technology. As there is no clear mapping between the European Commission's definition and CPC codes, classifications are chosen to the author's best knowledge.

Table 1: *Selected CPC Codes*

Class	CPC
Machine Learning	G06N20/00, G06N20/10, G06N20/20
Supervised Learning	G06N3/09
Unsupervised Learning	G06N3/088
Reinforcement Learning	G06N3/092
Deep Learning	G06N3/08

3.2 Data Acquisition

In order to retrieve data from the European Patent Office's Open Patent Services (OPS) API, queries were composed to link retrieved patents to their respective industry. The query composition is based on the

[1]The European Commission's proposal for the "Artificial Intelligence Act" is currently in the legislative process. At the time of writing, the European Parliament has made amendments to this proposal, one of which - unfortunately - is the removal of the list of technologies classified as AI from the initial proposal's annex (**?**(202023, p. 326f.). For the time being, the AI Act now lacks a clear defintion which is why the European Comission's initial proposal's definition will be used.

selected CPC codes displayed in Table 1 as well as keywords from the list of NACE codes that have been retrieved from Eurostat. Each NACE code is composed of section (alphabetical), division (numerical), group (numerical) and class (numerical) of a particular economic activity. Sections relate to the overall industry, while divisions, groups and classes relate to more specific activities within the industry (**?**). For each industry, keywords are extracted from the NACE code's description. To ensure only relevant keywords are used, each description is cleaned of common characters and unrelated words (e.g., ",", "and", "or", "to"). Descriptions for each industry are then split into lists of single keywords that will be used in the API query.

Because some industries contain a variety of different activities (e.g., NACE industry (section) "A" relates to "Agriculture, forestry and fishing" (Eurostat, 2023c)), main keywords that relate to the section as a whole are manually selected (see Table 10 in the Appendix). For each industry and main keyword, queries are then build using the (manuallly selected) main keyword, the description keywords, as well as the chosen CPC codes. The resulting query is then used to retrieve patents from the OPS API. To limit query results to the relevant market, only patents filed with the European Patent Office (EP) will be retrieved. This approach disregards patents filed with national patent offices. However, since the query language is english, and many national patents are only filed in their native language, retrieving patents for all EU markets was unfeasible. The query is composed of the following elements:

> **(ta = Main Keyword) AND (ta = ANY Description Keywords) AND (cpc ANY CPC Codes) AND (ap = "EP")**
> *Note: ta = title or abstract; ap = Application Number, referring to the Patent Office the patent was filed at. In this case, "EP" refers to the European Patent Office. See Table 11 for example queries*

The queries are then posted to the OPS API's Published Data Keywords Search with Variable Constituents endpoint (European Patent Office, n.d.) and data from the responses - which are provided in JSON format - extracted. Initially, queries were created not only for the European Patent Office but all patent offices within the European Union to retrieve patent data on a national level. This approach would have resulted in a much richer dataset, enabled better aggregates while also allowing for between-country comparissons. However, initial tests showed that most of the patents filed with a national patent office contain only patent titles and absracts in their native language, which renders the chosen keywords in the query language (english) ineffective. As a result, the decision was made to only retrieve patents filed with the European Patent Office.

3.3 Preprocessing

Since Eurostat's SBS data only includes codes to refer to given indicators as well as industries, data retrieved from Eurostat (SBS, and Code Lists about NACE and SBS codes) is merged. This is done by matching the NACE codes and SBS indicator codes to the respective NACE code and indicator in the SBS data. The economic indicators "Enterprises" and "Persons employed" are reported as totals. "Wage adjusted labour productivity (Apparent labour productivity by average personnel costs)" and "Share of personnel costs in production" are reported as percentages, and "Gross value added per employee" is reported in Euros.

Because the number of employees is rather large for each industry, the number of employees is divided by 1000 to reduce the scale of the data. This increases readability of tables in the following regression results while also still large enough that it is unlikely for coefficients (coeff.) and standard errors (SE) to fall too far into the decimals.[2]

Next, patent data retrieved from the OPS API, which returns data in JSON format, is converted into a pandas DataFrame (The pandas development team, 2023) (i.e., a table). As multiple queries for the same industry - but with different keywords - have been posted to the API, duplicates in the patent data are removed. Specifically, duplicate patent data (indicated by the patent application number) are removed in each industry subset of the data. This ensures that each industry only contains unique patents while patents can still appear in more than one industry (as their applicable usage may not be restricted to only one industry). Furthermore, as the SBS data only spans from 2011 to 2020, patents that have been filed before or after this period are removed from the data. As a next step, patents are grouped by their respective industry and year of application and the patent count for each subgroup is recorded. Furthermore, industries for which patents have been retrieved in less than four years within 2011-2020 are removed from the data to ensure a minimum sample size for the following statistics. The sum of patents for each industry and year composes the exogenous variable “Sum patents” that will be used in the regression analyses.

Furthermore, the SBS data is merged with the patent data by matching the industry and year of application with the industry and year of the SBS data. This ensures that each industry and year combination in the SBS data has a corresponding patent count. Furthermore, the data is once more grouped for each industry to retrieve the earliest and latest year for which patent counts are available. For each industry, SBS data is removed for the years before and after the first and last patent retrieval for the respective industry. This is done to ensure that the regression analyses are only conducted for years in which patent counts are available.[3] However, in some cases, patents were discontinuously retrieved for industries. In other words, if patents are retrieved for an industry in 2016, 2018, 2019, and 2020, but not in 2017, the SBS data for 2017 and the respective do not have a corresponding patent count. In order to account for missing values within a series of definite patent retrieval, the patent count for the missing year is set to zero. This is done for each industry and year combination in which patent counts are missing.

Lastly, in some rare cases, SBS data is missing for a given year and industry. In these cases, rows of the respective year and industry are removed from the data. This is done to ensure that the regression analyses are only conducted for years in which SBS data is available. The resulting data is then used for the regression analyses. In summary, data for each year and industry will be used further if the following conditions are

[2] Note that this is done to ensure readability and does not affect the regression results. Defactoring data by more than a thousand might lead to coefficients and standard errors falling into the decimals, which in turn may show up - due to rounding - as zeros despite having large scale effects (when rescaled with the original factor).

[3] There are valid arguments to be made for and against excluding these data. For once, the lack of patent retrieval for any given year implies no patent filing within that year, making null values a good control instance to check for variation in SBS data that is definitely not affected by patent filings. On the other hand, for a few industries, this would result in many null values, giving the data series of patent counts a definite trend. Furthermore, patent counts have also been removed for years in which SBS data is unavailable. To reduce potential bias produced by imputing and keeping the data’s integrity, removing the missing values has been chosen over the data’s accuracy.

met.

1. Patents have been retrieved for the industry in at least four years within 2011-2020
2. Patents have been retrieved for this or an earlier year
3. Patents have been retrieved for this or a later year
4. SBS data is available for this year and industry

The resulting data contains 211 data points across 6 Industries, each with 5 economic indicators. However, given the relatively short time period in which data could have been collected, paired with the fact that the retrieved patents are aggregated for each year, the resulting data size for each industry and economic indicator is relative small. The average number of years in which patent counts have been recorded -according to the methods above- is only 7 years, ranging from a minimum of 4 years up to 10 years. Since each year per industry and indicator will be used as a data point in the following regression analyses, it is neccesary to note that results may be biased due to the small sample size. Furthermore, given the small dataset -which makes diminishes the accuracy with which a regression can be fitted (i.e., fewer "anchor points"), assumptions about the extend to which patent counts affect the chosen economic indicators will not be made. Instead, the regression analyses will be used to assess whether a relationship between the number of patents and the chosen economic indicators exists at all. That is, the interest lies whether AI patent counts yield any explanatory power over the chosen economic indicators.

Because the collected data comrpises a time-series, as each industry's patent application counts as well as the SBS data have been retrieved for multiple years. As shown in Figure 1, the collected data on SBS indicators (blue) as well as the number of patents retrieved each year (red) clearly does not exhibit stationarity. In order to account for any trends in the data, the collected data is transformed using linear detrending method. This is done by utilicing scipy's detrending method (Virtanen et al., 2020), which fits a linear least-squares regression to the data and subtracts the resulting trend of the regression line from the data (The SciPy community, 2023). Note that other detrending options, such as logarithmic transformation or differencing have been considered but deemed insifficient. Logarithmic transformation is not applicable as the data contains zero values. While there are methods to circumvent this, for example taking the logarithm $log(x+1)$, this would lead to non-null values where null values are expected to control for variance in the endogenous variable in the absence of patent counts. Furthermore, as seen in Figure 1, many data series exhibit a continues positive or negative trend (a lack of fluctuation). In this case, differencing would merely reverse the trend, and logarithmic detrending would lead to a compression of the y-scale. Resulting data transformed by either of these methods, however, would still exhibit a definite trend. The resulting data, of which an example is shown in Figure 2, is then used for the regression analyses.

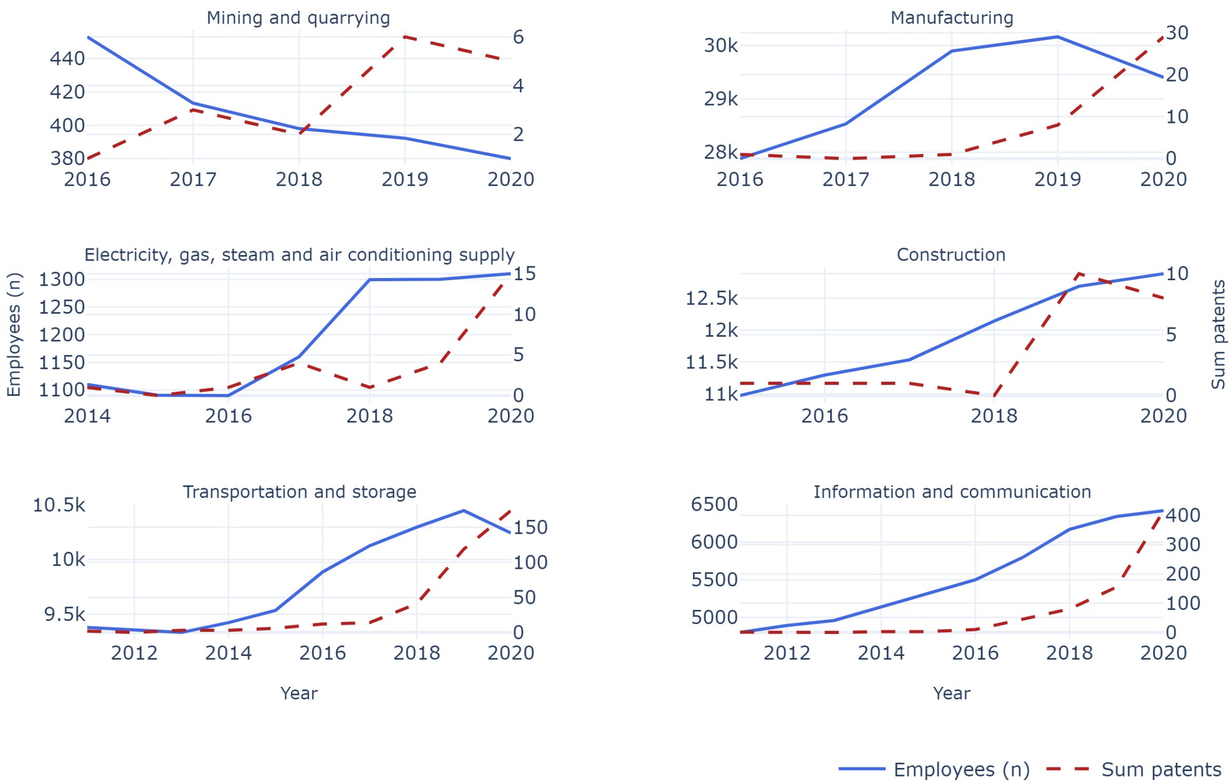

Figure 1: *Example of untransformed data for all Industries and NACE Code 'Number of Employees' plotted over years*

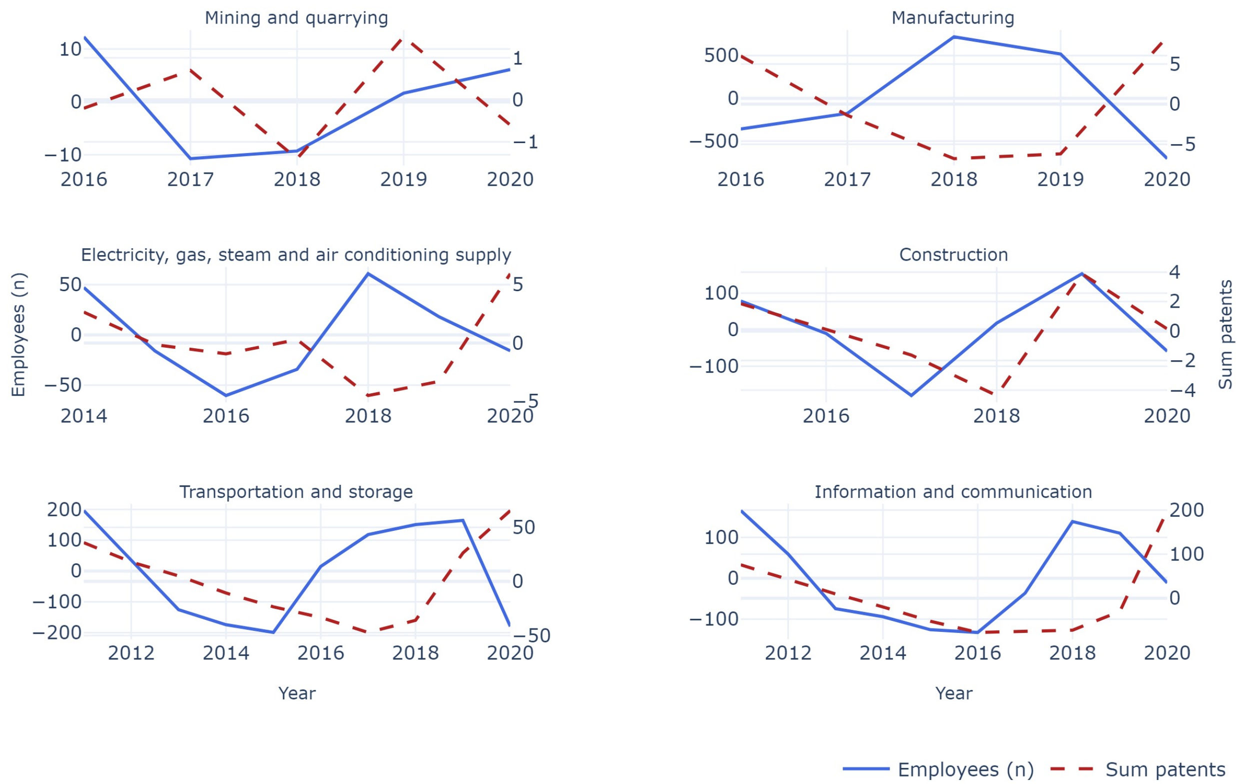

Figure 2: *Example of linear detrended data for all Industries and NACE Code 'Number of Employees' plotted over years*

Because data has been linearily detrended, to account for any remaining trend left in the data, the control variable "Year" is added to the regression analyses. This is done to ensure that any remaining trend in the data is accounted for and does not bias the regression results. Furthermore, the control variable "Year" is also added to the regression analyses to account for any time dependent macroeconomic effects that may have affected the chosen economic indicators but are not considered in the model.

3.4 Hypothesis

Do determine whether a relationship between the number of patents and the chosen economic indicators exists, the following hypotheses are tested. Given a standard multiple linear regression model of the form $y_{i,j} = \beta_0 + \beta_1 x_i + \beta_2 x_t$, where $i =$ industry, $j =$ economic indicator and $t =$ time, the coefficient β_1 is assumed to be 0. Specifically, the following assumptions are tested.

$$H_{0,i,j} : \beta_1 = 0 \text{ for } j = e = \text{number of enterprises} \tag{1}$$

$$H_{0,i,j} : \beta_1 = 0 \text{ for } j = L = \text{number of employees} \tag{2}$$

$$H_{0,i,j} : \beta_1 = 0 \text{ for } j = l = \text{wage adjusted labor productivity} \tag{3}$$

$$H_{0,i,j} : \beta_1 = 0 \text{ for } j = v = \text{gross value added per employee} \tag{4}$$

$$H_{0,i,j} : \beta_1 = 0 \text{ for } j = c = \text{personnel costs in production} \tag{5}$$

4 Results

The following section presents the main findings from the regression analyses. Results are summarized by industry, allowing a sectional comparrison of patent counts' influence on economic indicators within an industry.

Table 2: *Regression results - Mining and Quarrying (B)*

	Enterprises (n)	Employees (n)	Labor prod. (%)	GVA/employee (€)	Personnel costs (%)
const	0.0000	0.0000	-0.0000	-3878292.0523	0.0000
	(119814.6603)	(8940.2460)	(32619.9298)	(20665152.7595)	(998.5155)
Sum patents	-88.6471	0.3627	-8.1451	-1340.9907	0.6451
	(83.1388)	(6.2036)	(22.6348)	(14778.9277)	(0.6929)
Year	-0.0000	-0.0000	0.0000	1921.6113	-0.0000
	(59.3730)	(4.4302)	(16.1645)	(10239.1414)	(0.4948)
R-squared	0.3624	0.0017	0.0608	0.0411	0.3024
R-squared Adj.	-0.2751	-0.9966	-0.8784	-1.8766	-0.3952

** p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in paranthesis*

For patents classified as industry "Mining and Quarrying" (NACE code "B"), depicted in Table 2, the regression results show no significant relation between the sum of patents retrieved for each year and the chosen indicators. Furthermore, the control variable "Year", too, does not exhibit any significant relationships with the economic indicators. It should be noted, howver, that the number of patents retrieved for this industry is very low. While, as discussed in the Data Acquisition, industries for which patents were retrieved in fewer than five years were eliminated from the data, for Mining and Quarrying only 17.0 patents in 5 years were retrieved. As a result, the nullhypotheses $H_{0,i,j}$ for $j \in \{e, L, l, v, c\}, i = B$ are not rejected.

Table 3: *Regression results - Manufacturing (C)*

	Enterprises (n)	Employees (n)	Labor prod. (%)	GVA/employee (€)	Personnel costs (%)
const	0.0000	0.0000	-0.0000	-0.0000	0.0000
	(15775939.2113)	(161822.0105)	(554.4693)	(306669.6590)	(133.4932)
Sum patents	-14.6066	-82.4888**	-0.1300	-203.3126**	0.0400
	(1778.5677)	(18.2437)	(0.0625)	(34.5737)	(0.0150)
Year	-0.0000	-0.0000	0.0000	0.0000	-0.0000
	(7817.6092)	(80.1893)	(0.2748)	(151.9671)	(0.0662)
R-squared	0.0000	0.9109	0.6839	0.9453	0.7790
R-squared Adj.	-0.9999	0.8218	0.3677	0.8907	0.5580

For patents classified as industry "Manufacturing" (NACE code "C"), depicted in Table 3, the regression results show a statistically significant negative relationship between the number of retrieved patents and the number of employees within the Manufacturing sector (coeff. -82.488.8, SE 18.243'). In particular, the regression result's coefficient estimates a decrease of 82489 employees for each additional patent retrieved[4]. Furthermore, the control variable "Year" does not exhibit a statistically significant relationship with the number of employees (coeff. 0). The adjusted R^2 of 0.82 indicates a high ratio of explainability for the model.

While there are no statistically significant relations between the number of patents retrieved and the number of enterprises, wage adjusted labor productivity (labor prod.) and the percentage of personnel costs in production, the relationship between the number of patents and the gross value added per employee is statistically significant and negative (coeff. -203.313, SE 34.574) with an adjustes R^2 of 0.89. Lastly, it should be noted that the control variable does not exhibit a statistically significant relationship with any of the economic indicators. In summary, As a result, the nullhypotheses $H_{0,i,j}$ for $j \in \{L, v\}, i = C$ are rejected and $H_{0,i,j}$ for $j \in \{e, l, c\}, i = C$ cannot be rejected.

Table 4: *Regression results - Electricity, gas, steam and air conditioning supply (D)*

	Enterprises (n)	Employees (n)	Labor prod. (%)	GVA/employee (€)	Personnel costs (%)
const	0.0000	0.0000	-0.0000	0.0000	0.0000
	(7090398.8328)	(19752.8953)	(2819.3693)	(1493538.8451)	(103.5974)
Sum patents	-2148.9258	-3.4280	1.4361	581.8882	0.0439
	(2160.2744)	(6.0182)	(0.8590)	(455.0454)	(0.0164)
Year	-0.0000	-0.0000	0.0000	-0.0000	-0.0000
	(3515.3175)	(9.7932)	(1.3978)	(740.4750)	(0.0513)
R-squared	0.1983	0.0750	0.4113	0.2902	0.8780
R-squared Adj.	-0.2025	-0.3875	0.1170	-0.0647	0.6341

Table 5: *Regression results - Construction (F)*

	Enterprises (n)	Employees (n)	Labor prod. (%)	GVA/employee (€)	Personnel costs (%)
const	-0.0000	-0.0000	-0.0000	0.0000	0.0000
	(21172494.6121)	(58291.4313)	(659.6442)	(476813.8504)	(117.8279)
Sum patents	9252.7110	23.4435	-0.2356	16.5014	0.0191
	(6911.8456)	(19.0295)	(0.2153)	(155.6578)	(0.0385)
Year	0.0000	0.0000	0.0000	-0.0000	-0.0000
	(10494.4174)	(28.8929)	(0.3270)	(236.3389)	(0.0584)
R-squared	0.3740	0.3359	0.2851	0.0037	0.0756
R-squared Adj.	-0.0434	-0.1068	-0.1915	-0.6604	-0.5407

For patents classified as industry "Electricity, gas, steam and air conditioning supply" (NACE code "D"), depicted in Table 4, as well as for patents falling into the "Construction" ("F") industry Table 5 the regression results show no statistically significant relationship between the number of patents retrieved and the

[4]Note that while the coefficient's implications are mentioned, this merely refers to the slope of the regression line and should not be interpreted as valid result with real-world implications. The regression model is not intended to be used for prediction.

chosen economic indicators. The control variable "Year", too, does not exhibit a statistically significant relationship with the chosen indicators. Furthermore, the adjusted R^2 is very low (and often even negative) across all dependent variables, indicating no explanatory power of the model. Therefore, the nullhypotheses $H_{0,i,j}$ for $j \in \{e, L, lv, c\}, i \in \{D, F\}$ cannot be rejected.

Table 6: *Regression results - Transportation and storage (H)*

	Enterprises (n)	Employees (n)	Labor prod. (%)	GVA/employee (€)	Personnel costs (%)
const	0.0000	0.0000	0.0000	0.0000	0.0000
	(3846872.7673)	(39250.0211)	(706.0715)	(376987.6682)	(124.3979)
Sum patents	594.7790***	-0.4540	-0.1021**	-33.0649*	0.0156**
	(159.0156)	(1.6225)	(0.0292)	(15.5833)	(0.0051)
Year	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
	(1908.6425)	(19.4741)	(0.3503)	(187.0441)	(0.0617)
R-squared	0.6665	0.0111	0.6362	0.3914	0.5673
R-squared Adj.	0.5712	-0.2715	0.5322	0.2175	0.4436

The regression models between number of patents allocated to the transportation and storage industry (H) and the chosen endogenous variables, depicted in Table 6, show a number of statistically significant relationships. First, the number of filed patents is statistically significant in predicting the number of Enterprises present in any given year. The coefficient of 94.78 (SE 159.016) implies a positive relationship between the number of AI patents and the number of Enterprises. The control variable remains statistically insignificant. This holds also true for the remaining indicators modelled within the transportation and storage industry. The adjusted R^2 of 0.57 indicates that over 50% of the predictors' variance is explained by the model. No statistically significant relationship can be reported between the industries retrieved annual patent counts and the number of employees and gross value added per employee. However, wage adjusted labor productivity exhibits a statistically negative relationship to increasing patent AI patent filings with a coefficient of -0.102 and a standard error of 0.029 (Adj. R^2 0.532). The same relationship occurs for the percentage of personnel costs in production is found to be significantly positively related to the number of patents filed (coeff. 0.005, SE 0.005). To conclude, hypotheses $H_{0,i,j}$ for $j \in \{e, l, c\}, i = H$ are rejected and $H_{0,i,j}$ for $j \in \{L, v\}, i = H$ cannot be rejected.

Table 7: *Regression results - Information and communication (J)*

	Enterprises (n)	Employees (n)	Labor prod. (%)	GVA/employee (€)	Personnel costs (%)
const	-0.0000	0.0000	0.0000	0.0000	-0.0000
	(2195739.8894)	(27234.7248)	(438.8221)	(225189.9648)	(69.7770)
Sum patents	2.9806	0.2962	0.0325***	22.0818***	-0.0028*
	(37.9587)	(0.4708)	(0.0076)	(3.8930)	(0.0012)
Year	0.0000	-0.0000	-0.0000	-0.0000	0.0000
	(1089.4258)	(13.5126)	(0.2177)	(111.7290)	(0.0346)
R-squared	0.0009	0.0535	0.7242	0.8213	0.4411

	Enterprises (n)	Employees (n)	Labor prod. (%)	GVA/employee (€)	Personnel costs (%)
R-squared Adj.	-0.2846	-0.2169	0.6454	0.7703	0.2815

Lastly, the regression models' results, as shown in Table 7, paint a somewhat similar picture for the information and communication industry (NACE code "J"). Here no significant relationship was found between the number of patents retrieved and the number of enterprises, with a negative adjusted R^2, showing independent variables yielding no explanatory power over the dependent variable. The same results can be reported for the model on the number of employees. However, the number of patents retrieved is found to be significantly and positively related to wage adjusted labor productivity (coeff. 0.033, SE 0.008) and results show an adjusted R^2 of 0.645. The same relationship can be reported for the gross value added per employee (coeff. 22.082, SE 3.893), which yields the highest adjusted R^2 (0.770) of all models in this analysis. The number of AI patents does not significantly explain the percentage share of personnel costs in production. Therefore, $H_{0,i,j}$ for $j \in \{l, v\}, i = J$ are rejected and $H_{0,i,j}$ for $j \in \{e, L, c\}, i = J$ are accepted.

In summary, the models' results paint a rather mixed picture with the majority of models tested showing statistically insignificant relationships between the number of AI patents retrieved for an industry, and the the chosen economic indicators reported within each industry. Only seven out of the 30 models tested exhibit statistically significant relationships. The results are further exacerbated, when one considers the fact that the chance of a rare event occuring increases with repeated exposure to that probability.[5] A common method to correct for the possibility of false positives is the Bonferroni Correction (Mittelhammer et al., 2000, p. 73f.). Given the above chosen α-level of 0.05, the Bonferroni Correction counterbalances the increased likelyhood of rare events (in this case, the Type I error) occuring when exposed to a plurality of situations in which they could occur (e.g., running a multitude of regressions). The Bonferroni Correction is calculated by dividing the chosen α-level by the number of tests conducted. In this case, the Bonferroni Correction would be $\frac{0.05}{30} = 0.00167$. This means that accounting for the number of models evaluated in this section, adjusted α-level would need to be set to 0.00167 to diminish the chance of false positives in the models' results.

Table 8: *Retrieved significant p values for coefficients of number of patents by industry and indicator*

Indicator	B	C	D	F	H	J
Employees (n)		0.04559				
Enterprises (n)					0.00726	
GVA/employee (€)		0.02772				0.00076
Labor prod. (%)					0.01001	0.00362
Personnel costs (%)					0.01913	

[5] A good analogy would be that the chance of winning the lottery increases with repeated playing. Or that the chance of rolling a six on a die is more likely in four rolls than in one roll.

Table 8, depicts only the number of patents' coefficient's p-values that lie beneath the unadjusted α threshold of 0.05. When considering the adjusted α value of 0.00167, one can see that merely one regression result's p value fulfills the new criterion (wage adjusted labor productivits in industry J). To conclude, the presented regression results vary in their significance and explanatory power to such extent, that is doubtful in how far relationships, while being statistically significant, actually exist. Furthermore, the Bonferroni Correction shows that the at least some of the presented results are likely to be false positives.

5 Discussion

5.1 Implications

This research aims to answer the question of if and how AI innovation impacts labor. Given the mixed results presented in Chapter 4, it is difficult to deduce clear implications of the findings.

Table 9: *Significant p values with coefficient sign, sample size and total number of patents*

	B	C	D	F	H	J	Patents (sum)	Sample size
Employees (n)		0.04559					1190.0	43.0
Enterprises (n)					0.00726*		1190.0	43.0
GVA/employee (€)		0.02772				0.00076*	1187.0	42.0
Labor prod. (%)					0.01001	0.00362*	1190.0	43.0
Personnel costs (%)					0.01913*		1188.0	40.0
Patents (sum)	17	39	26	21	374	713		
Sample size	5	5	7	6	10	10		

5.2 Limitations

5.2.1 Small data set

5.2.2 Patent weighting

As pointed out by Trajtenberg (1990), the plain number of patent counts disregard the fact that patents to not carry equal economical weight, i.e., the effect a patent might have on a market or industry cannot be inferred by the presence of a patent without incorporating weights.

5.2.3 Additional CPC classes

5.2.4 Further control variables

5.3 Further research

6 Conclusion

References

Acemoglu, D., 2021. Harms of AI. Working Paper Series. https://doi.org/10.3386/w29247

Acemoglu, D., Autor, D., Hazell, J., Restrepo, P., 2020a. AI and Jobs: Evidence from Online Vacancies. Working Paper Series. https://doi.org/10.3386/w28257

Acemoglu, D., Lelarge, C., Restrepo, P., 2020b. Competing with Robots: Firm-Level Evidence from France. AEA Papers & Proceedings 110, 383–388. https://doi.org/10.1257/pandp.20201003

Acemoglu, D., Restrepo, P., 2022. Tasks, Automation, and the Rise in U.S. Wage Inequality. Econometrica 90, 1973–2016. https://doi.org/10.3982/ECTA19815

Acemoglu, D., Restrepo, P., 2020a. Robots and Jobs: Evidence from US Labor Markets. Journal of Political Economy 128, 2188–2244. https://doi.org/10.1086/705716

Acemoglu, D., Restrepo, P., 2020b. Unpacking Skill Bias: Automation and New Tasks. AEA Papers & Proceedings 110, 356–361. https://doi.org/10.1257/pandp.20201063

Acemoglu, D., Restrepo, P., 2019. Automation and New Tasks: How Technology Displaces and Reinstates Labor. Journal of Economic Perspectives 33, 3–29. https://doi.org/10.1257/jep.33.2.3

Acemoglu, D., Restrepo, P., 2018a. The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment. American Economic Review 108, 1488–1542. https://doi.org/10.1257/aer.20160696

Acemoglu, D., Restrepo, P., 2018b. Artificial Intelligence, Automation, and Work, in: The Economics of Artificial Intelligence: An Agenda. University of Chicago Press, pp. 197–236.

Acemoglu, D., Restrepo, P., 2017. Low-Skill and High-Skill Automation. https://doi.org/10.2139/ssrn.3083552

Anderson, A.S. and M., 2017. Automation in Everyday Life. Pew Research Center: Internet, Science & Tech.

Arntz, M., Gregory, T., Zierahn, U., 2016. The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis. OECD, Paris. https://doi.org/10.1787/5jlz9h56dvq7-en

Autor, D.H., Dorn, D., 2013. The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. American Economic Review 103, 1553–1597. https://doi.org/10.1257/aer.103.5.1553

Autor, D.H., Dorn, D., Hanson, G.H., 2015. Untangling Trade and Technology: Evidence from Local Labour Markets. Economic Journal 125, 621–646. https://doi.org/10.1111/ecoj.12245

Beaudry, P., Green, D.A., Sand, B.M., 2016. The Great Reversal in the Demand for Skill and Cognitive Tasks. Journal of Labor Economics 34, S199–S247. https://doi.org/10.1086/682347

Boustan, L.P., Choi, J., Clingingsmith, D., 2022. Automation After the Assembly Line: Computerized Machine Tools, Employment and Productivity in the United States. Working Paper Series. https://doi.org/10.3386/w30400

Brynjolfsson, E., Mitchell, T., 2017. What can machine learning do? Workforce implications. Science 358, 1530–1534. https://doi.org/10.1126/science.aap8062

Brynjolfsson, E., Mitchell, T., Rock, D., 2018. What Can Machines Learn and What Does It Mean for Occupations and the Economy? AEA Papers & Proceedings 108, 43–47. https://doi.org/10.1257/pandp.20181019

Cirera, X., Sabetti, L., 2019. The effects of innovation on employment in developing countries: Evidence from enterprise surveys. Industrial and Corporate Change 28, 161–176. https://doi.org/10.1093/icc/dty061

Dauth, W., Findeisen, S., Südekum, J., Woessner, N., 2017. German Robots - The Impact of Industrial Robots on Workers. IAB Discussion Paper 30/2017.

Dauth, W., Findeisen, S., Suedekum, J., Woessner, N., 2021. The Adjustment of Labor Markets to Robots. Journal of the European Economic Association 19, 3104–3153. https://doi.org/10.1093/jeea/jvab012

European Commission, 2021. Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL.

European Commission, 202AD. ANNEXES to the Proposal for a Regulation of the European Parliament and of the Council.

European Patent Office, n.d. Published Data Keywords Search with Variable Constituents | EPO Developer Portal.

European Patent Office, 2023. Open Patent Services (OPS).

Eurostat, 2023a. Annual enterprise statistics for special aggregates of activities (NACE Rev. 2).

Eurostat, 2023b. Statistical classification of economic activities in the European Community (NACE Rev. 2).

Eurostat, 2023c. Economical indicator for structural business statistics.

Felten, E.W., Raj, M., Seamans, R., 2019. The Effect of Artificial Intelligence on Human Labor: An Abilitybased Approach. Academy of Management Annual Meeting Proceedings 2019, 791–796. https://doi.org/10.5465/AMBPP.2019.140

Frank, M.R., Autor, D., Bessen, J.E., Brynjolfsson, E., Cebrian, M., Deming, D.J., Feldman, M., Groh, M., Lobo, J., Moro, E., Wang, D., Youn, H., Rahwan, I., 2019. Toward understanding the impact of artificial intelligence on labor. Proceedings of the National Academy of Sciences 116, 6531–6539. https://doi.org/10.1073/pnas.1900949116

Frey, C.B., Osborne, M.A., 2017. The future of employment: How susceptible are jobs to computerisation? Technological Forecasting and Social Change 114, 254–280. https://doi.org/10.1016/j.techfore.2016.08.019

Graetz, G., Michaels, G., 2018. Robots at Work. Review of Economics & Statistics 100, 753–768. https://doi.org/10.1162/rest_a_00754

Karabarbounis, L., Neiman, B., 2014. The Global Decline of the Labor Share. The Quarterly Journal of Economics 129, 61–103. https://doi.org/10.1093/qje/qjt032

Legg, S., Hutter, M., 2007. A Collection of Definitions of Intelligence.

Mann, K., Püttmann, L., 2018. Benign Effects of Automation: New Evidence from Patent Texts. https://doi.org/10.2139/ssrn.2959584

Mittelhammer, R.C., Judge, G.G., Miller, D.J., 2000. Econometric foundations, 1. publ. ed. Cambridge Univ.

Press, Cambridge.

Mokyr, J., Vickers, C., Ziebarth, N.L., 2015. The History of Technological Anxiety and the Future of Economic Growth: Is This Time Different? Journal of Economic Perspectives 29, 31–50. https://doi.org/10.1257/jep.29.3.31

Pajarinen, M., Rouvinen, P., Ekeland, A., 2015. Computerization Threatens One-Third of Finnish and Norwegian Employment. ETLA Brief.

The pandas development team, 2023. Pandas-dev/pandas: Pandas. https://doi.org/10.5281/ZENODO.3509134

The SciPy community, 2023. Scipy.signal.detrend — SciPy v1.11.3 Manual.

Trajtenberg, M., 1990. A Penny for Your Quotes: Patent Citations and the Value of Innovations. The RAND Journal of Economics 21, 172. https://doi.org/10.2307/2555502

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., Van Der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., Van Mulbregt, P., SciPy 1.0 Contributors, Vijaykumar, A., Bardelli, A.P., Rothberg, A., Hilboll, A., Kloeckner, A., Scopatz, A., Lee, A., Rokem, A., Woods, C.N., Fulton, C., Masson, C., Häggström, C., Fitzgerald, C., Nicholson, D.A., Hagen, D.R., Pasechnik, D.V., Olivetti, E., Martin, E., Wieser, E., Silva, F., Lenders, F., Wilhelm, F., Young, G., Price, G.A., Ingold, G.-L., Allen, G.E., Lee, G.R., Audren, H., Probst, I., Dietrich, J.P., Silterra, J., Webber, J.T., Slavič, J., Nothman, J., Buchner, J., Kulick, J., Schönberger, J.L., De Miranda Cardoso, J.V., Reimer, J., Harrington, J., Rodríguez, J.L.C., Nunez-Iglesias, J., Kuczynski, J., Tritz, K., Thoma, M., Newville, M., Kümmerer, M., Bolingbroke, M., Tartre, M., Pak, M., Smith, N.J., Nowaczyk, N., Shebanov, N., Pavlyk, O., Brodtkorb, P.A., Lee, P., McGibbon, R.T., Feldbauer, R., Lewis, S., Tygier, S., Sievert, S., Vigna, S., Peterson, S., More, S., Pudlik, T., Oshima, T., Pingel, T.J., Robitaille, T.P., Spura, T., Jones, T.R., Cera, T., Leslie, T., Zito, T., Krauss, T., Upadhyay, U., Halchenko, Y.O., Vázquez-Baeza, Y., 2020. SciPy 1.0: Fundamental algorithms for scientific computing in Python. Nature Methods 17, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Webb, M., 2019. The Impact of Artificial Intelligence on the Labor Market. https://doi.org/10.2139/ssrn.3482150

Appendix

Table 10: *Selected main keywords for NACE industries*

NACE Section	Keywords
A	AGRICULTURE, FORESTRY, FISHING
B	MINING, QUARRYING
C	MANUFACTURING
D	ELECTRICITY, GAS, STEAM, AIR CONDITIONING SUPPLY
E	WATER SUPPLY, SEWERAGE, WASTE MANAGEMENT, REMEDIATION ACTIVITIES
F	CONSTRUCTION
G	WHOLESALE, RETAIL, REPAIR OF MOTOR VEHICLES AND MOTORCYCLES
H	TRANSPORTATION, STORAGE
I	ACCOMMODATION, FOOD SERVICE ACTIVITIES
J	INFORMATION, COMMUNICATION
K	FINANCIAL ACTIVITIES, INSURANCE ACTIVITIES
L	REAL ESTATE
M	PROFESSIONAL ACTIVITIES, SCIENTIFIC ACTIVITIES, TECHNICAL ACTIVITIES
N	ADMINISTRATIVE ACTIVITIES, SUPPORT SERVICE ACTIVITIES
O	PUBLIC ADMINISTRATION, DEFENCE, COMPULSORY SOCIAL SECURITY
P	EDUCATION
Q	HUMAN HEALTH, SOCIAL WORK ACTIVITIES
R	ARTS, ENTERTAINMENT, RECREATION
S	OTHER SERVICE ACTIVITIES
T	ACTIVITIES OF HOUSEHOLDS AS EMPLOYERS, UNDIFFERENTIATED GOODS, SERVICES PRODUCING ACTIVITIES OF HOUSEHOLDS FOR OWN USE
U	ACTIVITIES OF EXTRATERRITORIAL ORGANISATIONS, ACTIVITIES OF EXTRATERRITORIAL BODIES

Table 11: *Example queries posted to the OPS API*

	Query examples
0	(ta ALL “agriculture” OR ta ALL “forestry” OR ta ALL “fishing”) AND (ta = “farming” OR ta = “raising” OR ta = “related” OR ta = “dairy” OR ta = “grapes” OR ta = “oleaginous” OR ta = “buffaloes”) AND cpc any “G06N20/00 G06N20/10 G06N20/20 G06N3/09 G06N3/088 G06N3/092 G06N3/08” AND AP=“EP”
1	(ta ALL “agriculture” OR ta ALL “forestry” OR ta ALL “fishing”) AND (ta = “leguminous” OR ta = “beverage” OR ta = “pharmaceutical” OR ta = “hunting” OR ta = “tree” OR ta = “sheep” OR ta = “pome”) AND cpc any “G06N20/00 G06N20/10 G06N20/20 G06N3/09 G06N3/088 G06N3/092 G06N3/08” AND AP=“EP”
2	(ta ALL “agriculture” OR ta ALL “forestry” OR ta ALL “fishing”) AND (ta = “propagation” OR ta = “camelids” OR ta = “crop” OR ta = “stone” OR ta = “tropical” OR ta = “oil” OR ta = “service”) AND cpc any “G06N20/00 G06N20/10 G06N20/20 G06N3/09 G06N3/088 G06N3/092 G06N3/08” AND AP=“EP”
3	(ta ALL “agriculture” OR ta ALL “forestry” OR ta ALL “fishing”) AND (ta = “roots” OR ta = “bush” OR ta = “post-harvest” OR ta = “tobacco” OR ta = “cereals” OR ta = “cattle” OR ta = “crops”) AND cpc any “G06N20/00 G06N20/10 G06N20/20 G06N3/09 G06N3/088 G06N3/092 G06N3/08” AND AP=“EP”
4	(ta ALL “agriculture” OR ta ALL “forestry” OR ta ALL “fishing”) AND (ta = “spices” OR ta = “perennial” OR ta = “subtropical” OR ta = “except” OR ta = “camels” OR ta = “support” OR ta = “plant”) AND cpc any “G06N20/00 G06N20/10 G06N20/20 G06N3/09 G06N3/088 G06N3/092 G06N3/08” AND AP=“EP”
5	(ta ALL “agriculture” OR ta ALL “forestry” OR ta ALL “fishing”) AND (ta = “melons” OR ta = “poultry” OR ta = “animals” OR ta = “vegetables” OR ta = “nuts” OR ta = “activities” OR ta = “trapping”) AND cpc any “G06N20/00 G06N20/10 G06N20/20 G06N3/09 G06N3/088 G06N3/092 G06N3/08” AND AP=“EP”
6	(ta ALL “agriculture” OR ta ALL “forestry” OR ta ALL “fishing”) AND (ta = “agriculture” OR ta = “production” OR ta = “processing” OR ta = “citrus” OR ta = “aromatic” OR ta = “rice” OR ta = “non-perennial”) AND cpc any “G06N20/00 G06N20/10 G06N20/20 G06N3/09 G06N3/088 G06N3/092 G06N3/08” AND AP=“EP”
7	(ta ALL “agriculture” OR ta ALL “forestry” OR ta ALL “fishing”) AND (ta = “horses” OR ta = “goats” OR ta = “fruits” OR ta = “seeds” OR ta = “equines” OR ta = “seed” OR ta = “drug”) AND cpc any “G06N20/00 G06N20/10 G06N20/20 G06N3/09 G06N3/088 G06N3/092 G06N3/08” AND AP=“EP”
8	(ta ALL “agriculture” OR ta ALL “forestry” OR ta ALL “fishing”) AND (ta = “fibre” OR ta = “growing” OR ta = “sugar” OR ta = “mixed” OR ta = “cane” OR ta = “tubers” OR ta = “pigs”) AND cpc any “G06N20/00 G06N20/10 G06N20/20 G06N3/09 G06N3/088 G06N3/092 G06N3/08” AND AP=“EP”
9	(ta ALL “agriculture” OR ta ALL “forestry” OR ta ALL “fishing”) AND (ta = “swine” OR ta = “animal”) AND cpc any “G06N20/00 G06N20/10 G06N20/20 G06N3/09 G06N3/088 G06N3/092 G06N3/08” AND AP=“EP”