Evidence of labour-saving technologies within European patents

Jacopo Staccioli[a b] Mariagrazia Squicciarini[c]

[a] *Department of Economic Policy, Università Cattolica del Sacro Cuore, Milan*

[b] *Institute of Economics, Scuola Superiore Sant'Anna, Pisa*

[c] *Directorate for Science, Technology and Innovation, OECD, Paris*

OECD – CIIE Breakout Webinar

7th December 2020

UNIVERSITÀ CATTOLICA del Sacro Cuore

Outline

Context

- the impact of automation on employment is a major topic in economic policy debate

Brynjolfsson and McAfee (2011, 2014) the root of current unemployment is not the Great Recession, but rather a 'Great Restructuring' characterised by an exponential growth in computers' processing power having an ever-bigger impact on jobs, skills, and the whole economy (*"This time is different"*)

Frey and Osborne (2017) 47% of the occupational categories are at high risk of being automated, including services and highly cognitive jobs

Arntz et al. (2016) 9% of jobs in OECD countries are automatable on average (6% in South Korea, 12% in Austria)

Nedelkoska and Quintini (2018) 14% of jobs in OECD countries have probability of automation above 70%

Motivation

- policymakers committed to ensure an orderly transition to the new industrial paradigm (e.g. avoiding episodes of considerable unemployment) need to know *in advance*
  - which occupational categories are likely to be substituted in the future
  - which skills are required by remaining (and possibly new) professions
- e.g. governments could set up targeted training infrastructure to make sure that workers possess the necessary capabilities demanded by employers
- to do so, a *leading* (not *lagging*) indicator of automation innovation is required
- patents are well suited in that they provide a proxy of innovative effort at the *invention* stage, well before actual implementation in production

Patents

"[R]obots [. . .] [satisfy] the demand for saving labor and rationalization of work in view of the current rise in labor cost." [EP0068026A1, 1980]

"The need for skilled labor, along with the attendant costs in training and replacement is reduced and, furthermore, if the skills involved constitute more an art than a skill, the call for such talent is avoided." [EP0778957B1, 1995]

"Automated machining stations can be used to manufacture large quantities of pieces quickly and completely without human intervention." [EP2475501B1, 2009]

"The main function of a robot arm is to act as a substitute for human arms and do repetitive or demanding works so as to increase productivity and reduce labor costs." [EP3379410A1, 2017]

Our contribution

- we use textual analysis on the universe (6m+) of 1978–2019 patent applications at EPO
- we investigate the presence of *explicit* labour-saving heuristics among robotics patents
- robotics patents are broadly defined, entailing robotic artefacts as a *product* but also as *process* and complementary technology
- we analyse innovative actors engaged in labour-saving technology and their economic environment (identity, location, industry)
- we provide an appraisal of possibly impacted occupations

- extension of another analysis on USPTO patents
  Montobbio, Fabio, Jacopo Staccioli, Maria Enrica Virgillito, and Marco Vivarelli (2020)
  Robots and the origin of their labour-saving impact. LEM Working Papers series n. 3/2020

Outline

Original data

- universe of EPO patent applications between 1978 and 2019
- 6,109,462 full-text patent applications, of which 4,382,445 in English

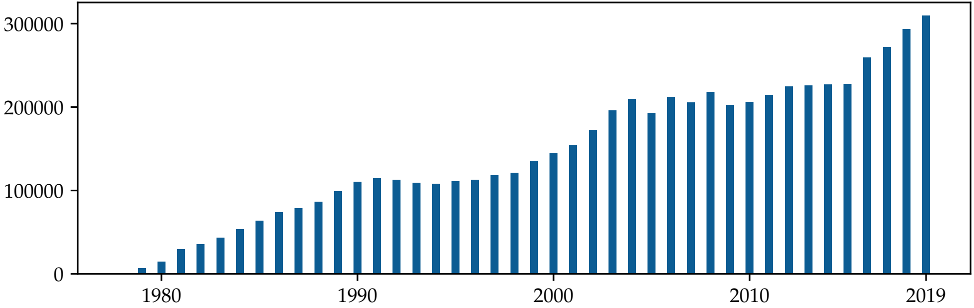


Figure: # of patents by year

Robotics patents

- identification of robotics[-related] patents
  1. via IPC codes
     - concordance table with USPC class 901 (*'Robots'*)
     - purely robotic technology
     - 13,852 `'IPC'` patents
  2. via keyword search
     - multiple occurrence ($\times 7$) of morphological root *'robot'*
     - process implementation and complementary technology
     - 8,125 `'K7'` patents (excluding those already found in 1)
- 21,977 total robotics patents

Robotics patents (cont'd)

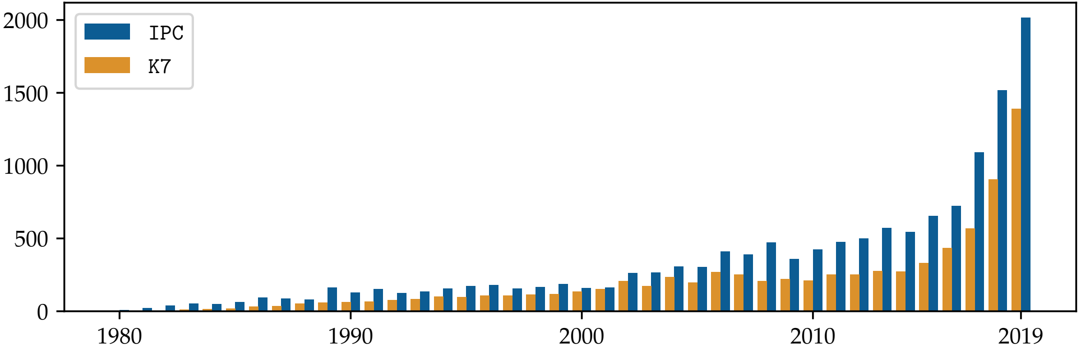


Figure: # of **robotics** patents by year

Text preprocessing

tokenisation

- each patent textual body is divided into *sentences* by means of a punctuation regexp
- patent text $\Longrightarrow$ list of sentences
- sentence $\Longrightarrow$ list of words

stemming

- each word in each sentence is reduced to its morphological root with the Porter2 stemming algorithm (an improved version of the original Porter (1980) algorithm)
- patent text $\Longrightarrow$ list of lists of stemmed words

- identification of labour-saving (LS) patents by means of a word-level text query per sentence

Labour-saving patents

$$\underbrace{\begin{bmatrix}\text{`reduc'}\\ \text{`replac'}\\ \text{`elimin'}\\ \text{`save'}\\ \text{`lower'}\\ \text{`substitut'}\\ \text{`autom'}\\ \text{`less'}\end{bmatrix}}_{\text{verbal predicate}} \times \underbrace{\begin{bmatrix}\text{`labor'}\\ \text{`worker'}\\ \text{`human'}\\ \text{`employe'}\\ \text{`manpow'}\\ \text{`job'}\\ \text{`manual'}\end{bmatrix}}_{\text{direct object}} \times \underbrace{\begin{bmatrix}\text{`cost'}\\ \text{`expenditure'}\\ \text{`expens'}\\ \text{`hour'}\\ \text{`intens'}\\ \text{`task'}\\ \text{`time'}\\ \text{`skill'}\\ \text{`effort'}\\ \text{`intervent'}\end{bmatrix}}_{\text{object attribute}}$$

- 640 combinations of triplets (**not** *trigrams*, as we do not require adjacency)
- a patent is flagged as *potentially* LS if contains at least one triplet within a sentence
- 1,662 potentially LS patents

Labour-saving patents (cont'd)

- all matched sentences are manually examined and flagged as *explicitly* LS if appropriate
- 1,545 explicitly LS patents ($\approx$ 93% of potentially LS; $\approx$ 7% of robotic patents)
- of which 814 ($\approx$ 52.7%) are IPC and 731 ($\approx$ 47.3%) are K7

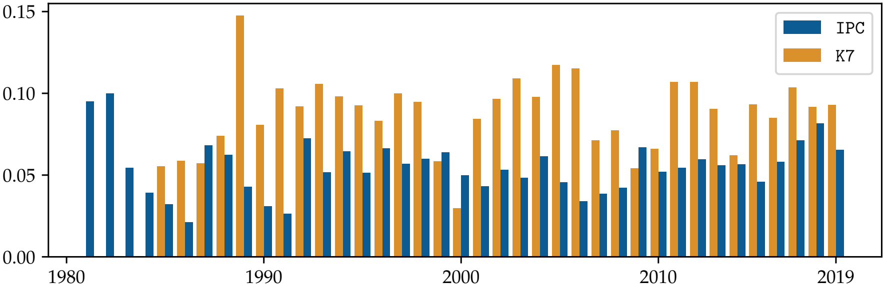


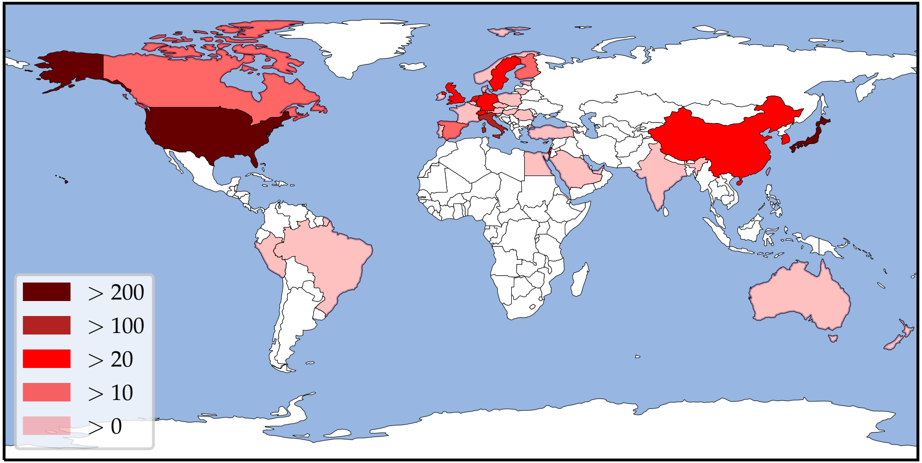
Figure: Fraction of **explicitly** LS patents over **robotic** patents by year

Firm level match

- LS patents are matched to their assignee via ORBIS (BvD)
- 1,322 ($\approx$ 85.6%) are matched to at least one firm
- there are 787 LS firms in total

Outline

LS patents by country



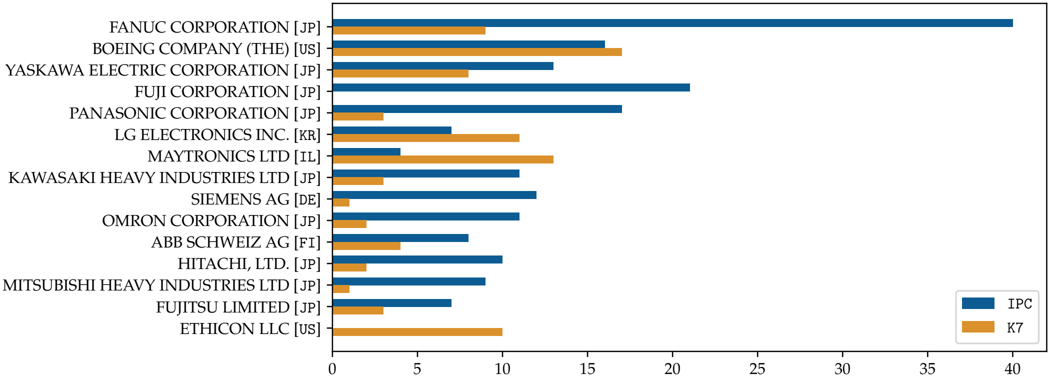
LS patents by EPO member states

Country	LS patents IPC	LS patents K7	LS patents Total
Italy	73	34	107
Germany	36	18	54
Netherlands	14	30	44
Switzerland	27	13	40
Sweden	18	19	37
United Kingdom	18	10	28
France	6	12	18
Spain	9	9	18
Finland	15	1	16
Belgium	5	4	9
Portugal	3	1	4
Denmark	0	3	3
Norway	1	2	3
…			

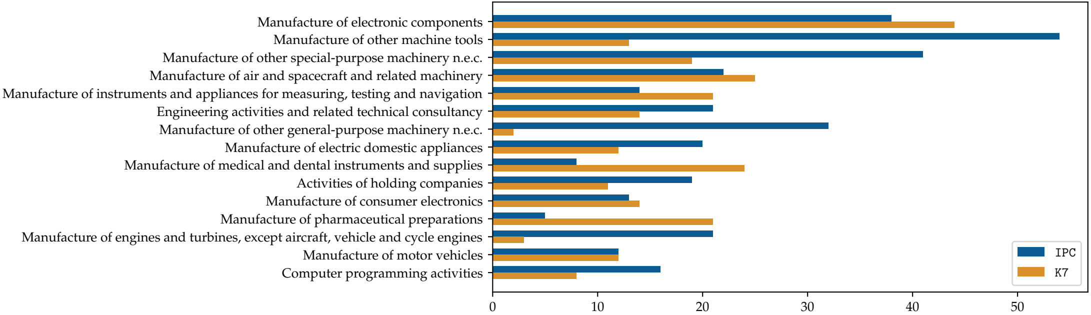
…

Country	LS patents IPC	LS patents K7	LS patents Total
Turkey	2	0	2
Ireland	0	2	2
Estonia	0	2	2
Slovenia	1	0	1
Austria	1	0	1
Poland	1	0	1
Romania	1	0	1
Lithuania	1	0	1
Czech Republic	1	0	1
Liechtenstein	0	1	1
Hungary	0	1	1
Malta	0	1	1
Total	234	169	403

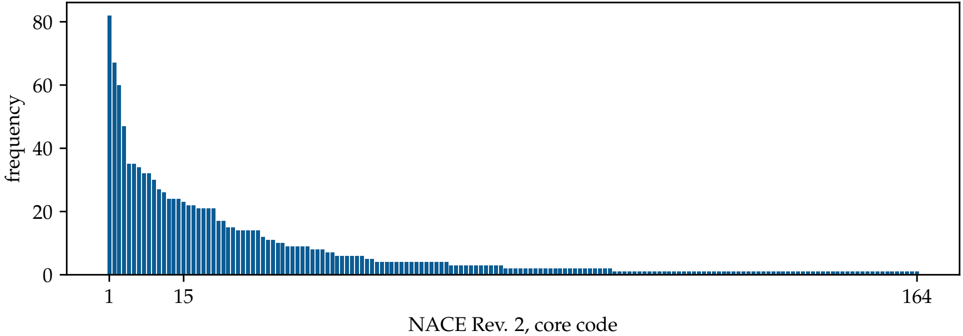
LS patents by firm



LS patents by industry



LS patents by industry (cont'd)



in the tail: automatic fruit picking, sushi/pizza preparation, agricultural pollination…

Outline

Occupations and LS patents

- we compute a text similarity measure between occupations and patents

Occupations corpus

- International Standard Classification of Occupations (ISCO)
- we exclude `major group 10` *"Armed forces occupations"*
- 430 definitions of 4-digit occupations (`unit group`)

Patents corpus

- 1545 LS patents previously found
- concatenation of `abstract`, `description`, and `claims`

Document-term matrix

1. construct the *document-term matrix* $\mathcal{D}_{ISCO}$ of the corpus D of ISCO occupations
   - each cell contains the frequency of term t in occupation d
   - tf-idf: term frequency–inverse document frequency

$$\text{tf-idf}(t, d, D) := \text{tf}(t, d) \cdot \text{idf}(t, D)$$

$$\text{tf}(t, d) := \mathbf{1}_d(t) = \begin{cases} 1 & \text{if } t \in d \\ 0 & \text{otherwise} \end{cases}$$

$$\text{idf}(t, D) := \log\left(\frac{|D|}{|\{d \in D : t \in d\}|}\right)$$

   - restrict to tf-idf > 0.4 to discard terms not specific enough to the underlying occupation
   - 430×5633 matrix

Cosine similarity

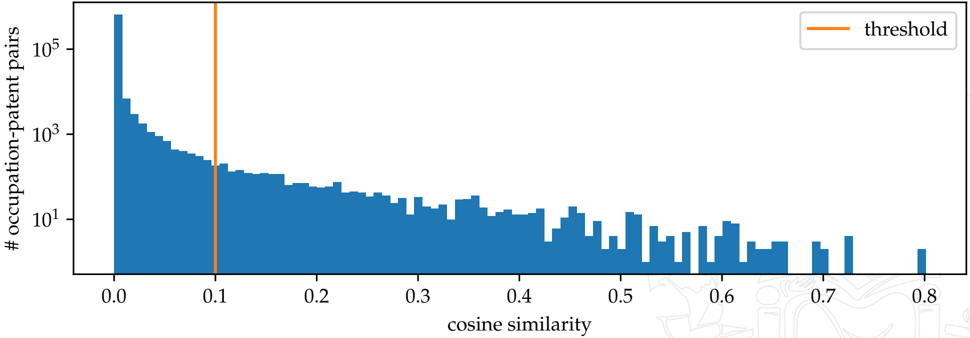
2. construct the *document-term matrix* $\mathcal{D}_{LS}$ of the corpus of LS patents
   - projected on the *vocabulary* of the occupations matrix
   - 1545×5633 matrix
3. construct the cosine similarity measure between the two corpora
   - for each couple of row vectors $X \in \mathcal{D}_{ISCO}$, $Y \in \mathcal{D}_{LS}$

$$\text{cosine}(X, Y) := \frac{\sum_t x_t y_t}{\sqrt{\sum_t x_t^2}\sqrt{\sum_t y_t^2}}$$

   - $\text{cosine}(X, Y) \in [0, 1]$ since vectors X and Y are non-negative valued
   - w.r.t. Euclidean distance, cosine similarity normalises for varying lengths of documents
   - 430×1545 cosine similarity matrix[1]
   - each occupation obtains a similarity score to each LS patent (664,350 occupation-patent pairs)

[1] under tf-idf, it is possible to show that $\text{cosine}(\mathcal{D}_{ISCO}, \mathcal{D}_{LS}) \equiv \mathcal{D}_{ISCO} \cdot \mathcal{D}'_{LS}$

Cosine similarity (cont'd)



- values above the threshold account for 2,413 occupation-patent pairs ($\approx 0.36\%$)
- robustness: no threshold and threshold $= 0.5$ ($\approx 0.02\%$ pairs) yield fairly similar results

Top 15 occupations by similarity

Similarity*	ISCO code	ISCO title
40.84	2511	Systems Analysts
33.89	4132	Data Entry Clerks
32.16	9331	Hand and Pedal Vehicle Drivers
32.16	9122	Vehicle Cleaners
21.58	7231	Motor Vehicle Mechanics and Repairers
13.34	2514	Applications Programmers
13.19	8156	Shoemaking and Related Machine Operators
13.19	8183	Packing, Bottling and Labelling Machine Operators
13.19	8159	Textile, Fur and Leather Products Machine Operators
13.19	8154	Bleaching, Dyeing and Fabric Cleaning Machine Operators
13.19	7223	Metal Working Machine Tool Setters and Operators
9.06	3254	Dispensing Opticians
8.67	3131	Power Production Plant Operators
8.60	7549	Craft and Related Workers
7.78	7213	Sheet Metal Workers

*sum of cosine similarity score across all LS patents

Focus: patents published between 2015 and 2019

Similarity	ISCO code	ISCO title
20.14	9122	Vehicle Cleaners
20.14	9331	Hand and Pedal Vehicle Drivers
19.33	2511	Systems Analysts
13.85	4132	Data Entry Clerks
12.32	7231	Motor Vehicle Mechanics and Repairers
6.6	3131	Power Production Plant Operators
6.12	5165	Driving Instructors
4.1	7223	Metal Working Machine Tool Setters and Operators
4.1	8154	Bleaching, Dyeing and Fabric Cleaning Machine Operators
4.1	8156	Shoemaking and Related Machine Operators
4.1	8159	Textile, Fur and Leather Products Machine Operators
4.1	8183	Packing, Bottling and Labelling Machine Operators
3.27	7311	Precision-instrument Makers and Repairers
3.22	6130	Mixed Crop and Animal Producers
3.22	6320	Subsistence Livestock Farmers

red titles were not present in the previous picture

Occupations and number of patents per year

Code	Title	#patents	2019	2018	2017	2016	2015	2014
2511	Systems Analysts	256	51	34	16	14	12	9
4132	Data Entry Clerks	161	26	21	8	11	5	8
9331	Hand and Pedal Vehicle Drivers	100	23	23	10	4	4	2
9122	Vehicle Cleaners	100	23	23	10	4	4	2
7231	Motor Vehicle Mechanics and Repairers	104	21	21	12	4	3	4
2514	Applications Programmers	63	2	5	3	4	1	2
8156	Shoemaking and Related Machine Operators	73	4	5	8	1	5	2
8183	Packing, Bottling and Labelling Machine Operators	73	4	5	8	1	5	2
8159	Textile, Fur and Leather Products Machine Oper...	73	4	5	8	1	5	2
8154	Bleaching, Dyeing and Fabric Cleaning Machine...	73	4	5	8	1	5	2
7223	Metal Working Machine Tool Setters and Operators	73	4	5	8	1	5	2
3254	Dispensing Opticians	34	2	3	5	0	1	2
3131	Power Production Plant Operators	37	8	7	6	5	2	0
7549	Craft and Related Workers	32	2	2	5	0	1	1
7213	Sheet Metal Workers	30	2	1	1	0	1	1

Outline

Wrap up

- while the number of robotics patent applications has increased over time, the presence of embedded labour-saving heuristics has been stable
- traditional robotics competence centres (Japan, US, Italy) dominate the picture, although developing countries are also present (China, India, Turkey, Brazil…)
- robots manufacturers account for most LS (*product*) innovations
- but LS heuristics emerge along the entire supply chain
- industries such as mining, retail, food processing etc. patent robotics technology as *process* innovation, which could lead to employment disruption once implemented
- wide range of involved occupations, from low-skilled/blue collar to highly specialised

Thank you very much!

jacopo.staccioli@unicatt.it

this presentation is available at www.staccioli.org