

# Robotic patents and the origin of labour-saving technology

Fabio Montobbio<sup>a b</sup>    **Jacopo Staccioli**<sup>a c</sup>    M. Enrica Virgillito<sup>a c</sup>    Marco Vivarelli<sup>a c</sup>

<sup>a</sup>*Department of Economic Policy, Università Cattolica del Sacro Cuore, Milano*

<sup>b</sup>*BRICK, Collegio Carlo Alberto, Torino*

<sup>c</sup>*Institute of Economics, Scuola Superiore Sant'Anna, Pisa*

This version: 25<sup>th</sup> September 2019



- 1 Context and motivation**
- 2 Data and analysis
- 3 Results
- 4 Topic modelling and technological taxonomy
- 5 Discussion

*“Automated systems, such as robotic systems, are used in a variety of industries to **reduce labor costs and/or increase productivity**. Additionally, the use of human operators can involve increased cost relative to automated systems.”* [US20170178485(A1)]

*“The use of [robotic] technology results in improved management of information, services, and data, increased efficiency, significant reduction of time, **decreased manpower requirements**, and substantial cost savings.”* [US20100223134(A1)]



## Motivation (cont'd)

- the impact of robotic automation upon employment has become a major topic of discussion both in the policy and academic 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

# Our contribution

- we aim at contributing by extending the analysis beyond the IFR dataset
- by using **textual** patent data on robotic artefacts, we are able to identify the knowledge generation patterns behind this technology
- we intend to analyse not only those patents entailing robotic artefacts as *products* but also as *processes*
- we aim at understanding the extent to which the heuristics behind the generation of robotic patents are *truly* labour-saving



# Research question

- 1 to what extent robotic patents make **explicit** reference to labour-saving technology?
- 2 what are the underlying patenting firms?
- 3 how are they distributed by country and by industry?
- 4 which technological realms does labour-saving innovation build upon?



# Beyond traditional patent classification

- we adopt a broad definition of robotic patents that goes beyond the traditional classification by patent examiners
- we perform an in depth semantic analysis in order to distinguish those patents that *explicitly* claim to have a direct labour-saving impact
- in so doing, we restrict the analysis to a lexical domain uncommon for inventors
- in particular, we look at a dictionary of words typical of the ‘economic slang’

- 1 Context and motivation
- 2 Data and analysis**
- 3 Results
- 4 Topic modelling and technological taxonomy
- 5 Discussion



# Original data

- universe of USPTO patent applications from 1st January 2009 to 31st December 2018
- 3,557,435 full-text applications (hereafter, patents)

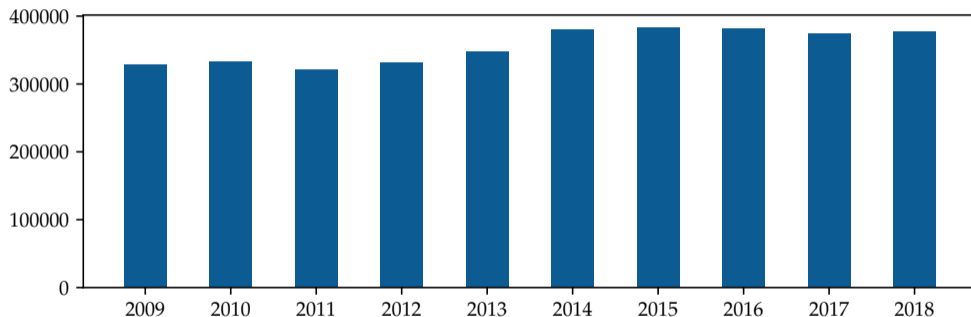


Figure: # of patents by year



- identification of robotics[-related] patents
  - 1 via CPC codes
    - USPTO concordance table with USPC class 901
    - purely robotic technology
    - 10,929 'CPC' patents
  - 2 via keyword search
    - multiple occurrence ( $\times 10$ ) of morphological root '*robot*'
    - process implementation and complementary technology
    - 18,860 'K10' patents (once those already in 1 have been discarded)
- 29,789 total robotic patents



# Data processing (cont'd)

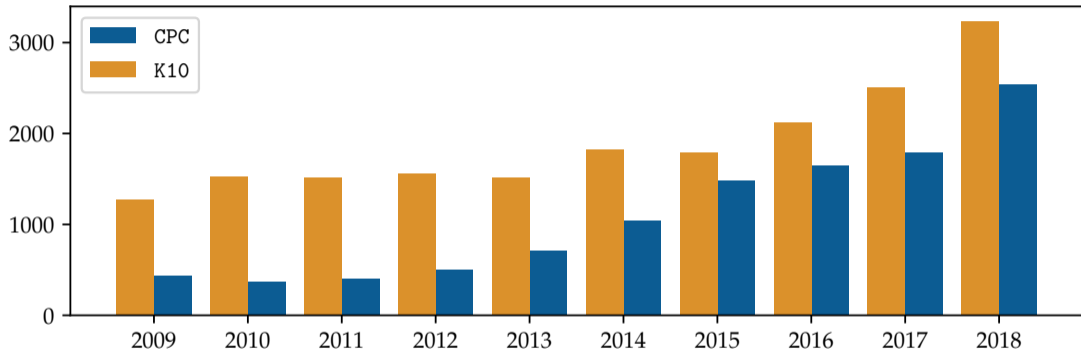


Figure: # of **robotic** patents by year



## tokenisation

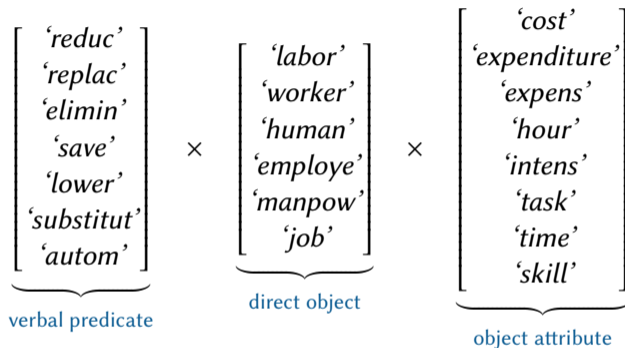
- each patent textual body is divided into *sentences* by means of a punctuation regexp
- patent text  $\implies$  list of sentences
- sentence  $\implies$  list of words

## 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  $\implies$  list of lists of stemmed words
- identification of labour-saving (LS) patents by means of a **word-level** text query **per sentence**



# Text query



- 336 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
- 1,666 potentially LS patents

# Validation

- all matched sentences are **manually** examined and flagged as *explicitly* LS if appropriate
- 1,276 explicitly LS patents ( $\approx 77\%$  of potentially LS;  $\approx 4.3\%$  of robotic patents)
- of which 461 ( $\approx 36.1\%$ ) are **CPC** and 815 ( $\approx 63.9\%$ ) are **K10**

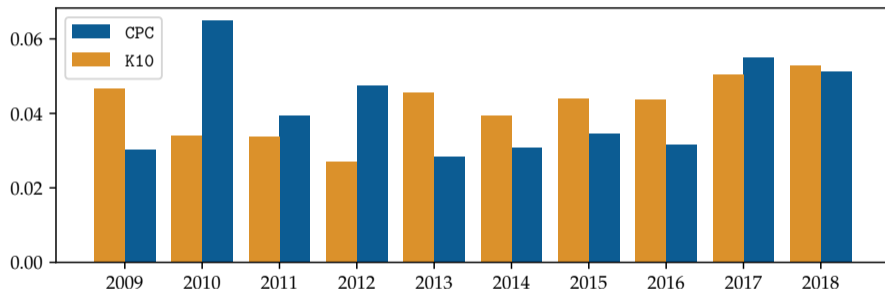


Figure: Fraction of **explicitly** LS patents over **robotic** patents by year

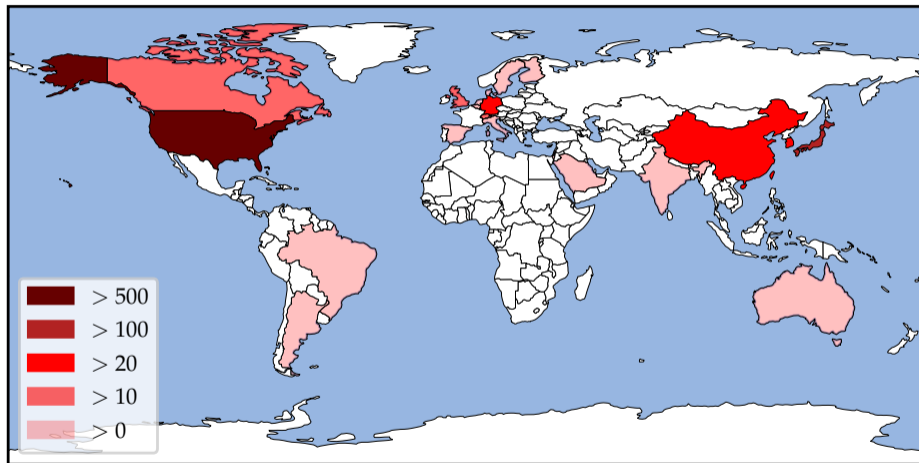


- LS patents are matched to their assignee via ORBIS (BvD)
- number reduces to 1,136 ( $\approx 89\%$ ) due to truncation on 31st July 2018 (140 discarded)
- of these, 903 ( $\approx 79\%$ ) are matched to at least one firm (233 find no match)
- there are 408 LS firms in total

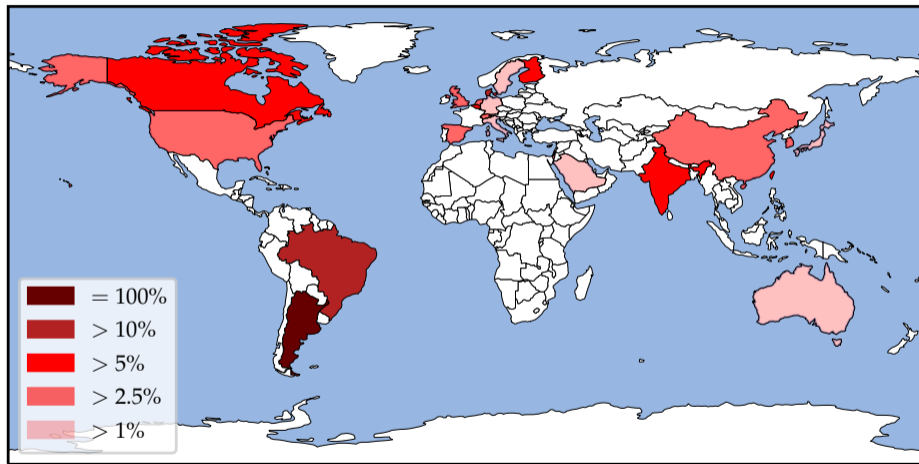
- 1 Context and motivation
- 2 Data and analysis
- 3 Results**
- 4 Topic modelling and technological taxonomy
- 5 Discussion



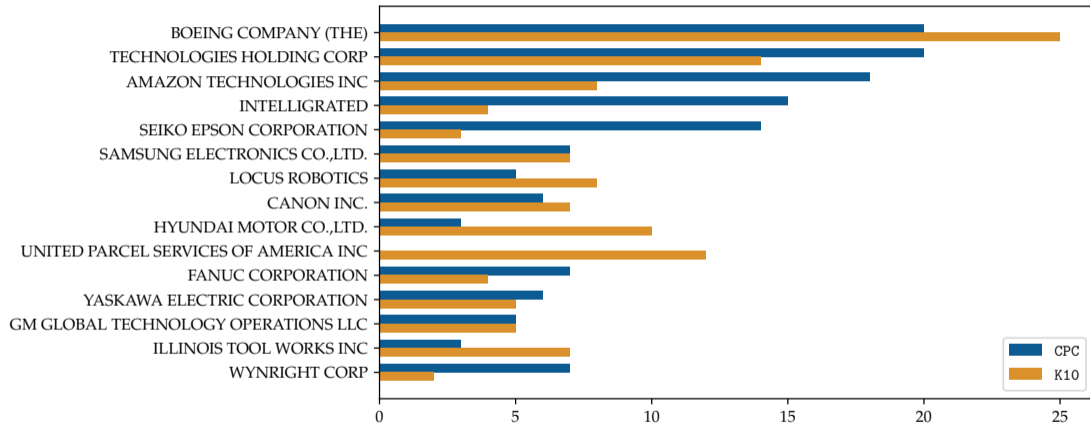
# LS patents by country – absolute value



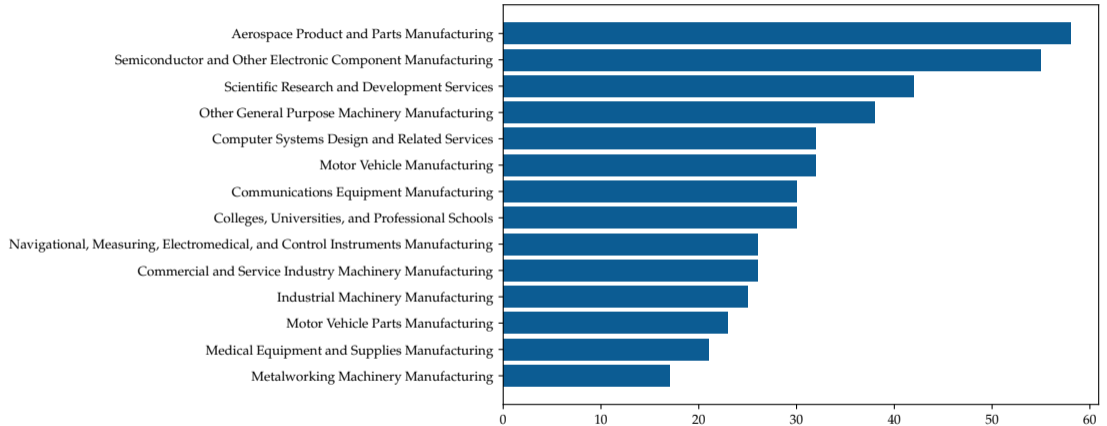
# LS patents by country – as % of robotic patents



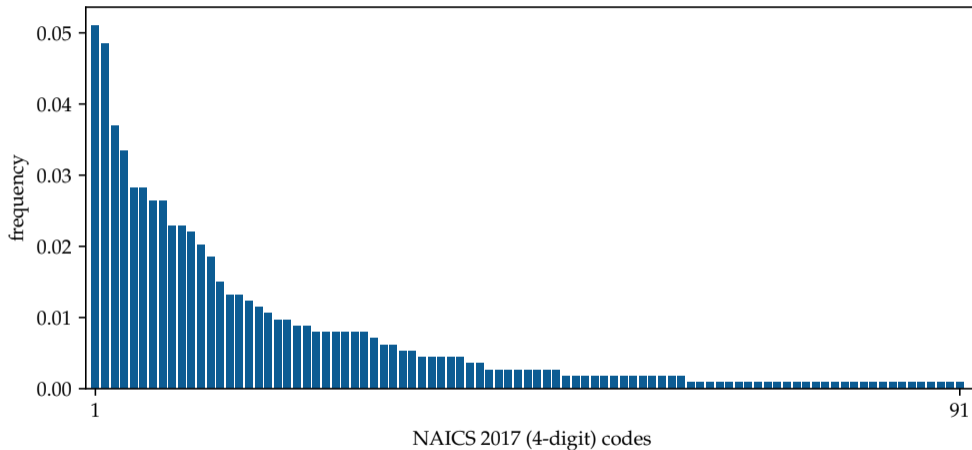
# LS patents by assignee



# LS patents by industry (cont'd)



# LS patents by industry



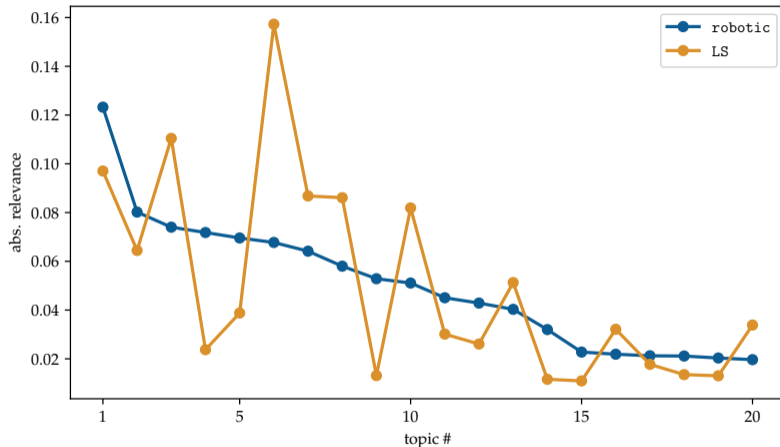
- 1 Context and motivation
- 2 Data and analysis
- 3 Results
- 4 Topic modelling and technological taxonomy**
- 5 Discussion

- 1 we estimate a topic model with  $K = 20$  topics on the whole collection of robotic patents  $D$ 
  - each topic  $k \in K$  is identified as a list of keywords ranked by frequency
  - each patent  $d \in D$  is assigned a distribution  $r(k)^d$  over the  $K$  topics
- 2 we assign a significance measure  $M(c)^k$  of CPC codes ( $c \in C$ ) originally attributed to patents to each topic  $k$  by leveraging on the *latent semantic structure* of the whole collection of patents, through relevance distributions  $r(k)^d$  obtained in **1**

$$m(c)^k = \sum_{d \in D} \mathbf{1}_{\{c \in d\}} \cdot r(k)^d \quad \forall k = 1, \dots, K; \quad \forall c \in C$$

- this brings useful information for labelling the topics
- 3 we compare the relevance of the  $K$  topics for robotic patents and the subset of LS patents

# Topic relevance for robotic and LS patents





# Technological taxonomy

Topic #	LS relev.	Words	CPC	Weight	Description
6	+132.2%	carrier	B65	24.4%	Conveying; packing; storing; handling thin or filamentary material
		conveyor			
		item	H01	6.8%	Basic electric elements
		gripper	G11	6.0%	Information storage
		tape	Y02	4.6%	Technologies or applications for mitigation or adaptation against climate change
			B23	4.3%	Machine tools; metal-working not otherwise provided for
...	...	...	...	...	...
14	-63.7%	effector	A61	55.6%	Medical or veterinary science; hygiene
		surgic	Y10T29	4.9%	Metal working
		stapl	H01	4.8%	Basic electric elements
		articul	Y10T74	4.2%	Machine element or mechanism
		fire	B23	3.5%	Machine tools; metal-working not otherwise provided for



- 1 Context and motivation
- 2 Data and analysis
- 3 Results
- 4 Topic modelling and technological taxonomy
- 5 Discussion**

# Main findings

- patenting firms are not only robots producers, but mainly users (archetypical cases are Boeing, Amazon, and UPS)
- the overall number of robotic patents is rapidly expanding (3-fold increase in a decade)
- conversely, LS patents do not exhibit a clear trend, supporting the idea that labour-saving is a rather mature heuristic
- LS robotic patents emerge along the entire supply chain, signalling pervasiveness
- LS patents are especially concentrated in labour intensive industries
- geographically, US and Japan still appear to largely dominate (possible bias from using USPTO patents), although China seems to be catching up
- but in relative terms this picture is reversed

# Conceivable extensions

- analysis of economic performance of LS firms vis-à-vis the rest of patenting robotic firms
  - productivity dynamics
  - employment dynamics
- grasp the underlying *skills* and *tasks* LS patents are likely to automate
  - policy implications in terms of retraining displaced workers
- refine and fully automate the information extraction query
  - extend the analysis to the *whole* patent universe (possibly including other patent offices)
  - look for different search heuristics (e.g. mitigation of climate change)

# Thank you very much!

[jacopo.staccioli@unicatt.it](mailto:jacopo.staccioli@unicatt.it)

<https://staccioli.org/>

