Labour-saving technologies and occupational exposure

Fabio Montobbio[a b] **Jacopo Staccioli**[a c] M. Enrica Virgillito[c a] Marco Vivarelli[a c]

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

[b] *BRICK, Collegio Carlo Alberto, Torino*

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

Workshop "Innovation, Firm Growth, and Employment"

23rd April 2021

UNIVERSITÀ CATTOLICA del Sacro Cuore

Outline

Context

- the impact of automation upon employment has become a major topic of discussion both in 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

  Acemoglu and Autor (2011) technology destroys occupation in the middle part of the wage distribution substituting repetitive and routinised tasks

Existing literature

- a few proxies have been used to measure the impact of technology on the labour market
  - share of computers in sectors of belonging (Autor, Levy, Murnane, 2003)
  - share of robots in sectors of belonging (Acemoglu and Restrepo, 2020)
  - automation probability constructed via Delphi method (experts judgment) and classifier systems (Arntz et al., 2016; Frey and Osborne, 2017; Nedelkoska and Quintini, 2018)

- these are all *indirect* measures which might confound firms and industry attributes and heterogeneous technological artefacts
- a more *direct* machine-task mapping is still missing

Existing literature (cont'd)

Webb (2020) proposes a direct measure of exposure via co-occurrence of verb-noun pairs in the **title** of AI patents and O*NET tasks

BUT

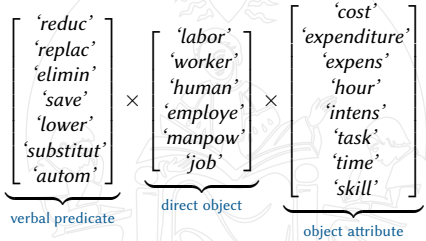
- titles of patents are hardly informative of the underlying functions executed by the technological artefact
- restricting to verb-noun pairs has high likelihood of false positives (e.g. [`develop`, `grid`], [`identify`, `area`])
- the measure of exposure is not constructed in terms of overall similarity of the two corpora
- arbitrary selection of relevant technologies (e.g. '`neural network`' only in titles)
- no way of distinguishing labour-saving from labour-augmenting technologies

Our starting point

Montobbio et al. (2020) identify labour-saving patents among USPTO robotic applications (2009–2018)

1. robotics patents identified by CPC and keyword search ($10 \times$ `'robots'`)
2. labour-saving patents identified by text query and manual validation (no false positives)

- 1,276 *truly* labour-saving patents



Examples of labour-saving patents

"Automated systems, such as robotic systems, are used in a variety of industries to ***reduce labo[u]r costs and/or increase productivity****. Additionally, the use of human operators can involve increased cost relative to automated systems."* [US20170178485A1]

"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."* [US20100223134A1]

Objective of the paper

- build a *direct* measure of occupational exposure to labour-saving technologies

- to study the machine-task relationship we need to look at functions and operations of both machines, in relation to humans, and humans themselves
- functions and operations are better described in CPC definitions than in patents full-text
- tasks executed by humans are well described in the O*NET questionnaire
- technological classification codes allow us to pinpoint truly labour-saving tasks

Outline

Occupations and labour-saving patents

- we compute a text similarity measure between technological codes and tasks

CPC corpus

- technological definitions from CPC `v.2019.08`
- 671 4-digit CPC codes

Task–Occupation corpus

- tasks description from O*NET `v.25.1`
- 19,231 tasks mapped to 923 8-digit SOC2018 occupations

preprocessing: every piece of text is *tokenised*, *stemmed*, and *stop words* are removed

Document-term matrix

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

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

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

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

   - 671×2309 matrix

Cosine similarity

2. construct the *document-term matrix* $\mathcal{D}_{ONET}$ of the corpus of task descriptions
   - projected on the *vocabulary* of the CPC matrix $\mathcal{D}_{CPC}$
   - 19231×2309 matrix
3. construct the cosine similarity (CS) measure between the two corpora
   - for each couple of row vectors $X \in \mathcal{D}_{CPC}$, $Y \in \mathcal{D}_{ONET}$ $\quad (X, Y \in \mathbb{R}_{+}^{2309})$

$$\cos(X, Y) := \frac{X \cdot Y}{\|X\| \, \|Y\|} = \frac{\sum_t x_t y_t}{\sqrt{\sum_t x_t^2} \sqrt{\sum_t y_t^2}}$$

   - $\cos(X, Y) \in [0, 1]$ since vectors X and Y are non-negative valued
   - w.r.t. Euclidean distance, cosine similarity normalises for varying lengths of documents
   - 671×19231 cosine similarity matrix[1]
   - each task obtains a similarity score to each CPC code (12,904,001 pairs)

[1] under tf-idf, it is possible to show that $\cos(\mathcal{D}_{CPC}, \mathcal{D}_{ONET}) \equiv \mathcal{D}_{CPC} \cdot \mathcal{D}'_{ONET}$

Cosine similarity (cont'd)

OCCUPATION	11-1011.00		...		...	53-7121.00	
TASK	8823	8824	...	...	...	12809	12810
CPC							
A01B	cos(A01B,8823)	cos(A01B,8824)	...	...	...	cos(A01B,12809)	cos(A01B,12810)
A01D	cos(A01D,8823)	cos(A01D,8824)	...	...	...	cos(A01D,12809)	cos(A01D,12810)
...	...	...	...	...	...	...	...
H05H	cos(H05H,8823)	cos(H05H,8824)	...	...	...	cos(H05H,12809)	cos(H05H,12810)
H05K	cos(H05K,8823)	cos(H05K,8824)	...	...	...	cos(H05K,12809)	cos(H05K,12810)

4. weight by CPC frequency in LS patents[2], sum across CPCs, and rescale between [0,1]

[2] codes B25*, G01*, G05*, G06*, and Y* are excluded because too general

From tasks to occupations

- each O*NET occupation consists of a number of *core* and *supplementary* tasks
- we attribute task CS to occupations with weights

$$\text{core}: \quad \frac{2/3}{\#\text{ tasks in the occupation}}$$

$$\text{supplementary}: \quad \frac{1/3}{\#\text{ tasks in the occupation}}$$

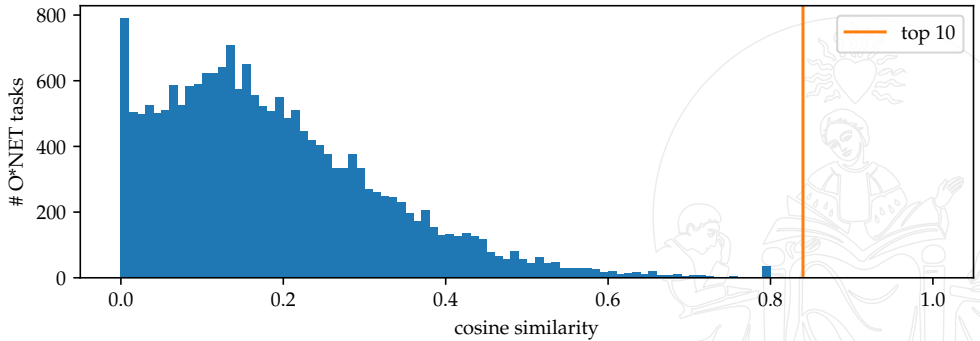
  - this weighting scheme reflects O*NET cutoff between core and supplementary tasks (based on a blend of *frequency*, *importance*, and *relevance* to underlying occupation)

Outline

Top tasks by similarity

#	Code	Description	CS
1	14587	Load materials and products into machines and equipment, or onto conveyors, using hand tools and moving devices	1.0
2	3202	Move levers or controls that operate lifting devices, such as forklifts, lift beams with swivel-hooks, hoists, or elevating platforms, to load, unload, transport, or stack material	0.96
3	3203	Position lifting devices under, over, or around loaded pallets, skids, or boxes and secure material or products for transport to designated areas	0.9
4	17928	Lift and move loads, using cranes, hoists, and rigging, to install or repair hydroelectric system equipment or infrastructure	0.89
5	15266	Manually or mechanically load or unload materials from pallets, skids, platforms, cars, lifting devices, or other transport vehicles	0.88
6	14584	Remove materials and products from machines and equipment, and place them in boxes, trucks or conveyors, using hand tools and moving devices	0.86
7	11839	Transport machine parts, tools, equipment, and other material between work areas and storage, using cranes, hoists, or dollies	0.85
8	3217	Load materials and products into package processing equipment	0.85
9	12805	Operate conveyors and equipment to transfer grain or other materials from transportation vehicles	0.85
10	12323	Communicate with systems operators to regulate and coordinate line voltages and transmission loads and frequencies	0.84

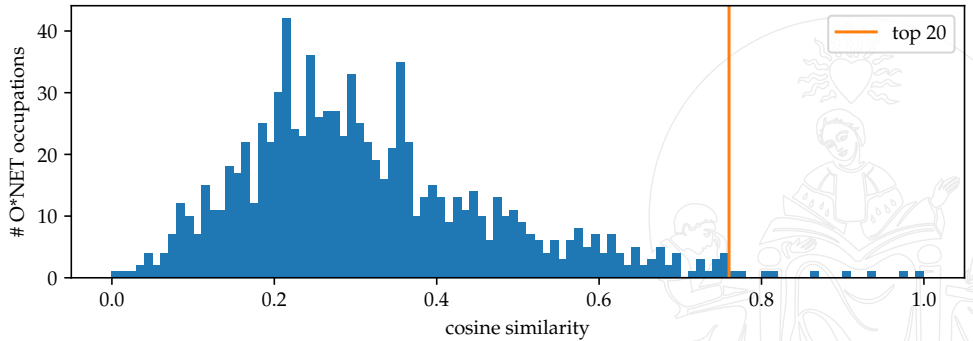
Tasks by similarity



Top occupations by similarity

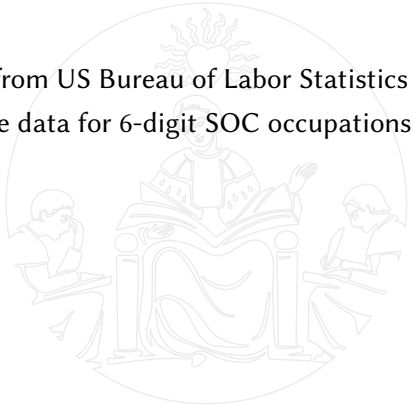
#	Code	Title	CS
1	53-7051.00	Industrial Truck and Tractor Operators	1.0
2	49-9043.00	Maintenance Workers, Machinery	0.97
3	53-7063.00	Machine Feeders and Offbearers	0.94
4	53-7064.00	Packers and Packagers, Hand	0.91
5	49-2091.00	Avionics Technicians	0.87
6	51-9111.00	Packaging and Filling Machine Operators and Tenders	0.81
7	49-3041.00	Farm Equipment Mechanics and Service Technicians	0.81
8	49-3092.00	Recreational Vehicle Service Technicians	0.78
9	49-3042.00	Mobile Heavy Equipment Mechanics, Except Engines	0.77
10	47-2111.00	Electricians	0.76
11	49-9098.00	Helpers–Installation, Maintenance, and Repair Workers	0.75
12	49-9041.00	Industrial Machinery Mechanics	0.75
13	51-9082.00	Medical Appliance Technicians	0.75
14	47-3011.00	Helpers–Brickmasons, Blockmasons, Stonemasons, and Tile and Marble Setters	0.75
15	51-9191.00	Adhesive Bonding Machine Operators and Tenders	0.75
16	51-9023.00	Mixing and Blending Machine Setters, Operators, and Tenders	0.74
17	13-1032.00	Insurance Appraisers, Auto Damage	0.73
18	51-4111.00	Tool and Die Makers	0.73
19	49-9081.00	Wind Turbine Service Technicians	0.72
20	51-8013.04	Hydroelectric Plant Technicians	0.72

Occupations by similarity

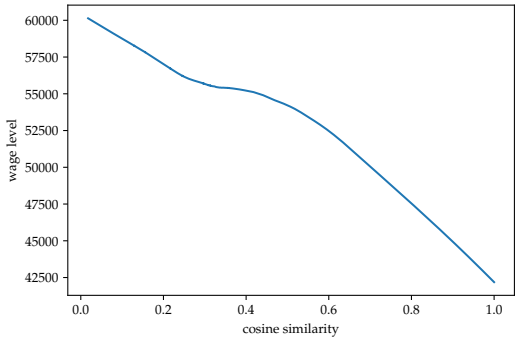


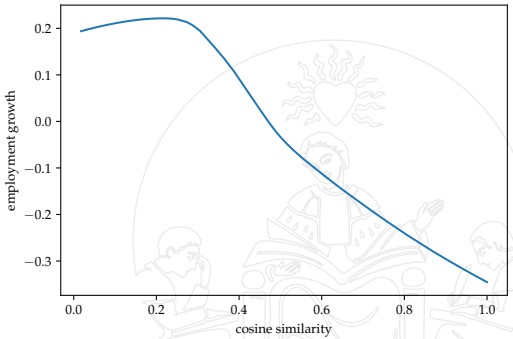
Occupational exposure and employment

- match with Occupational Employment Statistics (OES) from US Bureau of Labor Statistics
- employment (excluding self-employed) and median wage data for 6-digit SOC occupations
  - 2019 for levels
  - 1999 for 20-year growth rates



Wage levels and employment growth



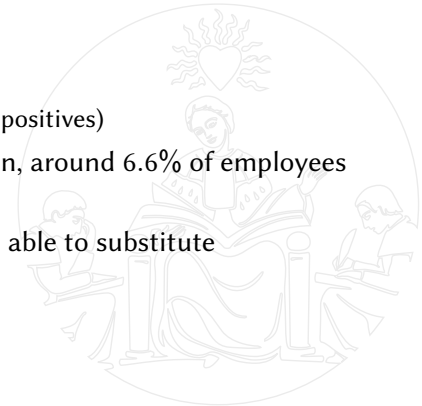


- robust LOWESS estimates of the underlying scatter plots (bandwidth $= 0.8$)

Outline

Discussion

- the cosine similarity matrix is overall very *sparse*
  - skewed distributions in both tasks and occupations
  - high similarity is a **rare event** (low probability of false positives)
- considering the top quartile of the similarity distribution, around 6.6% of employees (≈10m) are exposed to substitution
- we do not know how many workers a single machine is able to substitute



Discussion (cont'd)

- exposure to substitution is monotonically decreasing in wage
  - no U-shaped pattern but rather a negative declining relationship
- most affected occupations (2-digit) include "transportation and material moving" (logistics), "installation, maintenance, and repair" (automotive), "food preparation and serving"
- exposure to substitution is decreasing in employment growth
  - innovative efforts towards the weakest and cheapest segment of the labour market.

Future developments

- extend to overall labour-saving technologies (beyond robotics)
- extend from occupations to sectors of economic activity
- investigate of determinants of employment change
  - wage vs. technological exposure
- characterise labour-saving robotic patents in terms of quality and innovative content
  - since innovative efforts are devoted to substitute cheap labour

Thank you very much!

jacopo.staccioli@unicatt.it

this presentation is available at www.staccioli.org

Patent-O*NET match: top tasks by similarity

#	Code	Description	CS
1	16596	Build or assemble robotic devices or systems	1.0
2	11944	Set up and operate computer-controlled machines or robots to perform one or more machine functions on metal or plastic workpieces	0.98
3	21057	Build, configure, or test robots or robotic applications	0.97
4	16523	Conduct research on robotic technology to create new robotic systems or system capabilities	0.93
5	16511	Provide technical support for robotic systems	0.91
6	16587	Assist engineers in the design, configuration, or application of robotic systems	0.86
7	16525	Conduct research into the feasibility, design, operation, or performance of robotic mechanisms, components, or systems, such as planetary rovers, multiple mobile robots, reconfigurable robots, or man-machine interactions	0.84
8	16593	Install, program, or repair programmable controllers, robot controllers, end-of-arm tools, or conveyors	0.81
9	16584	Modify computer-controlled robot movements	0.8
10	16579	Maintain service records of robotic equipment or automated production systems	0.8

Patent-O*NET match: top occupations by similarity

#	Code	Title	CS
1	17-2199.08	Robotics Engineers	1.0
2	17-3024.01	Robotics Technicians	0.96
3	47-2231.00	Solar Photovoltaic Installers	0.49
4	17-2072.01	Radio Frequency Identification Device Specialists	0.46
5	15-1299.08	Computer Systems Engineers/Architects	0.45
6	15-1299.02	Geographic Information Systems Technologists and Technicians	0.42
7	51-9161.00	Computer Numerically Controlled Tool Operators	0.41
8	17-2199.11	Solar Energy Systems Engineers	0.4
9	49-2091.00	Avionics Technicians	0.39
10	15-1243.01	Data Warehousing Specialists	0.38
11	17-1022.01	Geodetic Surveyors	0.38
12	15-1244.00	Network and Computer Systems Administrators	0.38
13	17-2061.00	Computer Hardware Engineers	0.37
14	15-1299.03	Document Management Specialists	0.37
15	15-1211.00	Computer Systems Analysts	0.36
16	51-4034.00	Lathe and Turning Machine Tool Setters, Operators, and Tenders, Metal and Plastic	0.36
17	17-2041.00	Chemical Engineers	0.36
18	49-9044.00	Millwrights	0.36
19	15-2051.02	Clinical Data Managers	0.36
20	17-3021.00	Aerospace Engineering and Operations Technologists and Technicians	0.35