

Covid resilience and digital readiness: An analysis using online company data

Elodie Andrieu*, **Oleksii Romanko**[✉], Mary O'Mahony[✉],

* King's College London

[✉]Economic Statistics Centre of Excellence

-
- **Introduction** - heterogeneous firms' response to COVID restrictions
 - **The study setup** - which factors determine a resilience of a company. Data sources
 - **Methodology** - how do we 'mine' the data and extract meaningful information
 - **Preliminary findings** - whether covid resilience is correlated with digital readiness
-

- Government restrictions (closure of non-essential activities, work from home and lockdowns):
 - lead to digitalization/'innovation' (Bai et al., 2020; Dingel and Neiman, 2020).
 - 'stickers' applied to fairlough schemes and did not invest (Baker et al., 2016)
- IT industries survived 1st covid wave better (BICS,ONS), Innovative companies have higher survival rates (Nures and Lopes, 2013; Jin and McElheran (2018)). **We think that companies with better technical capabilities and technologically skilled human capital are better able to adapt and survive during and after the COVID-19 crisis.**
- Covid pandemic has driven the use of online data, e.g.: transaction data, mobility data, COVID-tracker, labour market data (BGT, Adzuna) etc. (Adams-Prassl, 2021, "*Real Time Research*" seminar)

The study setup

Which factors determine company's resilience?

- financial background(Ding et al., 2020; Song et al. (2020), Joseph et al., 2020).
- size of the firm matters but have heterogeneous effect (Guo et al. (2020));
- pre-crisis technological readiness is an important factor:
 - based on size and industry we may observe continuous improvement or rapid innovation (Archibugi et. al (2012)). Innovation is procyclical in the UK
 - Innovation: Cloud facilitates remote journey,rapid transformation and analytics*
 - Already online firms are less impacted (Bloom et al., 2021, Arighetti et. al., 2016)
 - Availability of skilled labour is a fuel of change
- Other unaccounted factors? (management practices, Lamorgese et al., 2021)

* [https://www.thelancet.com/journals/landig/article/PIIS2589-7500\(20\)30291-0/fulltext](https://www.thelancet.com/journals/landig/article/PIIS2589-7500(20)30291-0/fulltext)

<https://www.infoworld.com/article/3614809/the-covid-pandemics-lasting-impact-on-cloud-usage.html>

The study setup (cont.)

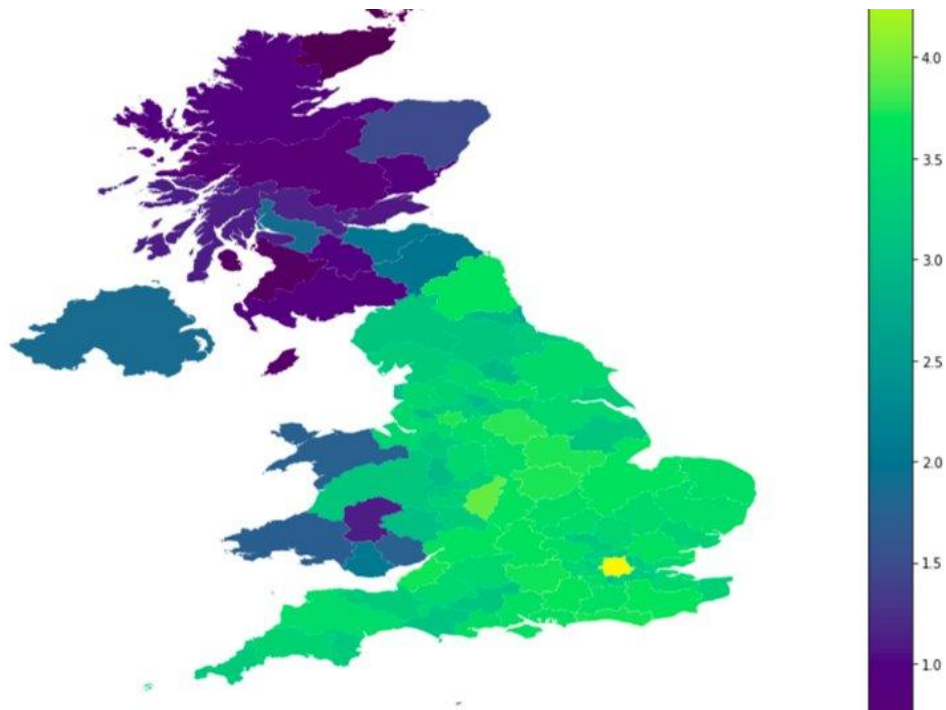
Resilience indicator	Resilience measure is based on scraping the websites of approximately 130,000 UK firms monthly since May 2020 Source: companies' web sites
Cloud indicator	Cloud indicator is used as a proxy to the usage of innovative technologies, monthly data from 2008 and ongoing. Source: DNS metadata
Tech hiring indicator	Occupational and skill patterns in hiring prior and during the COVID-19 crisis, monthly data from 2012. Source: BGT
Financial viability	Background information about businesses and their financial performance, yearly from 2008 Source: FAME dataset.

Methodology: FAME dataset

Financial Analysis Made Easy (FAME) dataset, that covers a population of businesses in the UK and derives information from Companies House records

Subset of 130,000 firms over the period from 2008 till 2020

There are financial indicators, employment statistics and other important indicators of business performance in the dataset



Number of companies depending on region (logs)

1. Take the list of companies from FAME (Financial Analysis Made Easy) dataset.

2. Collect main page from their websites using web scraping techniques (on monthly basis, starting from March 2020)

3. Transform and cleanse html and arrive at clean text version of the website

4. Automated content analysis (Nunez-Mir et al., 2016)
4a) manual keywords detection

Innovators: firms that actively react to COVID crisis

(continue to operate as usual, continue to provide services, office is open, online delivery)

Informers: firms that inform about the current situation and wait until things come normal

('close our offices', 'closed due to COVID 19', 'activities are cancelled')

Undecided: third group of firms who do not react

No COVID-related information



E S
C O
E
ECONOMIC
STATISTICS
CENTRE
OF
EXCELLENCE

4c) use machine learning model (Latent Dirichlet Allocation) to cluster websites text into three groups (Stephany et. al, 2020; Yang and Han, 2020; Kinne et. al (2020))

covid cases

covid update

Vaccination

covid test

long covid

Symptom

covid news

covid tracker

attestation covid

covid symptoms

covid 19

update

covid deaths

covid testing

covid 19 cases

[illegible][illegible]

7

We use the indicator of cloud technologies as a proxy to firm's innovation process

1. Use the same list of companies from FAME.
2. Do a parsing of Domain Name Server (DNS) records, find domain vendors using Security Trails API
3. Scrape websites of Domain vendors
4. Extract cloud related words
5. Assign cloud indicator to website hosting providers
6. Assign cloud indicator to companies based on vendors they use
7. Construct a panel data



Cloud-related words

Idea: appropriate occupations/skills to adapt 1) your work (remote work) and 2) the activity of the firm (e.g. shift to online services)

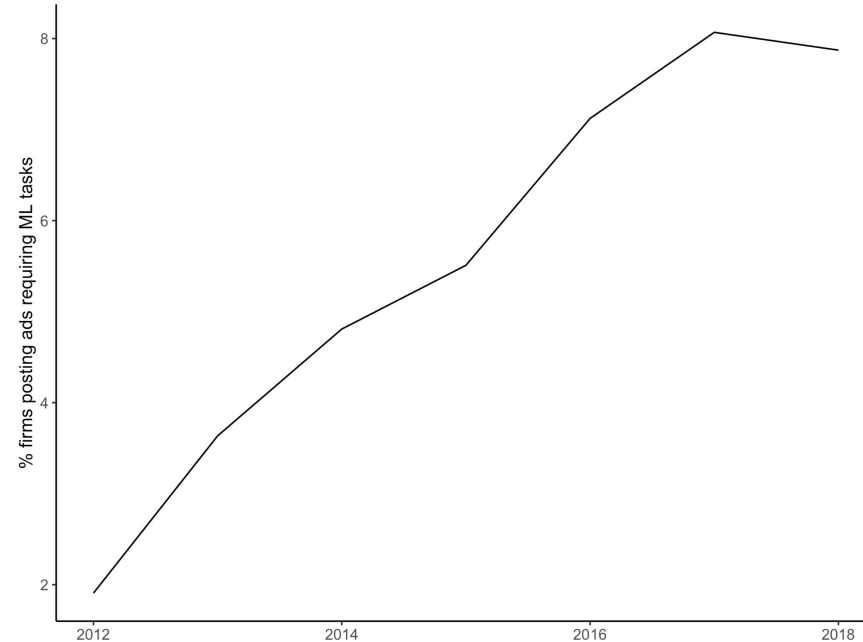
→ Create 2 indicators using the information in online job vacancies from BGT at the firm level

1. Occupation dimension:

- Use “techies” definition (Harrigan et al., 2018): occupations related to R&D and ICT adoption

2. Skill dimension:

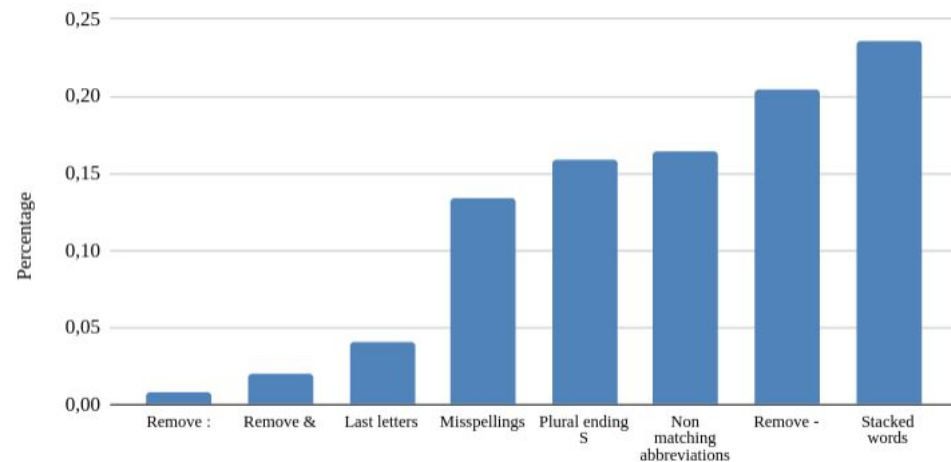
- AI-related skills has grown across all industry sectors and occupation groups (Alekseeva et al., 2020)
- Use skill content of the ad, identify “AI skills” (*based on various already existing classification: Deming et al., OECD etc.*)



Issues:

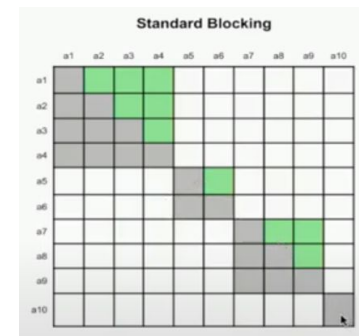
- Manually entered information: human factor, misspellings, grammatical mistakes
- Thus, we have used some NLP techniques to do a 'fuzzy matching' of companies' names across datasets.
- Complexity of the data linkage is $n*m$. Unrealistic (timing and computing resources; ~1,2 trillion 1s each = 38 051 years on a single machine).

Frequent typos in companies' names

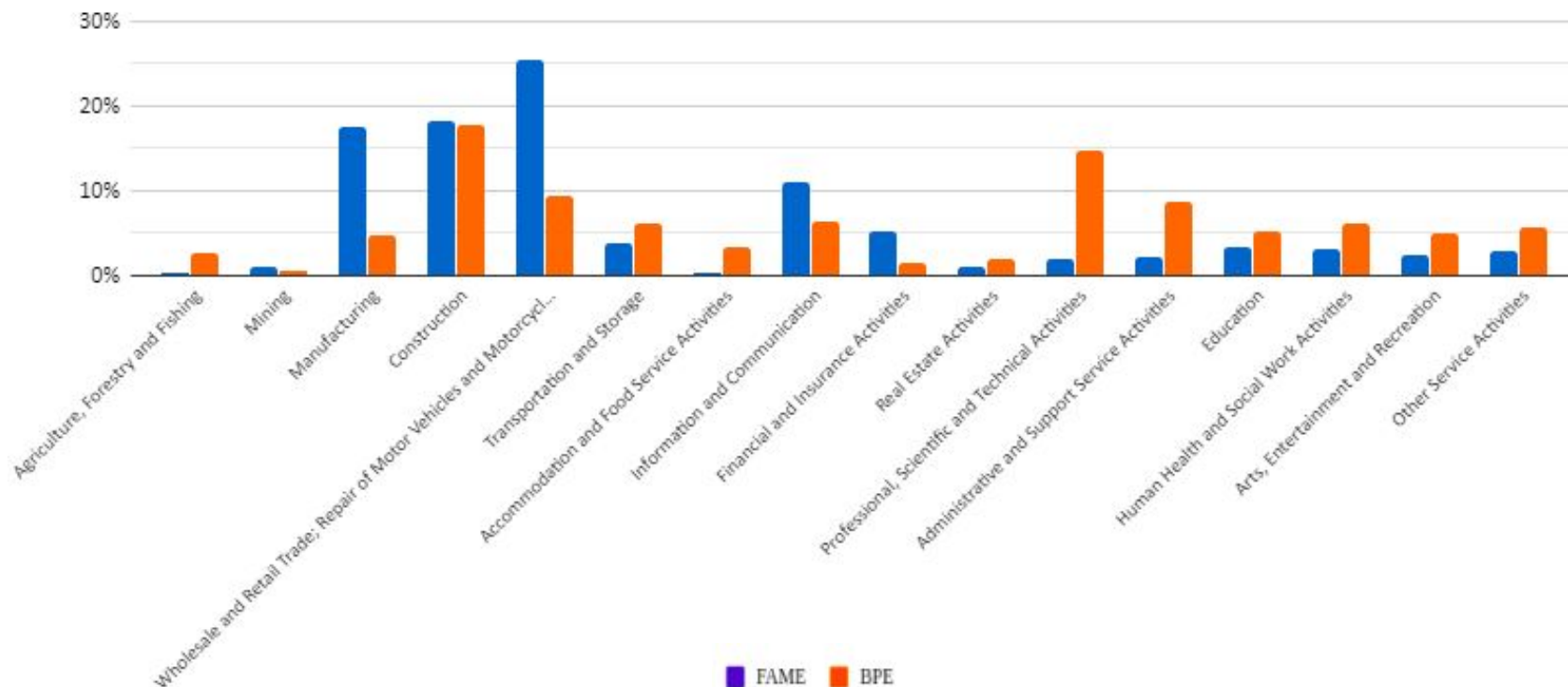


Solution - blocking and sorted neighborhood indexing (Christen, 2011):

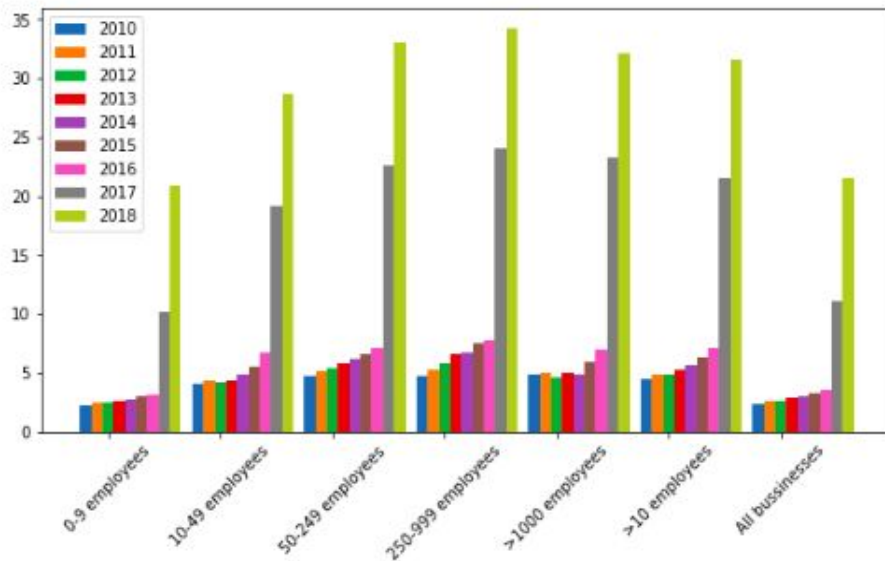
- Extract the first five letters from the company name. Create blocks.
- Combine and sort records in the block.
- N-records moving window comparison catches most of misspellings
- Iteratively rename companies pairs, (A-B, B-C, C-D) -> A.



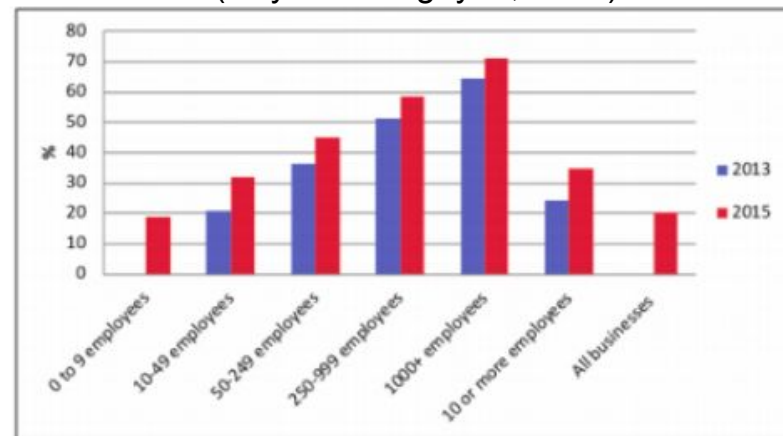
Share of industries in our subset versus Business Population Estimates



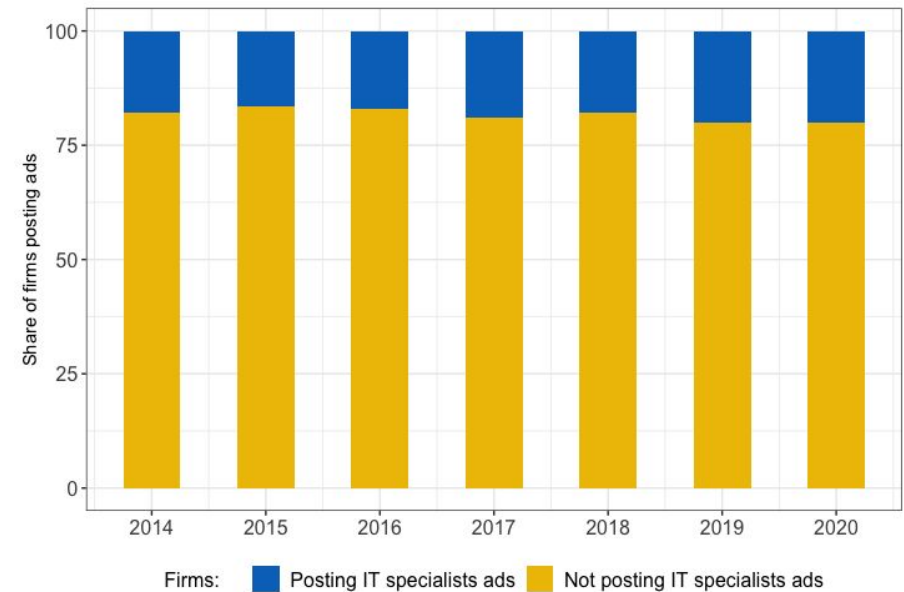
Share of cloud users, own data



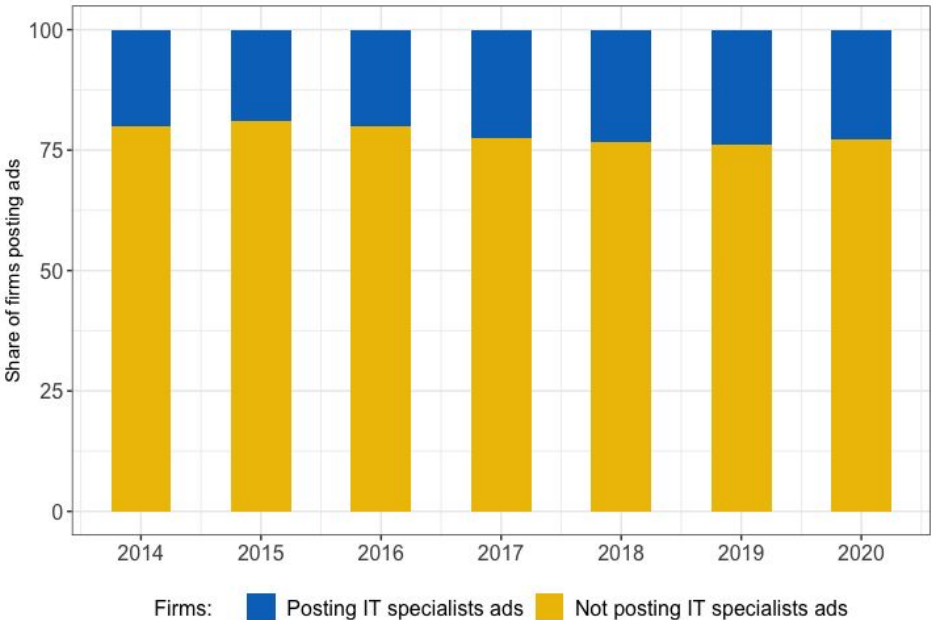
Share of cloud users, ONS e-commerce survey
(Coyle and Nguyen, 2018)



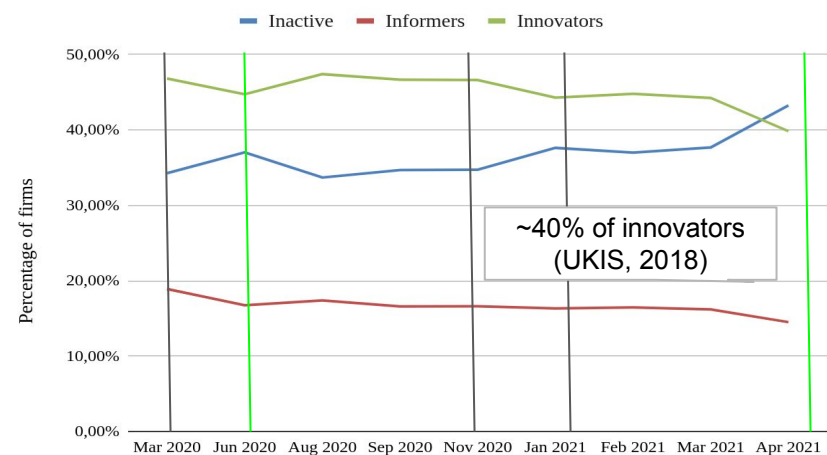
Panel (a): all firms in BGT



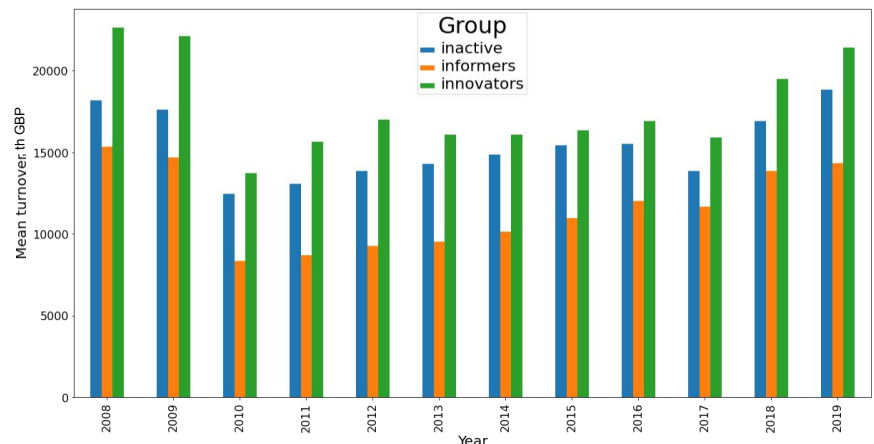
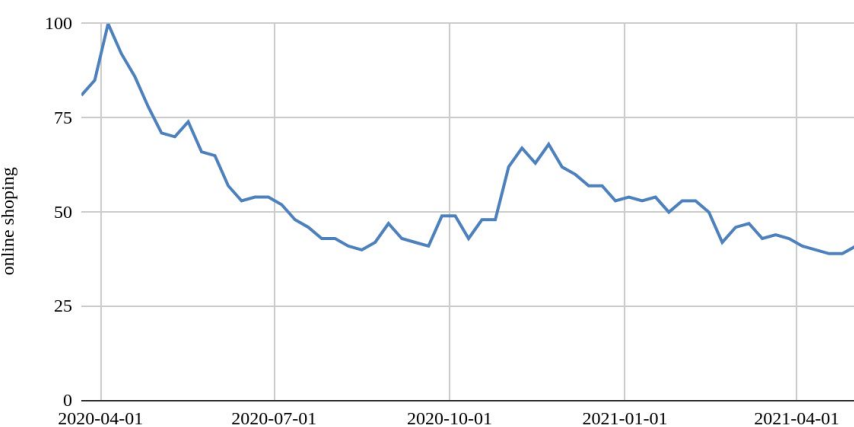
Panel (b): matching sample with FAME



Descriptives & Preliminary findings



Google Trends



Questions and hypotheses

- Hypothesis:
 - Our digital readiness indicators is an explanatory variable for our resilience measure, i.e. importance of hiring IT specialists, AI skilled workers and to use cloud computing ahead of the crisis
 - Financial situation of firms is also necessary for the innovative activity of the firm during the crisis
 - To what extent the pre-crisis digital readiness is important? Can firms be resilient by actively hiring the necessary workers and start using new technologies during the crisis?
 - To what extent is the technological implementation complemented with labour hiring?
- Preliminary findings:
 - hiring IT specialists ahead of the crisis is positively correlated with being innovator firm (logistic regressions show significant coefficient)
 - Financial variable: also positive (not significant)
 - Cloud variable: significantly negative

Conclusions

- Importance for the research: shift from passive advice to proactive advocacy (Greene, Rosiello 2020) starts with timely and accessible data sources
- Policy implications: help firms to access digital resources and support R&D grants during and **before** downturns *

Further work

- Sectoral analysis, timing of hiring, innovation and success
- Skills analysis

And thank you!

* <https://www.gov.uk/government/news/billion-pound-support-package-for-innovative-firms-hit-by-coronavirus>
https://ec.europa.eu/info/sites/default/files/research_and_innovation/strategy_on_research_and_innovation/documents/ec_rtd_covid19-recovery.pdf