

Essays on the Application of

Statistical Learning

in

Empirical Economic Research

Julian Dörr^{1,2}

¹Justus Liebig University Giessen

²former: ZEW - Leibniz Centre for European Economic Research

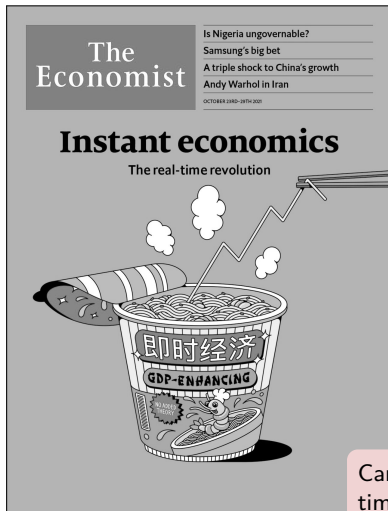
October 30, 2022

An integrated data framework for policy guidance during the coronavirus pandemic: Towards real-time decision support for economic policymakers

joint with Jan Kinne, David Lenz, Georg Licht, Peter Winker

published in: PLoS ONE

Motivation - lack of real-time economic data



Source: The Economist (2021a)

Does anyone really understand what is going on in the world economy? The pandemic has made plenty of observers look clueless.

Especially in times of rapid change, policymakers have operated in a fog.

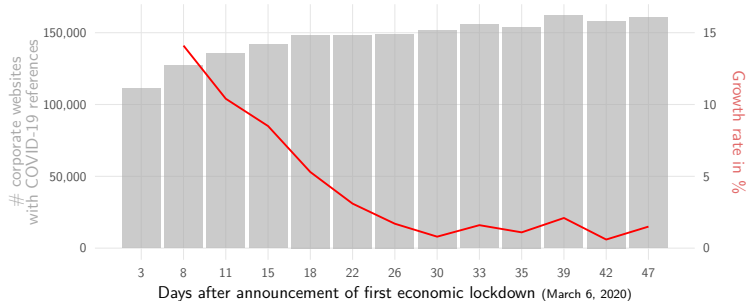
The gap between official data and what is happening in the real economy can still be glaring.

The Economist (2021a, 2021b)

Can we assist policy makers with **timely** and **insightful** firm level data in times of dynamic economic shocks such as COVID-19?

Early firm communication and corporate websites

- ▶ accessed corporate websites of ~ 1.18M German companies from Mar 20 - May 20 twice a week searching for references related to the pandemic
- ▶ finding: companies used their websites intensively to report about the pandemic



Turn website references into knowledge

But: context of Corona references greatly differed across firms:

'The Corona pandemic is not only affecting ongoing [REDACTED] projects, but also the current selection rounds of the 13th and 14th funding seasons.'

* * *

'We have therefore decided to adapt our services to the current situation and to limit them until further notice. Although we want to continue to provide you with all indispensable services, we also want to meet the recommendations of the federal government on how to deal with the corona virus.'

* * *

'Your [REDACTED] advisor stands by your side - also in times of COVID-19.'

Turn website references into knowledge

But: context of Corona references greatly differed across firms:

'The Corona pandemic is not only affecting ongoing [REDACTED] projects, but also the current selection rounds of the 13th and 14th funding seasons.'

* * *

'We have therefore decided to adapt our services to the current situation and to limit them until further notice. Although we want to continue to provide you with all indispensable services, we also want to meet the recommendations of the federal government on how to deal with the corona virus.'

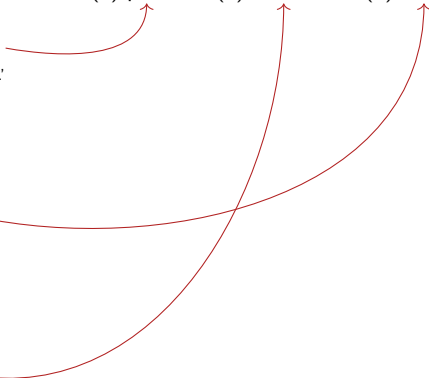
* * *

'Your [REDACTED] advisor stands by your side - also in times of COVID-19.'

Statistical learning approach:

1. introduced 5 meaningful & distinguishable classes

(1) problem, (2) confidence, (3) adaption, (4) information, (5) unclear



Turn website references into knowledge

But: context of Corona references greatly differed across firms:

'The Corona pandemic is not only affecting ongoing [REDACTED] projects, but also the current selection rounds of the 13th and 14th funding seasons.'

* * *

'We have therefore decided to adapt our services to the current situation and to limit them until further notice. Although we want to continue to provide you with all indispensable services, we also want to meet the recommendations of the federal government on how to deal with the corona virus.'

* * *

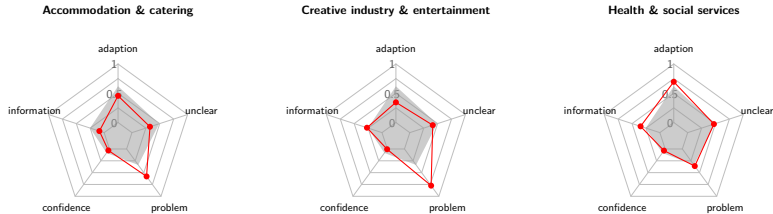
'Your [REDACTED] advisor stands by your side - also in times of COVID-19.'

Statistical learning approach:

1. introduced 5 meaningful & distinguishable **classes**
(1) problem, (2) confidence, (3) adaption, (4) information, (5) unclear
2. manually annotated ~ 4,000 references
3. fine-tuned pre-trained **language model** (XLM-R by Conneau et al. (2019))

Early insights from website analysis

- ▶ classified *firm level* communication on websites revealed impact heterogeneity at *sector level*
- ▶ insights generated at near-real time (right after shutdown announcement in Mar 20)



Classified website references as leading indicators

for later credit rating updates

$$\Delta r_{i,\bar{t}+z} = \alpha + \beta_1 \text{Problem}_{i,\bar{t}} + \beta_2 \text{Confidence}_{i,\bar{t}} + \beta_3 \text{Adaption}_{i,\bar{t}} + \beta_4 \text{Information}_{i,\bar{t}} + \beta_5 \text{Unclear}_{i,\bar{t}} + \gamma r_{i,\bar{t}-x} + \delta FE_i + \epsilon_i$$

	(1) $\Delta r_{\bar{t}+z}$	(2) $\Delta r_{\bar{t}+z}$	(3) $\Delta r_{\bar{t}+z}$	(4) $\Delta r_{\bar{t}+z}$
Problem $_{\bar{t}}$	+1.66***	+1.68***	+1.62***	+0.42**
Confidence $_{\bar{t}}$	-1.70***	-1.69***	-1.73***	-0.69
Adaption $_{\bar{t}}$	-0.46***	-0.47***	-0.33***	-0.13
Information $_{\bar{t}}$	-0.24***	-0.24***	-0.23***	-0.17***
Unclear $_{\bar{t}}$	-0.42***	-0.42***	-0.10	-0.08
$r_{\bar{t}-x}$	-0.09***	-0.10***	-0.11***	-0.13***
Age FE	No	Yes	Yes	Yes
Size FE	No	No	Yes	Yes
Sector FE	No	No	No	Yes
N	61,228	61,138	57,343	57,343

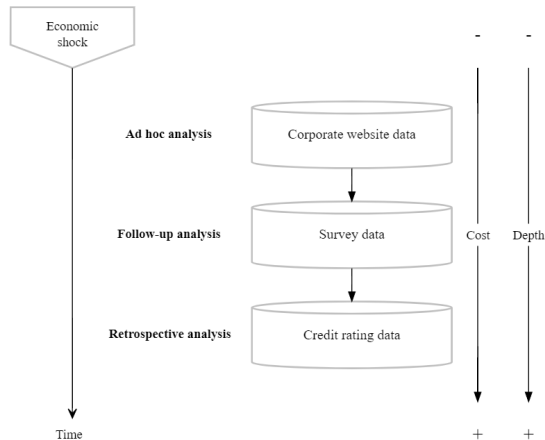
Δr_i : credit rating update (downgrade, upgrade) of firm i

\bar{t} : 01/03/20 - 31/05/20, $\bar{t} + z$: z days after 01/06/20, $\bar{t} - x$: x days before 01/03/20

FE: fixed effects. Significance levels: *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$

Main contributions

- ▶ **proposed a data framework for policy guidance** in times of economic shocks
 - Follow-up surveys
 - Outcome analysis
- ▶ to **overcome information deficits** policy makers are confronted with in highly dynamic situations
- ▶ possibly allowing more targeted liquidity injections to support affected companies instead of choosing the 'bazooka' as policy instrument



- Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., & Stoyanov, V. (2019). Unsupervised cross-lingual representation learning at scale. arXiv 1911.02116.
<https://doi.org/10.48550/arXiv.1911.02116>
- The Economist. (2021a). A real-time revolution will up-end the practice of macroeconomics.
- The Economist. (2021b). Enter third-wave economics. Retrieved October 25, 2021, from
<https://www.economist.com/briefing/2021/10/23/enter-third-wave-economics>

Appendix

Data Framework for Policy

Classification of COVID-19 web references

Categories	Description	Examples (translated)
Problem	Firm reports about adverse impacts of the pandemic on its business operations.	<p>Due to the Corona pandemic, [REDACTED] & [REDACTED] are closed.</p> <p>[REDACTED] has been cancelled due to the increasing concerns and escalated circumstances surrounding the recent coronavirus (COVID-19) outbreak.</p>
Confidence	Firm indicates that the pandemic has no negative impacts on its business operations.	<p>We are there for you 24/7 as usual despite Corona!</p> <p>Your [REDACTED] advisor stands by your side - also in times of COVID-19.</p>
Adaption	Firm reports that it is adapting to the new economic circumstances.	<p>We have also upgraded our IT and telecommunications system. Our employees are now also able to ensure that you are looked after from home, should this be necessary. Since we receive new information on the development of the coronavirus, the measures and the safety precautions every day, we will continue to monitor the development and react to it.</p> <p>Within our emergency opening times, we particularly take care of those who are currently performing at their best for our society in view of the coronavirus crisis and who depend on their glasses for their work.</p>
Information	Firm reports generally, not necessarily in a business-context, about the pandemic.	<p>The corona pandemic affects each of us now and in the near future. There are many uncertainties and resulting (insurance) issues. What about entitlement to holiday cancellations, health protection abroad and coverage in the event of business interruption are just a few of the questions.</p> <p>In cooperation with the software provider [REDACTED], the Bundesverband Pflegemanagement (Federal Association of Care Management) is launching a platform to recruit former care professionals to cope with the currently dramatic challenges facing care against the background of the Corona crisis.</p>
Unclear	COVID-19 reference does not come with further clearly distinguishable content.	<p>Current situation COVID-19.</p> <p>COVID-19 and how it affects us.</p>

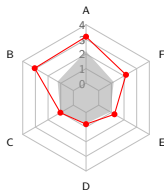
Add a slide with details about the language model

Follow up surveys

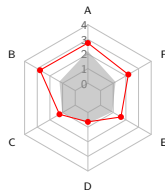
- ▶ based on the early findings, construct targeted businesses surveys
- ▶ gain more detailed understanding about the sort of impact in order to design counter measures most effectively
- ▶ here: surveyed ~ 1,500 companies consecutively (Apr, Jun, Sep 2020) with targeted impact questions

Figure: Targeted impact questions at sector level

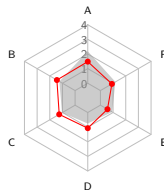
Accommodation & catering



Creative industry & entertainment



Health & social services



A: Drop in demand
D: Staffing shortages

B: Temporary closing
E: Logistical sales problems

C: Supply chain interruption
F: Liquidity shortfalls

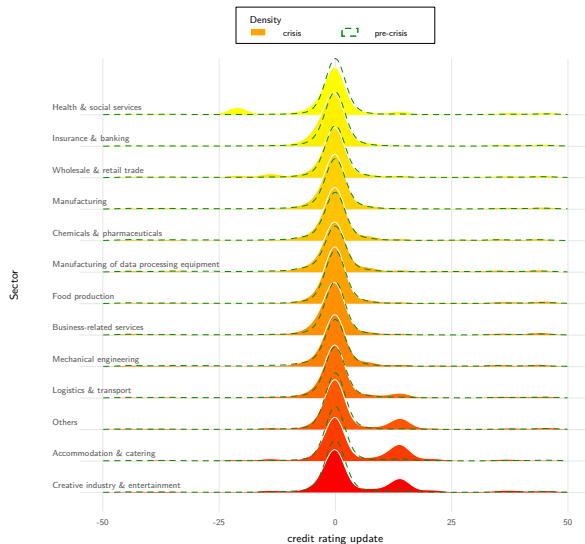
Retrospective analysis of firm outcomes

- ▶ after economic shock has materialized in economy, analyze firm outcomes
- ▶ understand possible long-term consequences and design stimulus programs
- ▶ here: examined credit rating updates of ~ 870,000 companies (between Jun 20 - Apr 21)

Retrospective analysis of firm outcomes

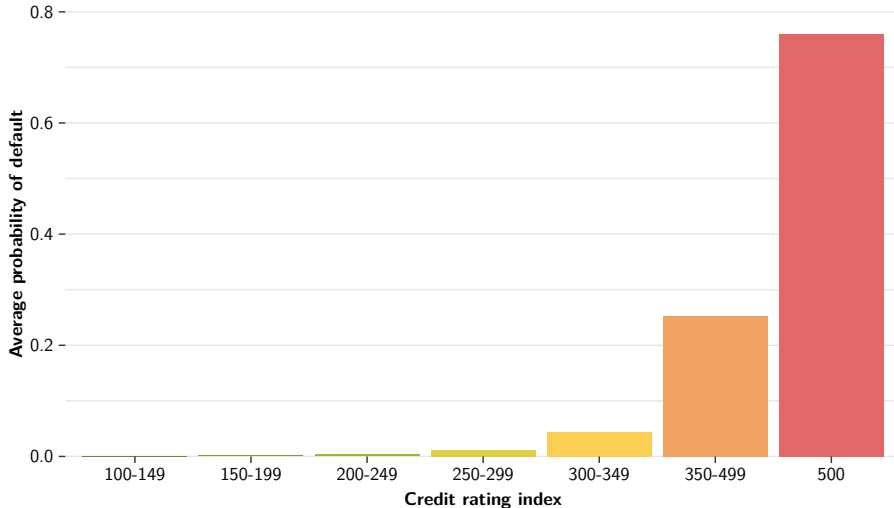
- ▶ after economic shock has materialized in economy, analyze firm outcomes
- ▶ understand possible long-term consequences and design stimulus programs
- ▶ here: examined credit rating updates of ~ 870,000 companies (between Jun 20 - Apr 21)

Figure: Credit rating movements at sector level



Credit Rating Data

Commonly used by banks (probability of default of debtors) and by research (insolvency risk estimation)



Source: Creditreform