



ORRG

Open and Reproducible
Research Group

Participation in Open Science Practices

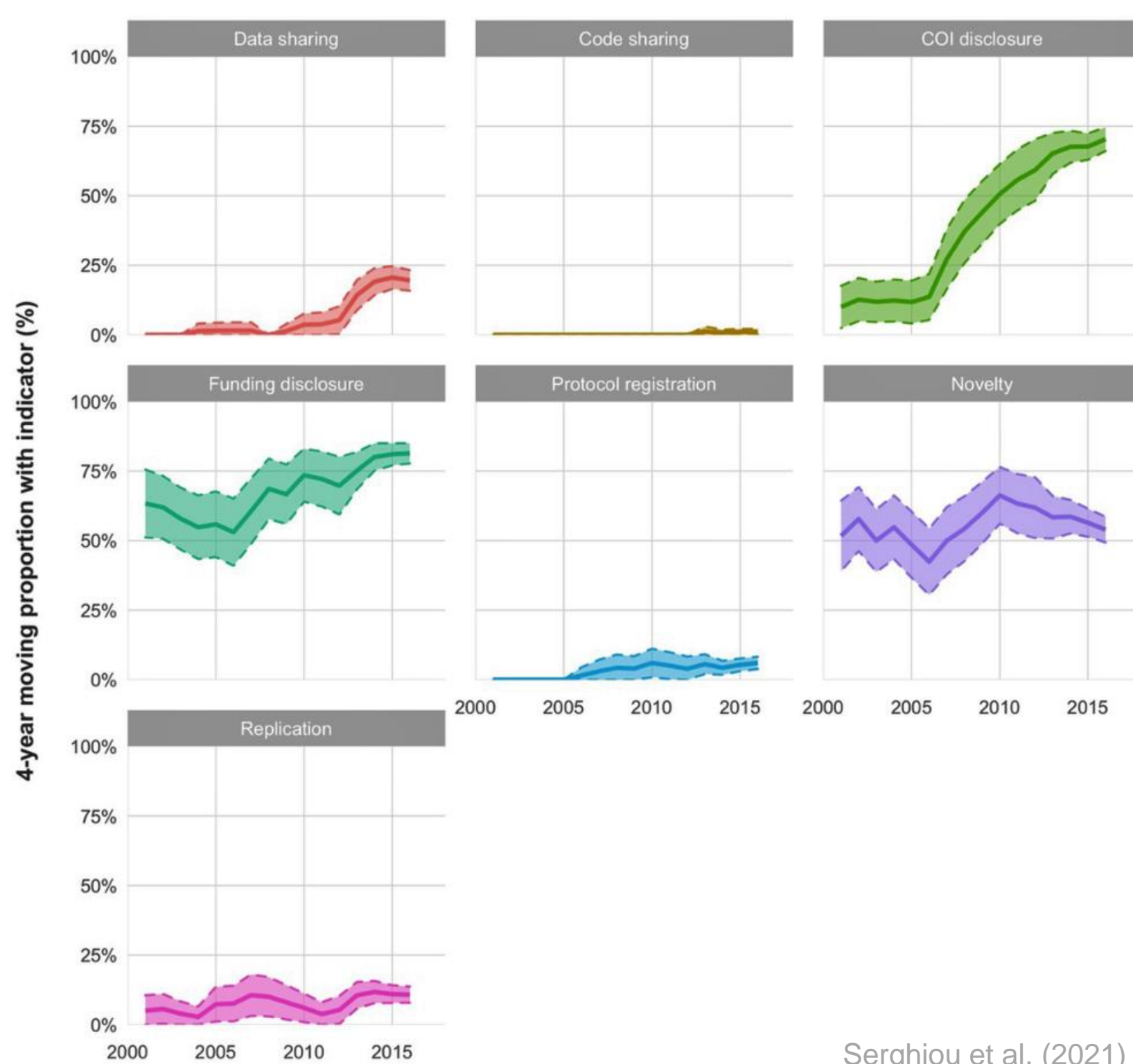
Thomas Klebel, Christian Fleck, Tony Ross-Hellauer

MOTIVATION

The Open Science (OS) movement aims at creating a more equitable and more open scientific process. Part of this puzzle are distinct practices, such as publishing research articles Open Access (OA), sharing research data, or code. Given that these practices are increasingly mandated by institutions, funders and institutions, we want to investigate, which types of institutions and individual researchers actually take up these practices. Is the uptake of OS spearheaded by established and well-funded actors, or early career researchers?

STATE OF THE ART

Practices of OS are gaining momentum, albeit at different rates. OA publishing is becoming mainstream, with about 45% of publications being OA in 2015, and an increase since (Piwowar et al. 2018). Data sharing and code sharing are rising too, but starting from much lower levels (Serghiou et al. 2021).



Serghiou et al. (2021)

METHODS/NEW APPROACH

The thesis will employ multiple computational approaches:

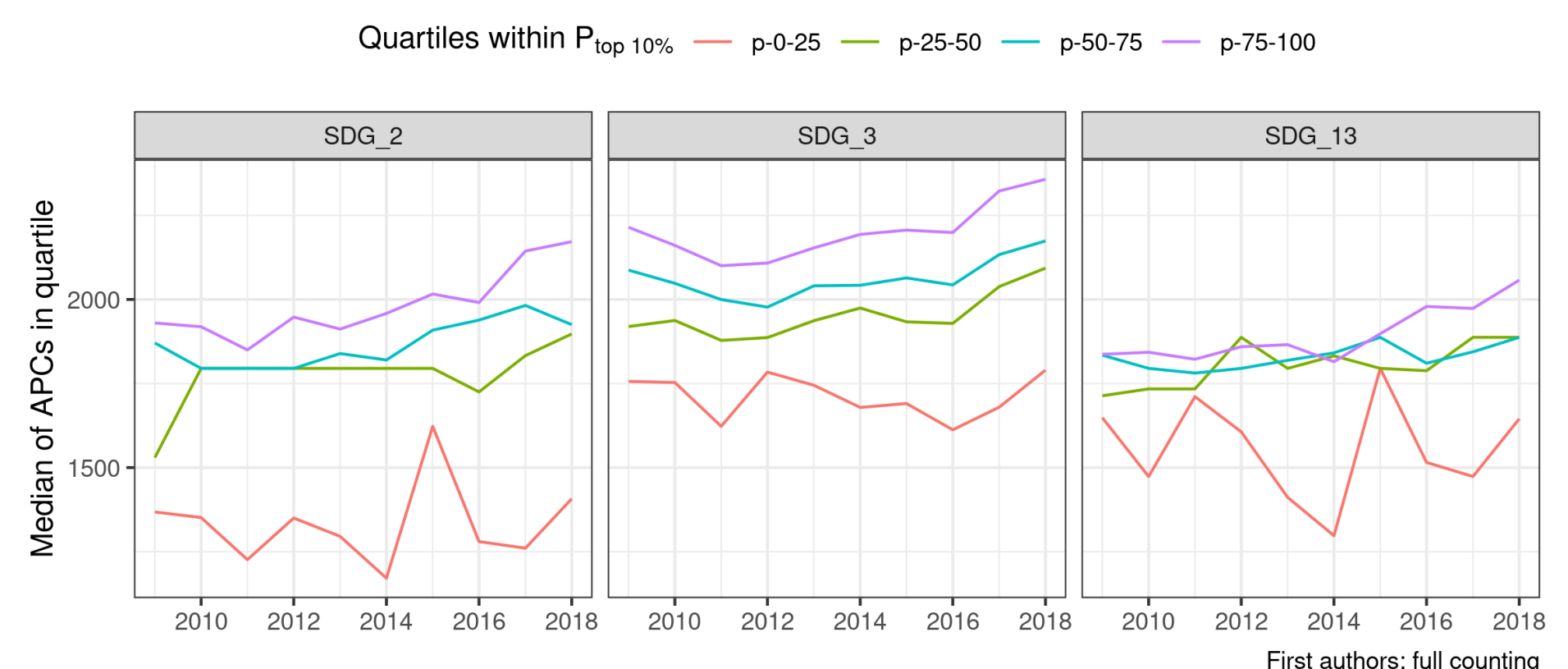
- Study 1 uses large-scale bibliometric data from Microsoft Academic Graph (MAG), joined with data from Unpaywall, genderize.io and the CWTS Leiden Ranking to investigate the uptake of OA publishing along common dimensions of stratification (institutional standing, gender, academic age, etc.).
- Study 2 leverages agent-based modeling to analyse potential mechanisms influencing the uptake of data sharing on the levels of institutions and academic journals.
- Study 3 uses full-text data from all publications hosted on JSTOR for the disciplines “economics” and “sociology” to investigate the uptake of code sharing. Recent approaches for detecting code sharing statements will be combined with bibliometric-based approaches from study 1 to provide novel insights into who is sharing their research code in the two social science disciplines.

RESULTS

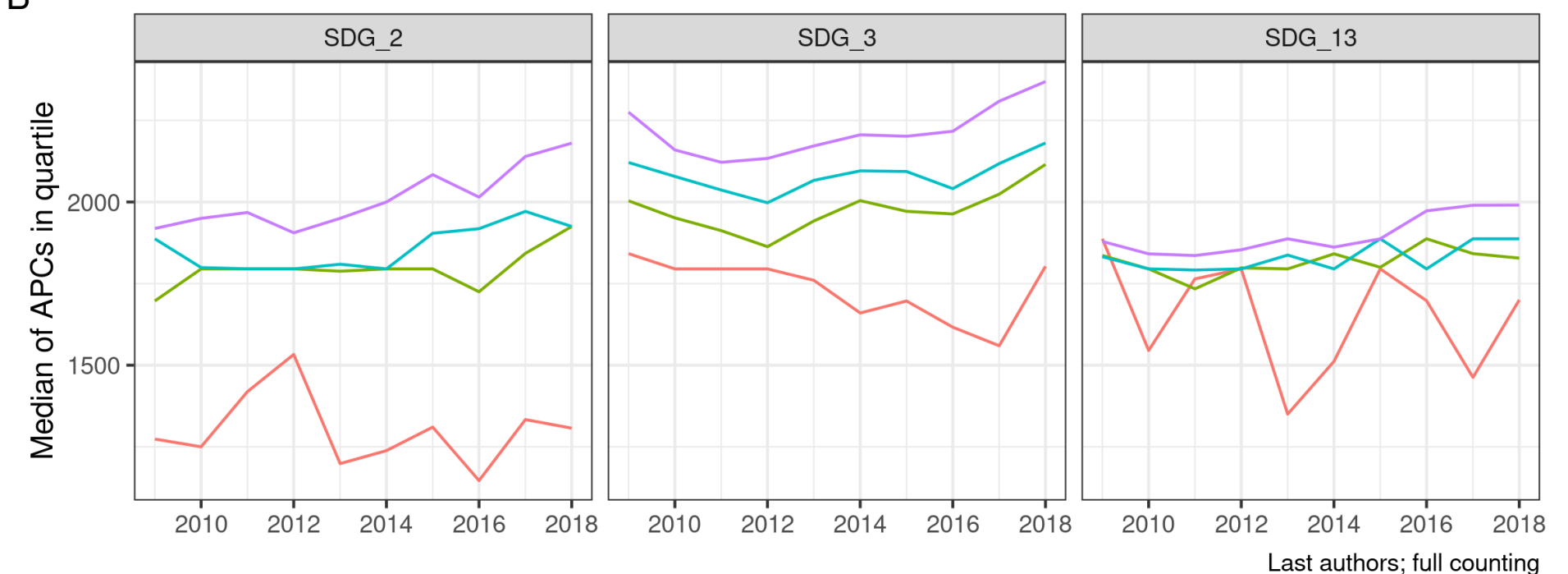
Study 1

We conducted our analysis based on a set of publications which are relevant to three key UN SDGs: Zero Hunger (SDG 2), Health/Well-Being (SDG 3) and Climate Action (SDG 13). Based on these, we find an increase in the share of publications that are OA (in line with the general literature), but also a marked increase in average article processing charges (APCs). Higher prestige institutions publish in journals with higher APCs, and this trend is intensifying. The APC-publishing model thus might actually further existing inequalities.

A



B



Plan for Study 2

We will create a model of academia, capturing the effect of cumulative advantage by rewarding scientists for published papers. Introducing mandates for data sharing on institutional and journal levels will allow us to investigate the interplay between individual success, data sharing policies and overall trends in terms of inequality. Potentially, the model would allow for journals and institutions to change policy, thus introducing a further dynamic element.

Plan for Study 3

The aim of this study is to provide an overview of who is currently sharing code among two social science disciplines (sociology and economics) with differing levels of quantitative research. By combining existing text mining approaches (R-package *rtransparent*) with a custom approach to distinguish quantitative from theoretical or qualitative research, this study will provide new insights into the structure of early adopters in these fields.

REFERENCES

- Piwowar et al., 2018, *PeerJ* 6: e4375.
<https://doi.org/10.7717/peerj.4375>.
 Serghiou et al., 2021 *PLOS Biology* 19 (3): e3001107.
<https://doi.org/10.1371/journal.pbio.3001107>.