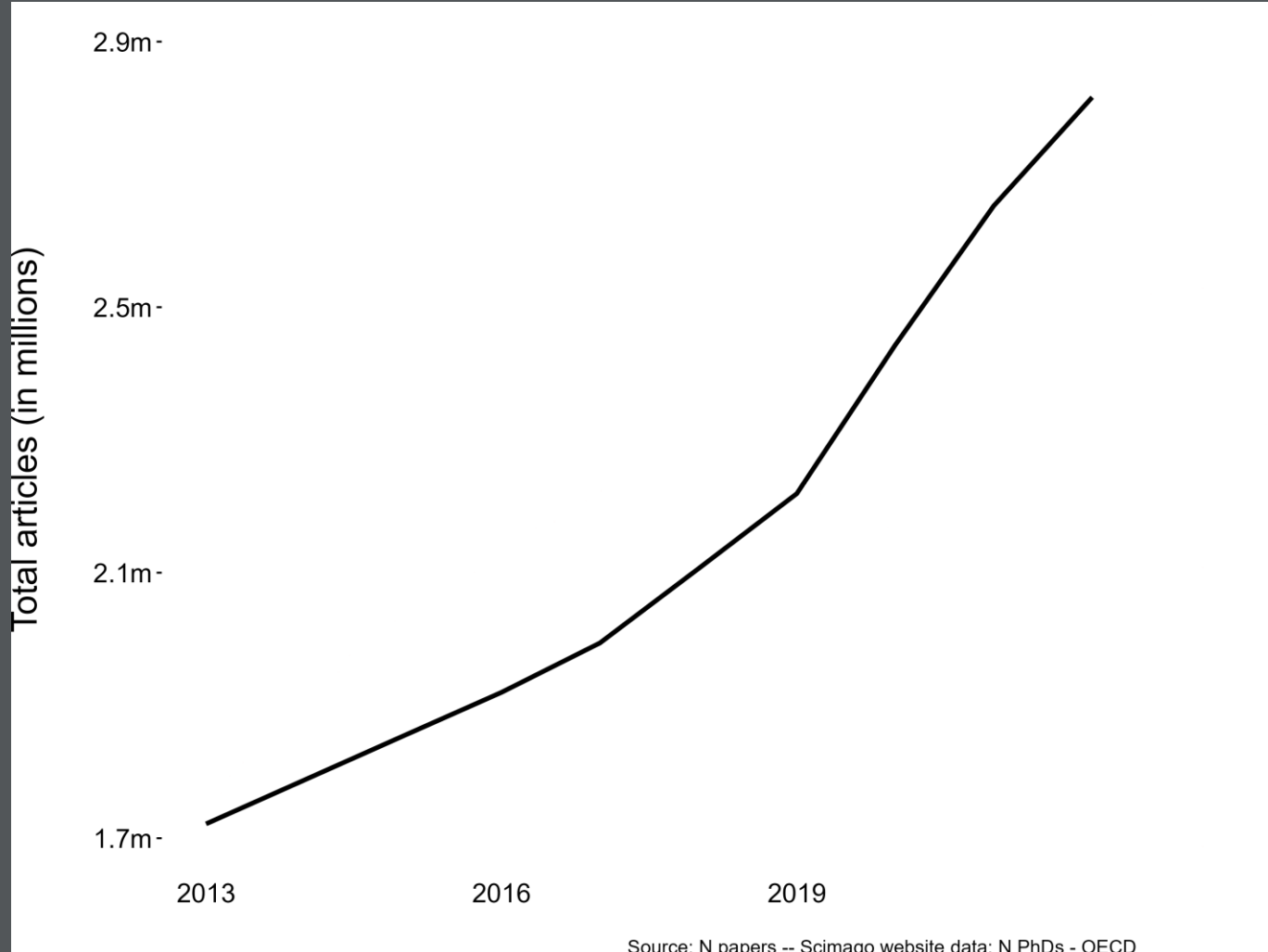


The **strain** on scientific publishing

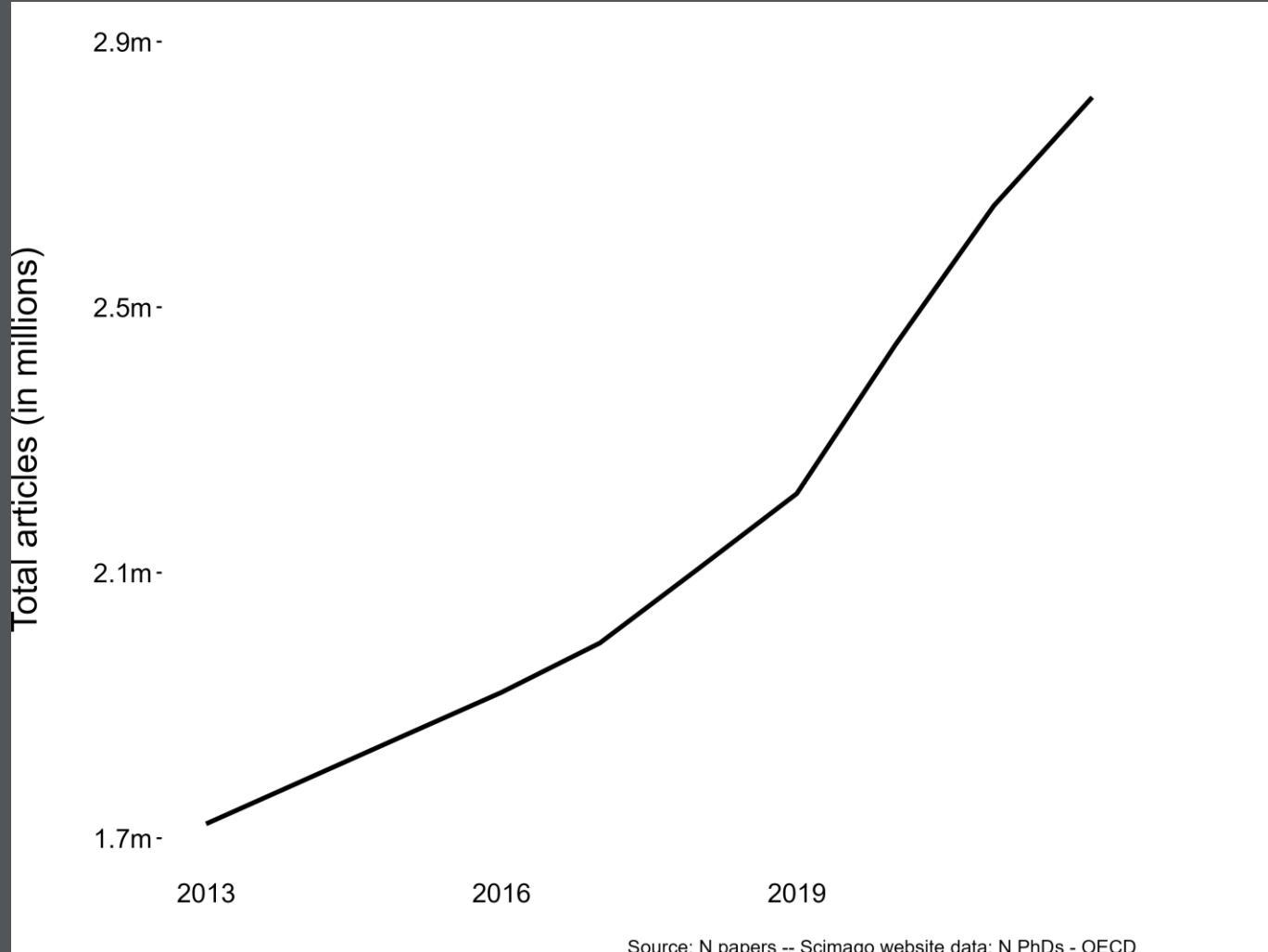
Mark Austin Hanson¹, Pablo Gomez-Barreiro², Paolo Crosetto³, Dan Brockington^{4,5}

The Growth in Publishing



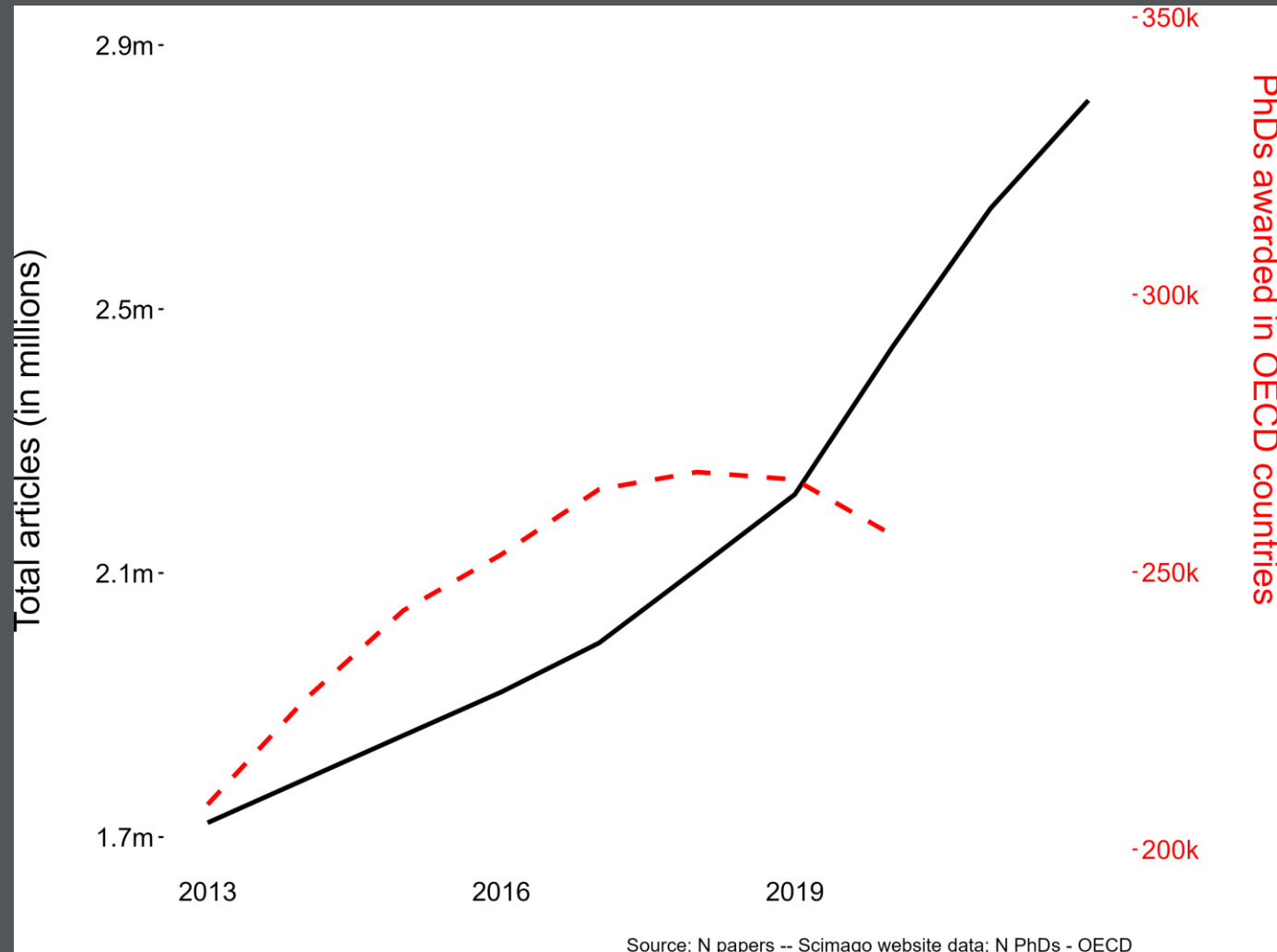
Growth in papers indexed by WoS and Scopus

The Growth in Publishing – welcome if . . .



- Circumvents gatekeepers and combats epistemic inertia
- Tackles positive result bias and file drawer problem
- Reflects broader global research investment

The Growth in Publishing – but causes strain



PhDs awarded in OECD countries

- Editors resigning
- Reviewers overworked
- Paper mills, fraud, retractions, multi-million dollar losses and 'sunsetting' of Hindawi brand

A semantic shift

"Publication"

used to mean

- A physical object with limited space
- a handful of journals
- long delays
- free for authors
- do it and thrive

→ good science got rejected?

A semantic shift

"Publication"

used to mean

- A physical object with limited space
- a handful of journals
- long delays
- free for authors
- do it and thrive

→ good science got rejected?

now it also means

- limitless named electronic repositories
- thousands of journals
- short delays
- authors pay
- don't do it and die

→ bad science accepted?

"Special issue"

used to mean

- A once-in-a-while issue
- About a special topic
- Strict editor control
- regular > special

"Special issue"

used to mean

- A once-in-a-while issue
- About a special topic
- Strict editor control
- regular > special

now it also means

- A many-a-day issue
- About any topic
- Relaxed editor control
- special > regular

"Publisher business model"

used to mean

- Many small journals
- Readers pay
- \$ through subscription
- "*Polish your gems*"

Incentive to ↑↑ **quality**,
quantity? ...

A semantic shift

"Publisher business model"

used to mean

- Many small journals
- Readers pay
- \$ through subscription
- "*Polish your gems*"

Incentive to ↑↑ **quality**,
quantity? ...

now it **also** means

- Few mega-journals
- Authors pay
- \$ through publication
- "*Get authors on board*"

Incentive to ↑↑ **quantity**,
quality? ...

Our aim:
understanding the strain on publishing

Our methods

- Scimago



+



Our methods

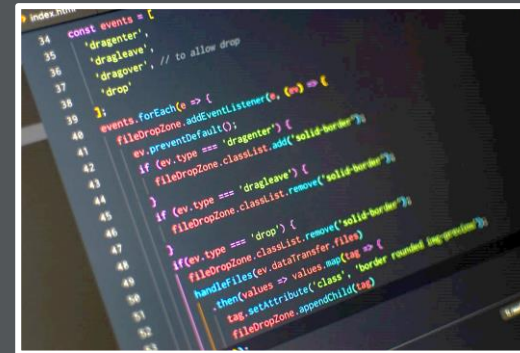
- Scimago



+



- Web-scraping



Our methods

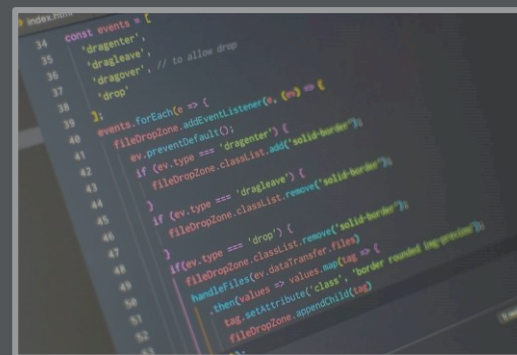
- Scimago



+



- Web-scraping



- Requests

Publishers



Hindawi



Springer



Taylor & Francis

nature
portfolio



frontiers

ELSEVIER

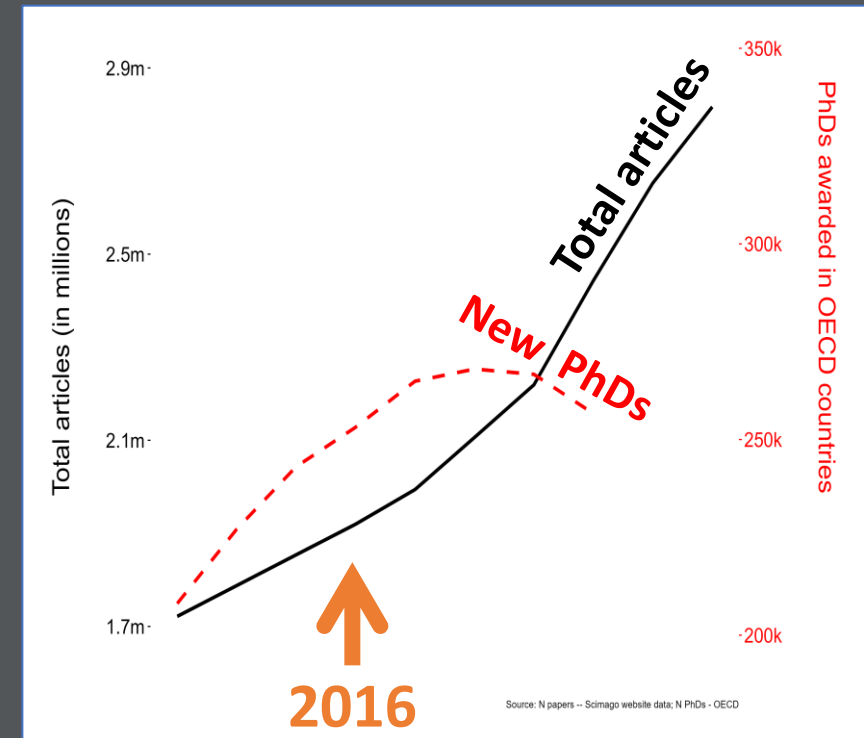
The *strain* on scientific publishing

Disproportionate growth in articles → “strain”

Annual articles (Web of Science*Scopus)

2016: ~ 1.9 million

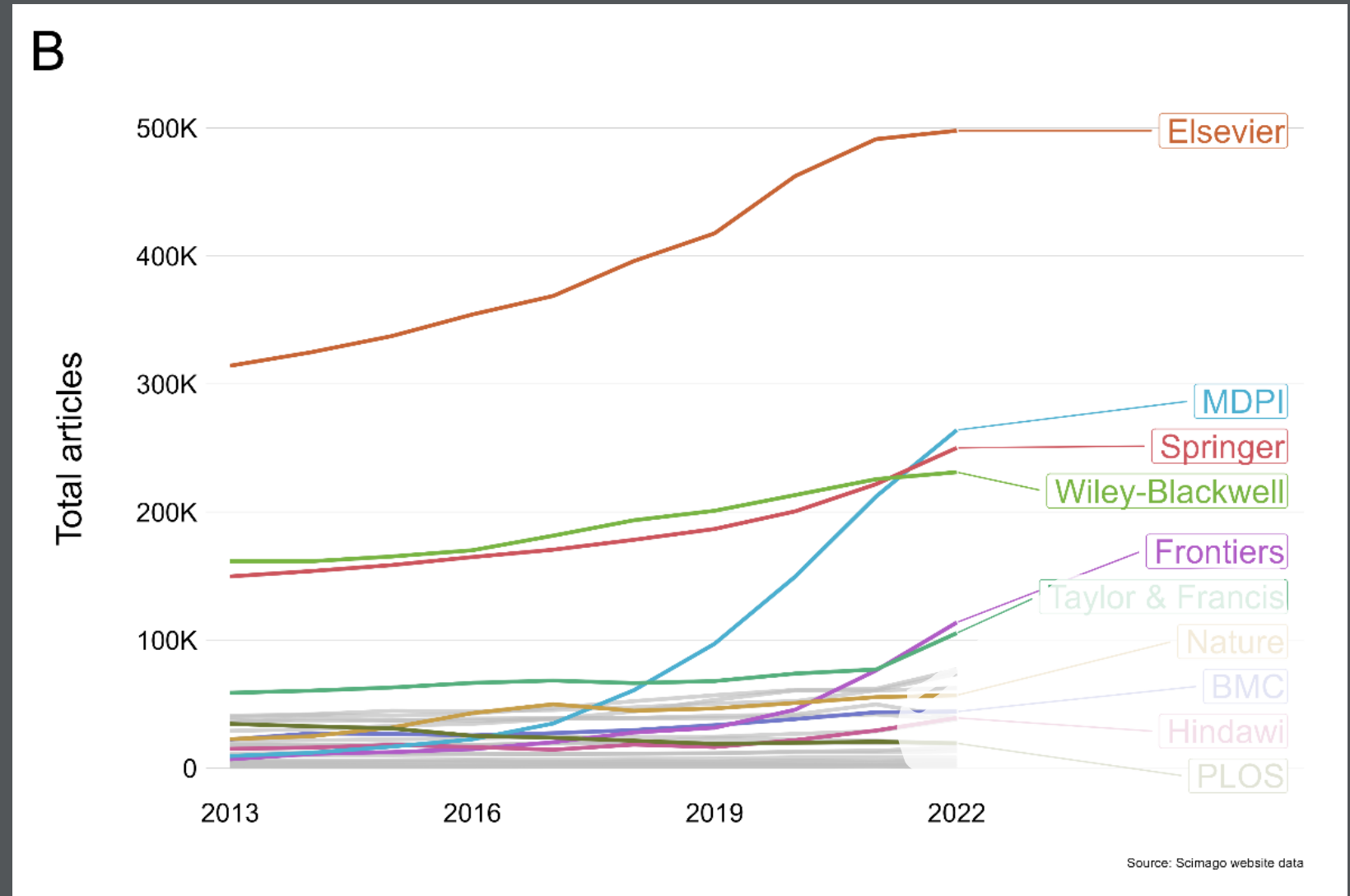
2022: ~ 2.8 million



Disproportionate growth in articles → “strain”

Five publishers:

- MDPI (27%)
- Elsevier (16%)
- Frontiers (10%)
- Springer (10%)
- Wiley (7%)

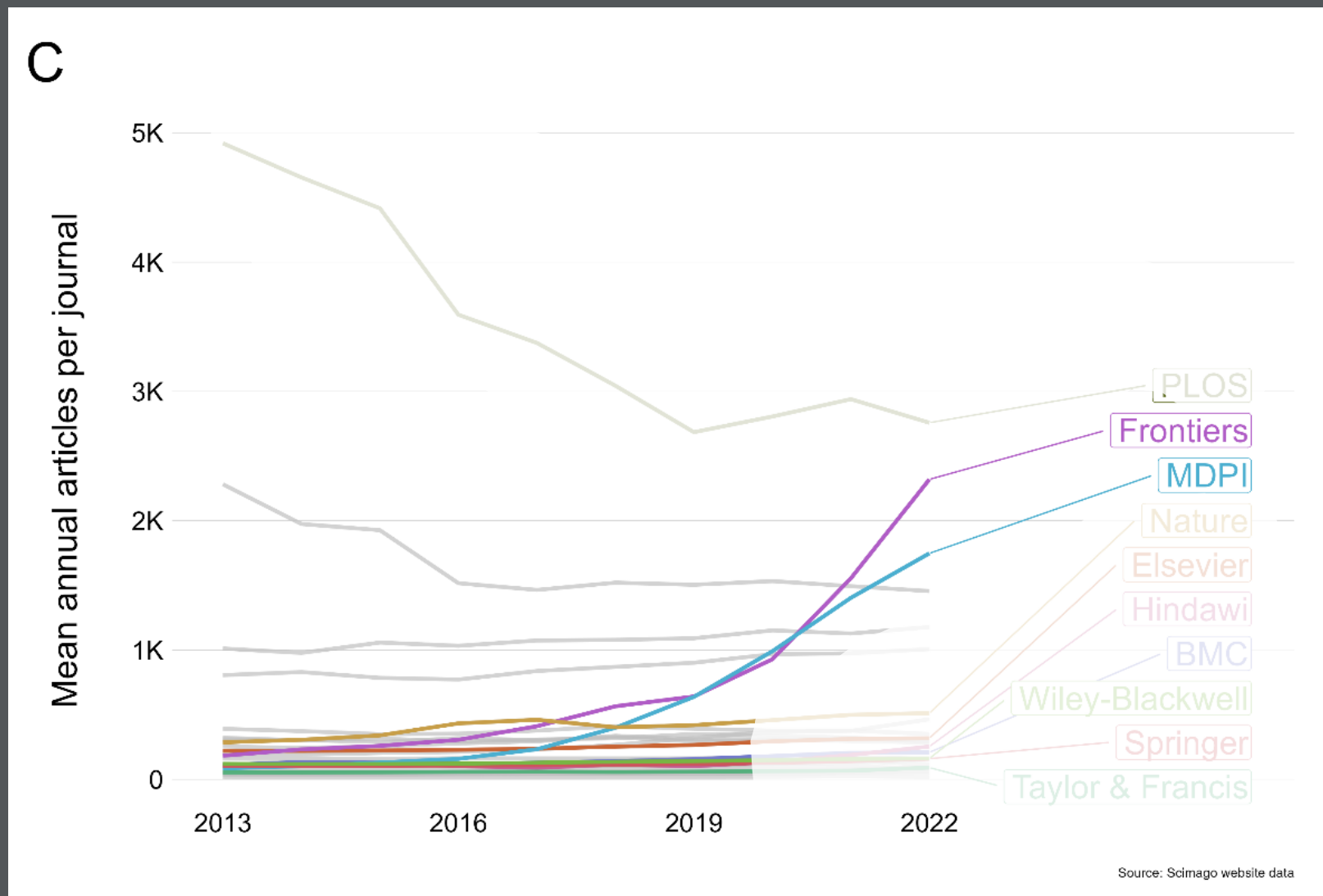


Growth in journal size

New!!!

“Mega journals”

- MDPI
- Frontiers



Strain: what's going on?

Trends

- Two strategies:
 - Growth in articles overall
 - Growth in *articles per journal*

Threats

- “publish or *perish*”
- Information overload...

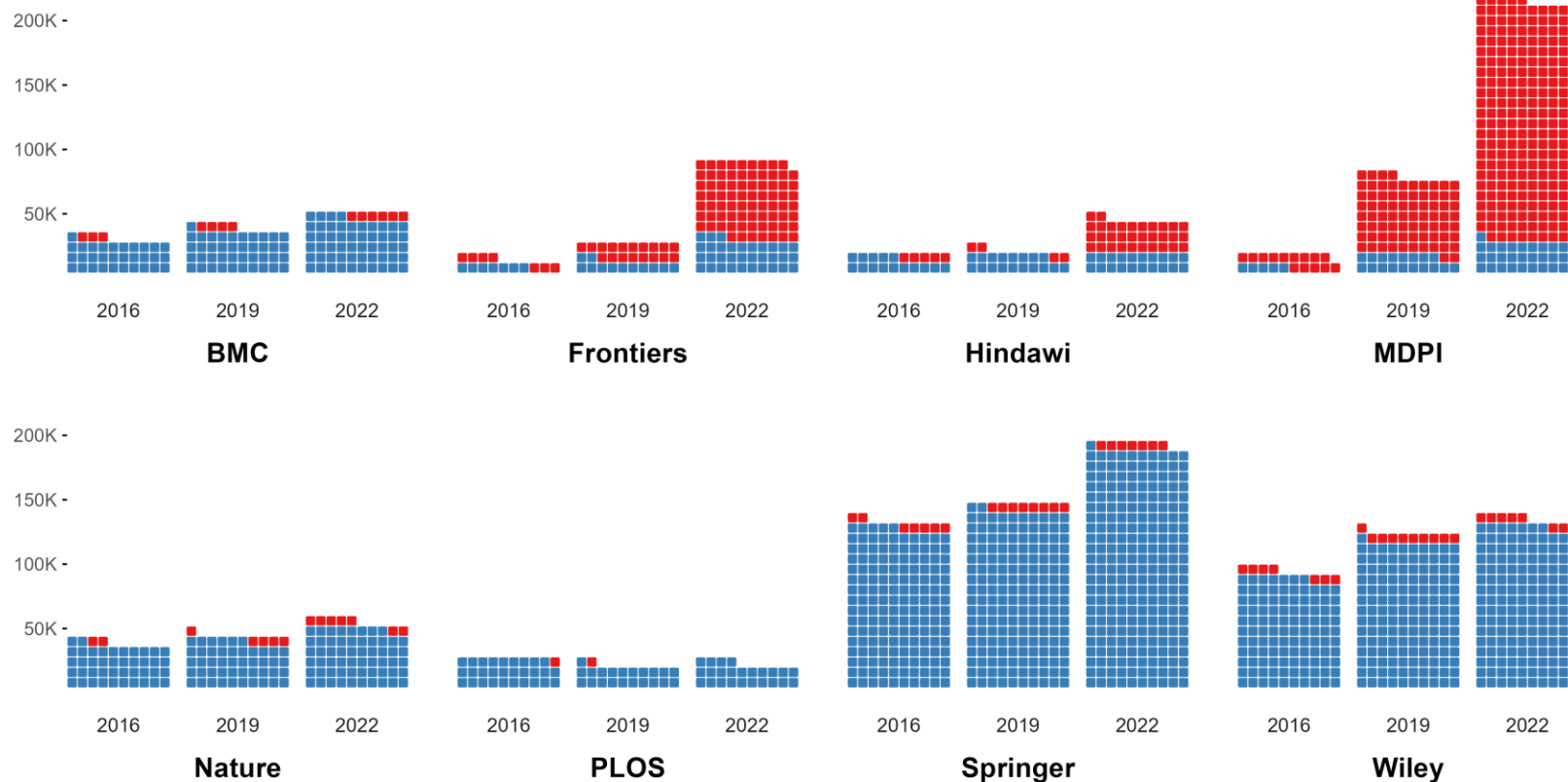
Not so special issues...

Three players:

MDPI
Frontiers
Hindawi

Number of papers published in **regular** vs **special** issues, 2016-22

One square = 800 articles



Source: data scraped from the publisher's website
Note: Special issues are called Collections at PLOS and Topics at Frontiers. For MDPI Collections, Sections and Topics not shown.

Not so special issues...

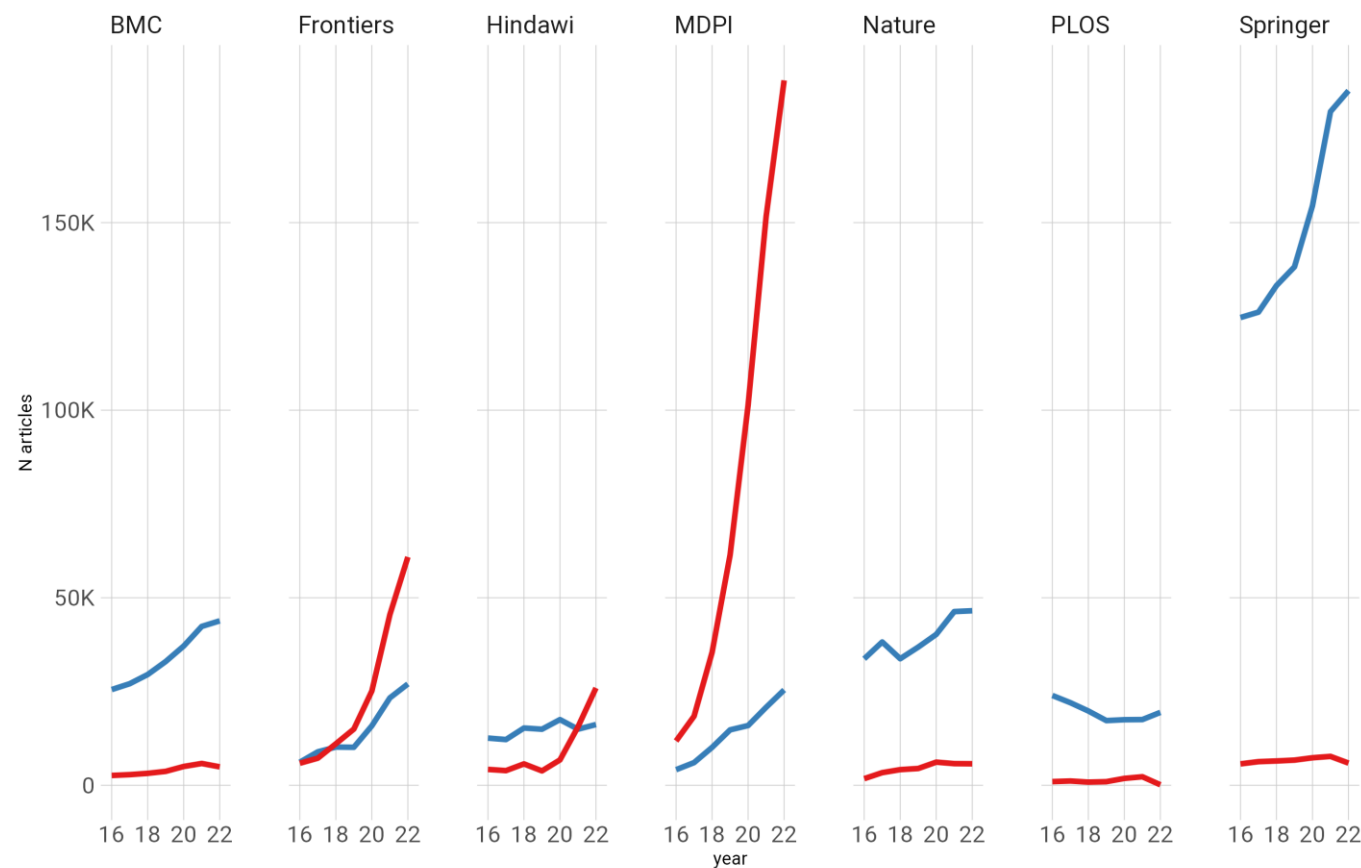
Three players:

MDPI

Frontiers

Hindawi

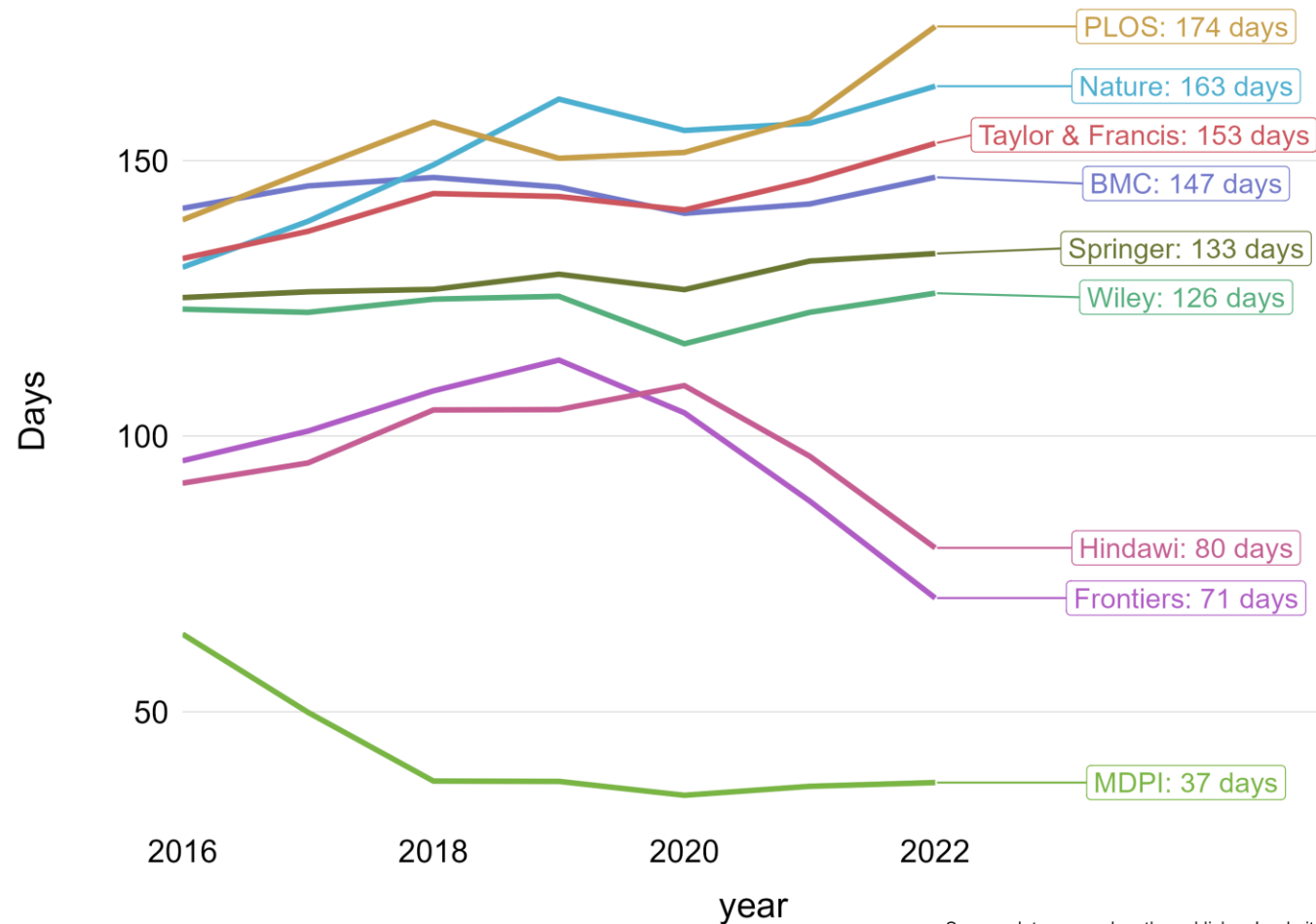
Number of papers published in **regular** vs **special** issues, 2016-22



Source: data scraped from the publisher's website
Notes: Special issues are called Collections at PLOS and Topics at Frontiers. For MDPI Collections, Sections and Topics not shown.

Turnaround Times

Submit → Accept
including revisions...

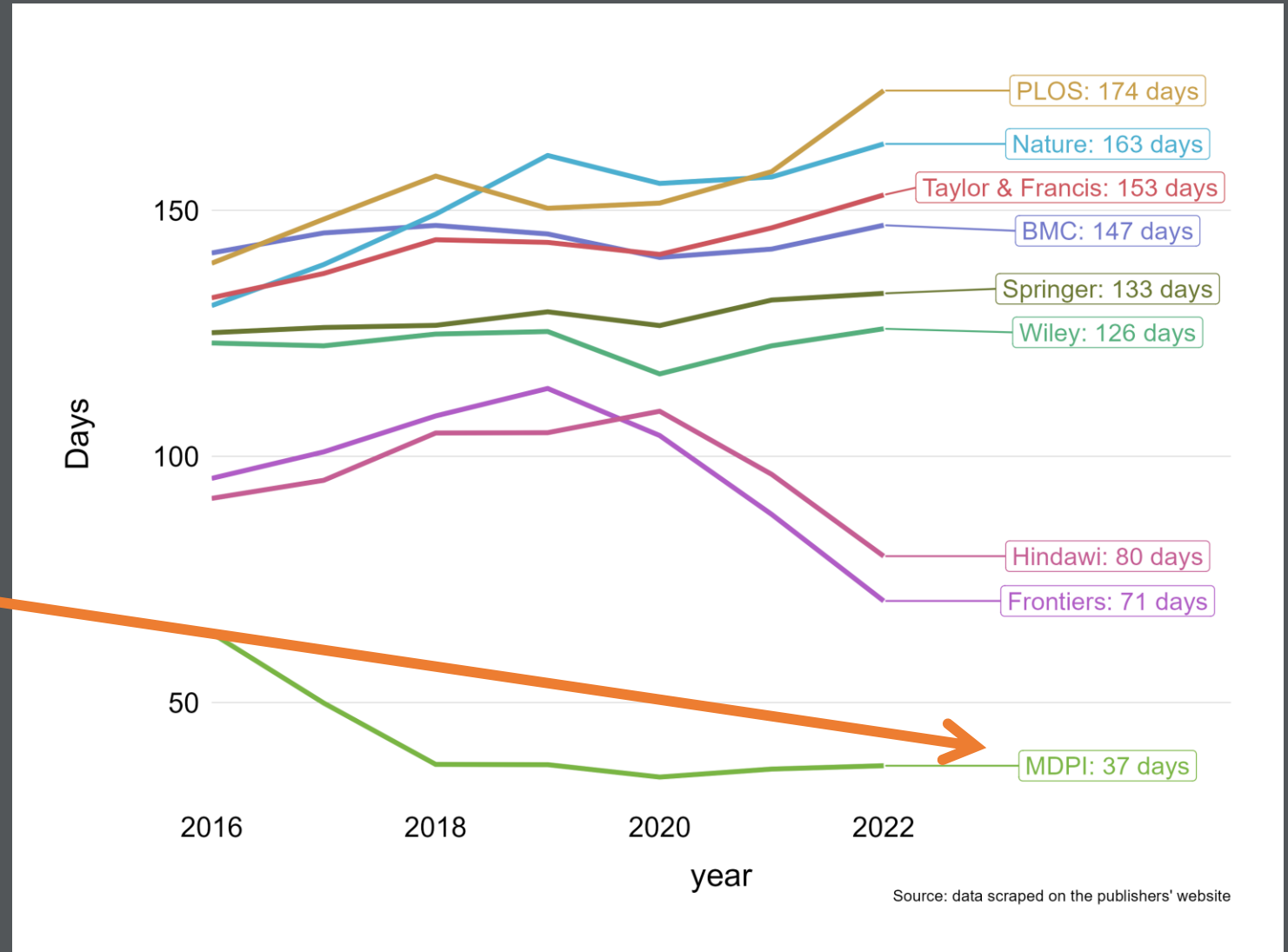


Source: data scraped on the publishers' website

Turnaround Times

Submit → Accept
including revisions...

Is it really possible
to review, revise
and re-review so
fast?

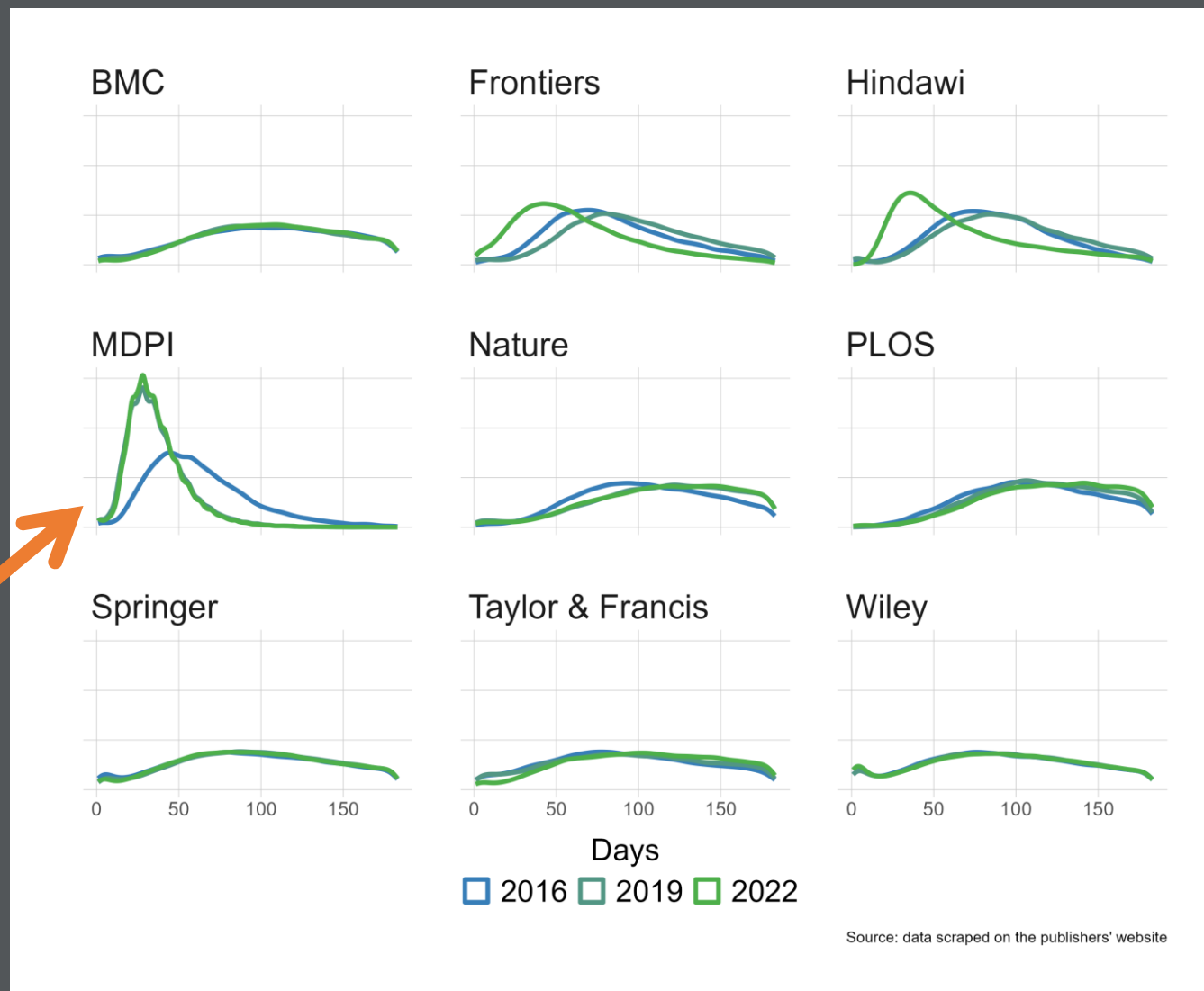


Turnaround Times

Submit → Accept
including revisions...

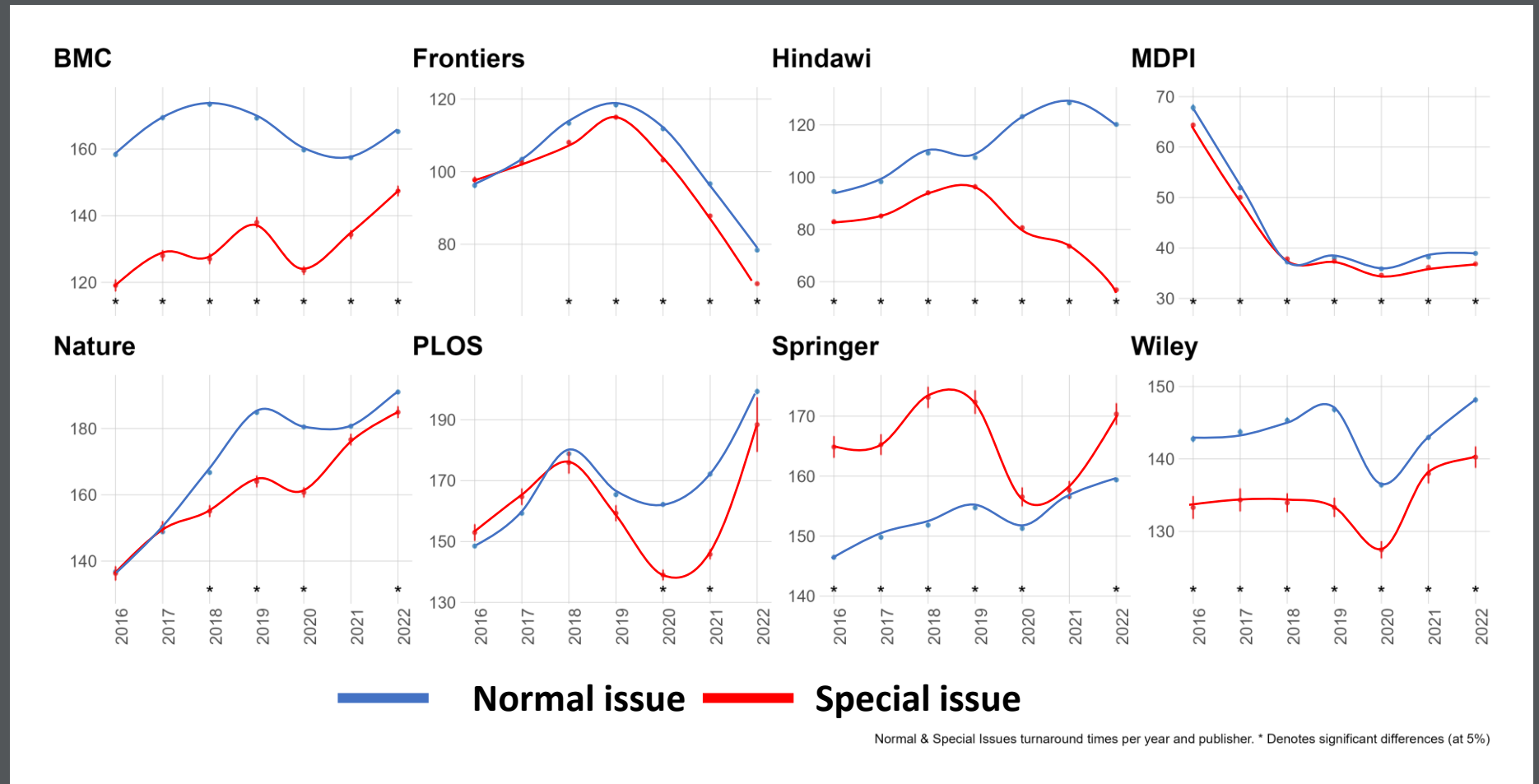
Is it really possible
to review so fast?

And so
consistently?



Turnaround Times in Special Issues

Special
issue
articles
published
faster



Special issues: what's going on?

Trends

- “Special” issues are a fantastic engine for growth

Why?

- Guest editors + reviewer networks → **scalable**

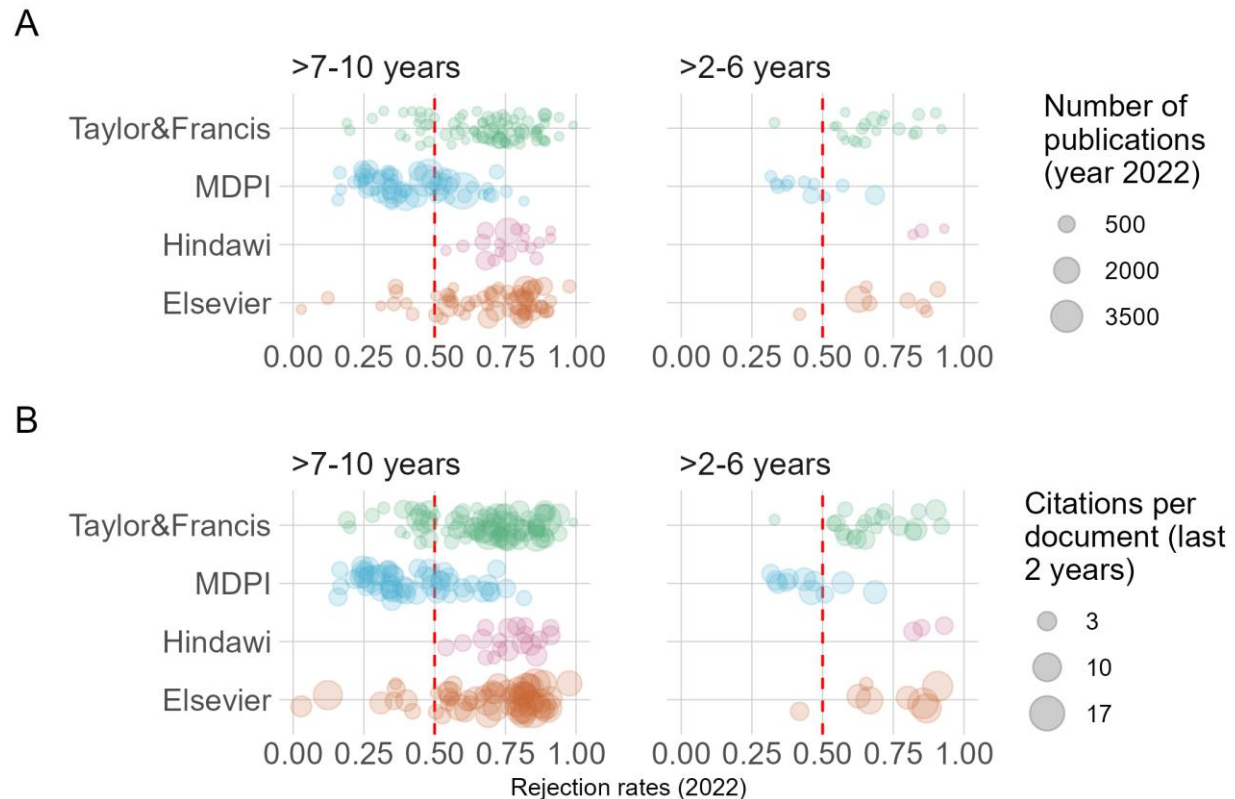
Threats

- Can one do **quality** peer review so fast?

Rejection rates: rarely public

Many caveats...

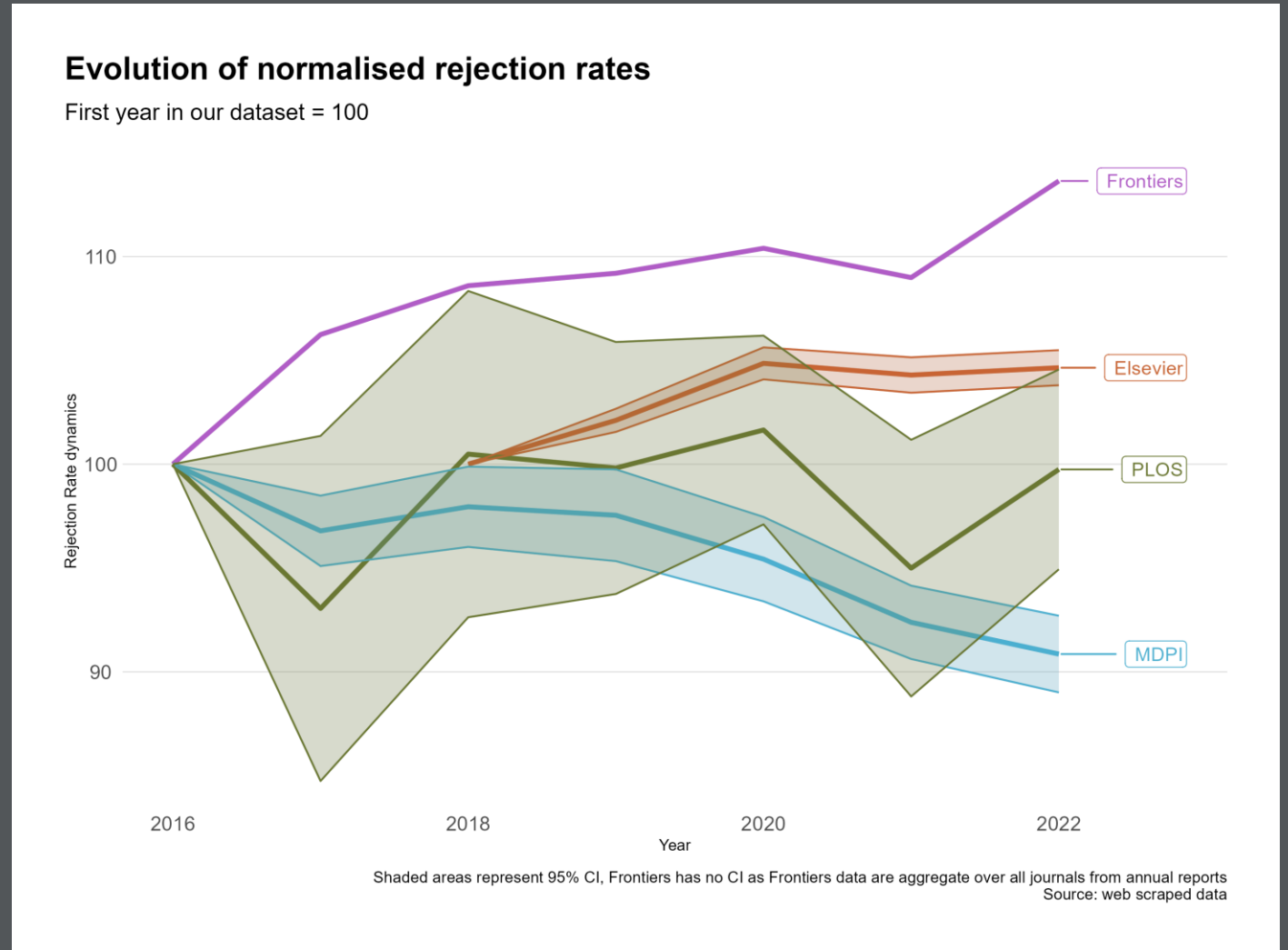
- No universal definition
- Reject + resubmit = ...?



Rejection rates vs total documents or citations per document by journal age and publisher.
Rejection rates obtained from diverse sources

Rejection rates

Trends best explained by publisher



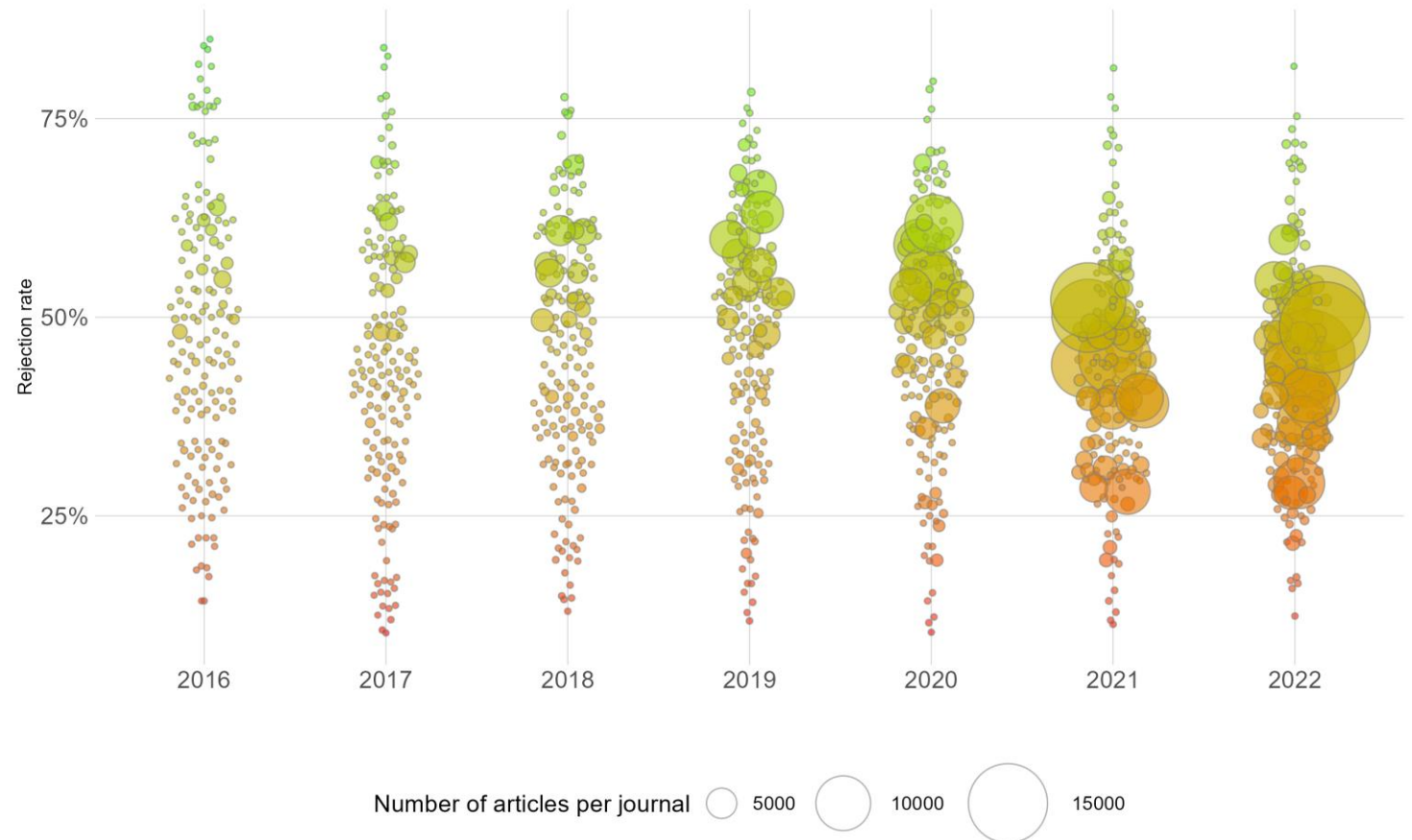
Rejection rates

MDPI trends

Recently reversed?

Evolution of rejection rates by relative size of the journal at MDPI, 2016-22

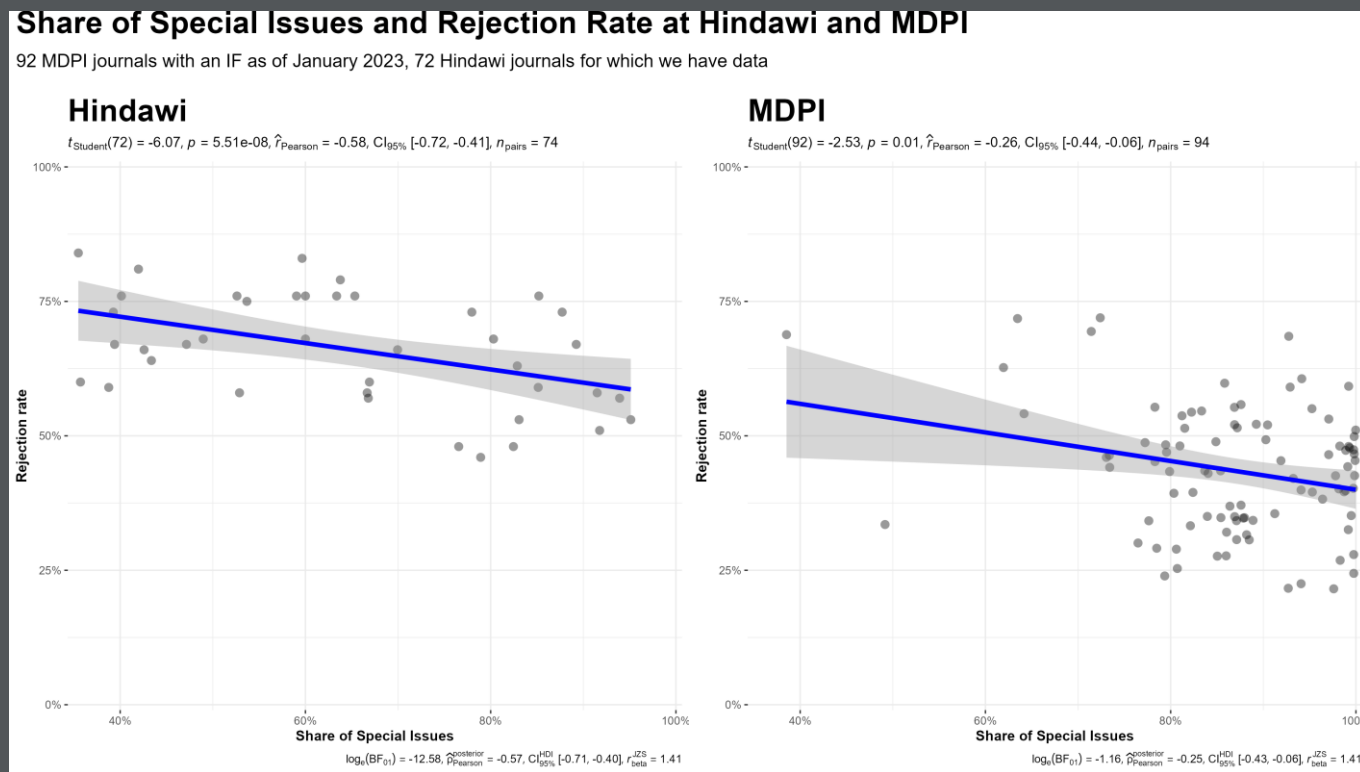
Only journals existing in 2016



Source: data scraped on the publisher's website

Rejection rates and Special Issues

Special issues,
special treatment



↑ Special issues = ↓ Rejection rates

Rejection rates: what's going on?

Trends

- Not transparent

Threats

- Publisher dictate **scope**

Inflation:

Inflation: *Impact Factor*

also housing... food... life...

Inflation: *Impact Factor*

Impact Factor (IF)

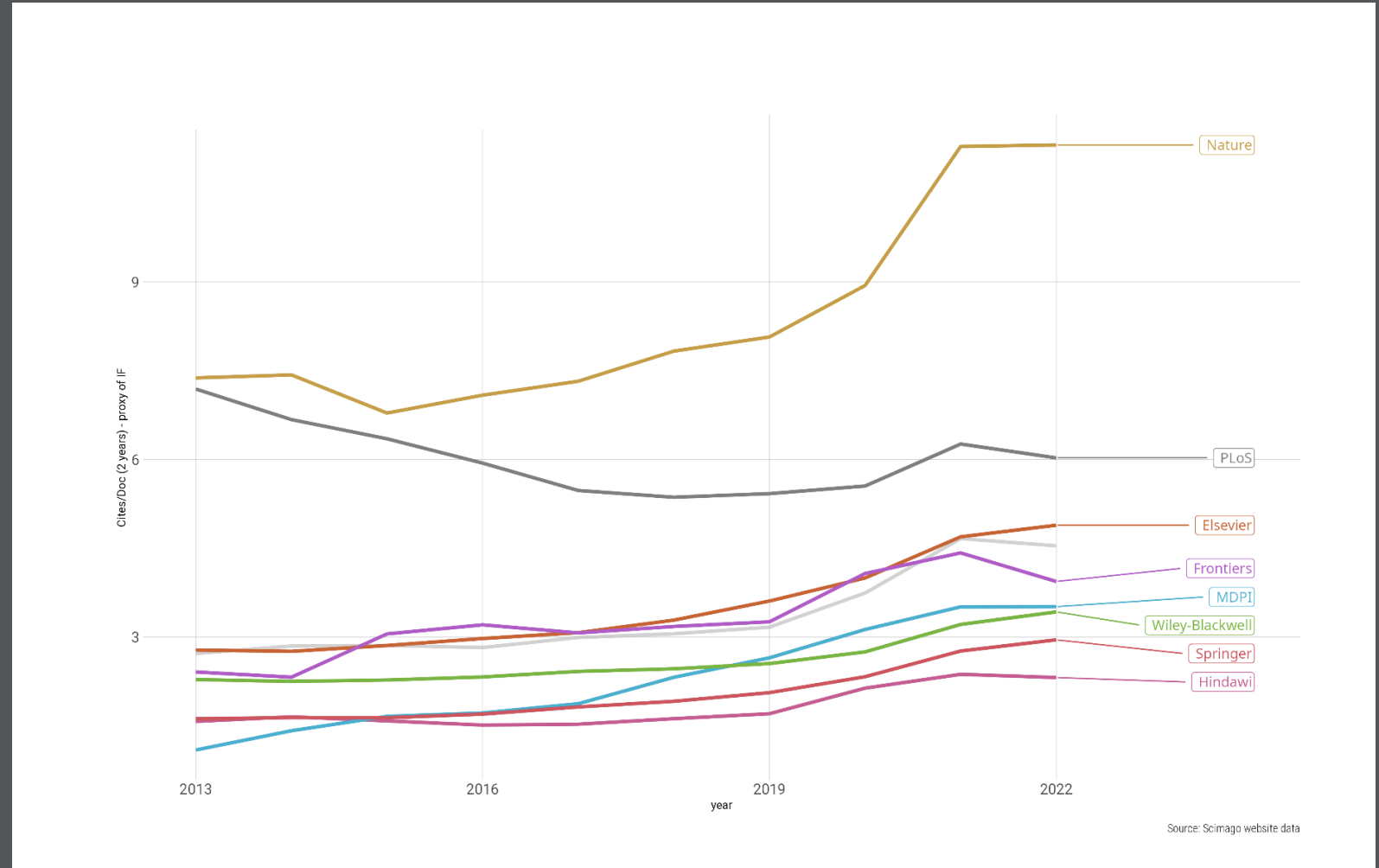
- Total citations (2 years)

N citations

Impact inflation

Impact factor

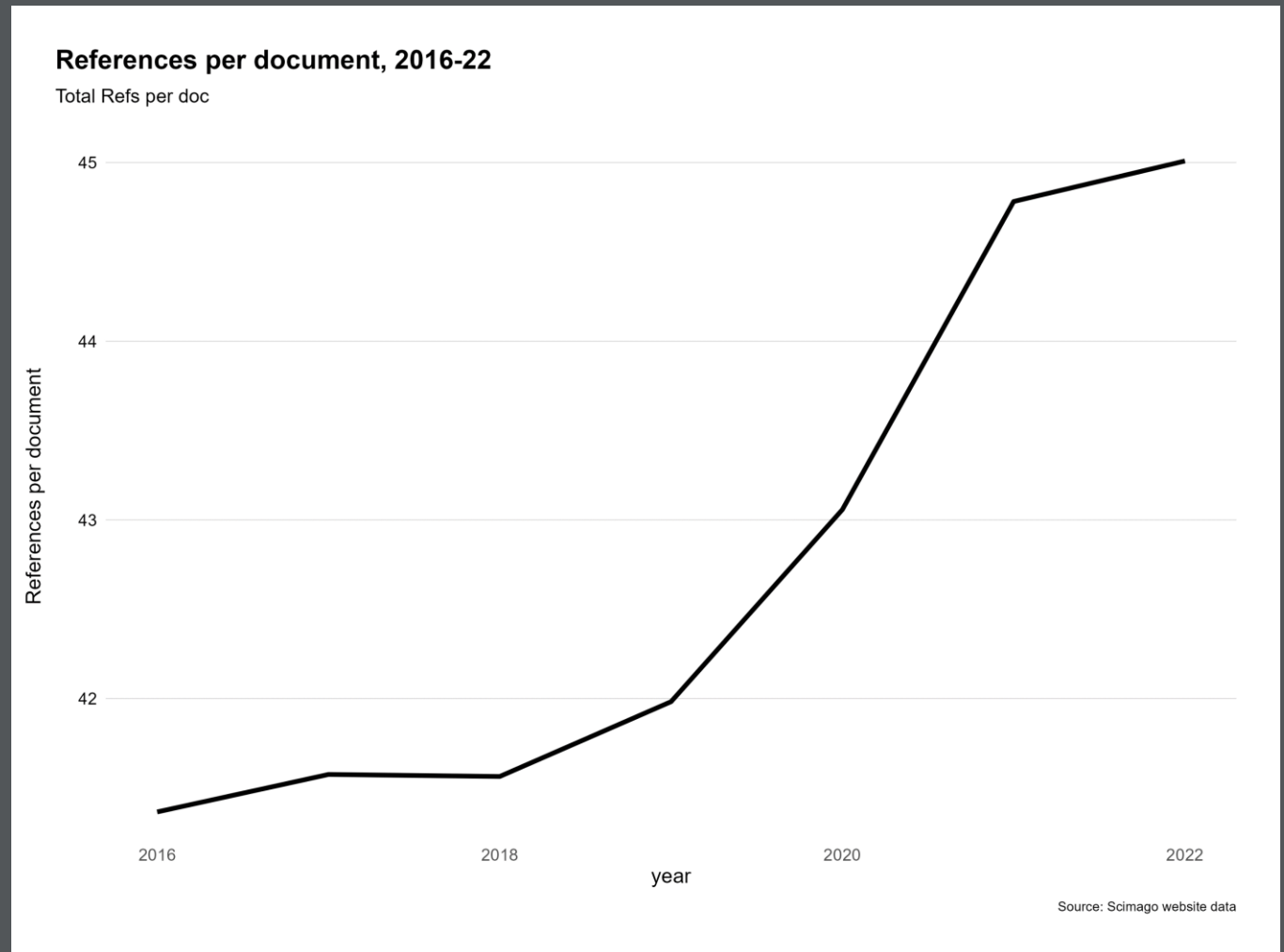
- Increased universally



Impact inflation

Impact factor

- Increased universally
- Refs per doc part of it...
- COVID (not whole story)



Impact inflation

Impact inflation

Goodhart's law

When a measure becomes a target, it ceases to be a good measure...

Impact inflation

Impact Factor (IF)

- Total citations (2 years)

N citations

Impact inflation

Impact Factor (IF)

- N citations

Scimago Journal Rank (SJR)

- Network... self-cites...

$$\frac{IF}{SJR}$$



“Impact Inflation”

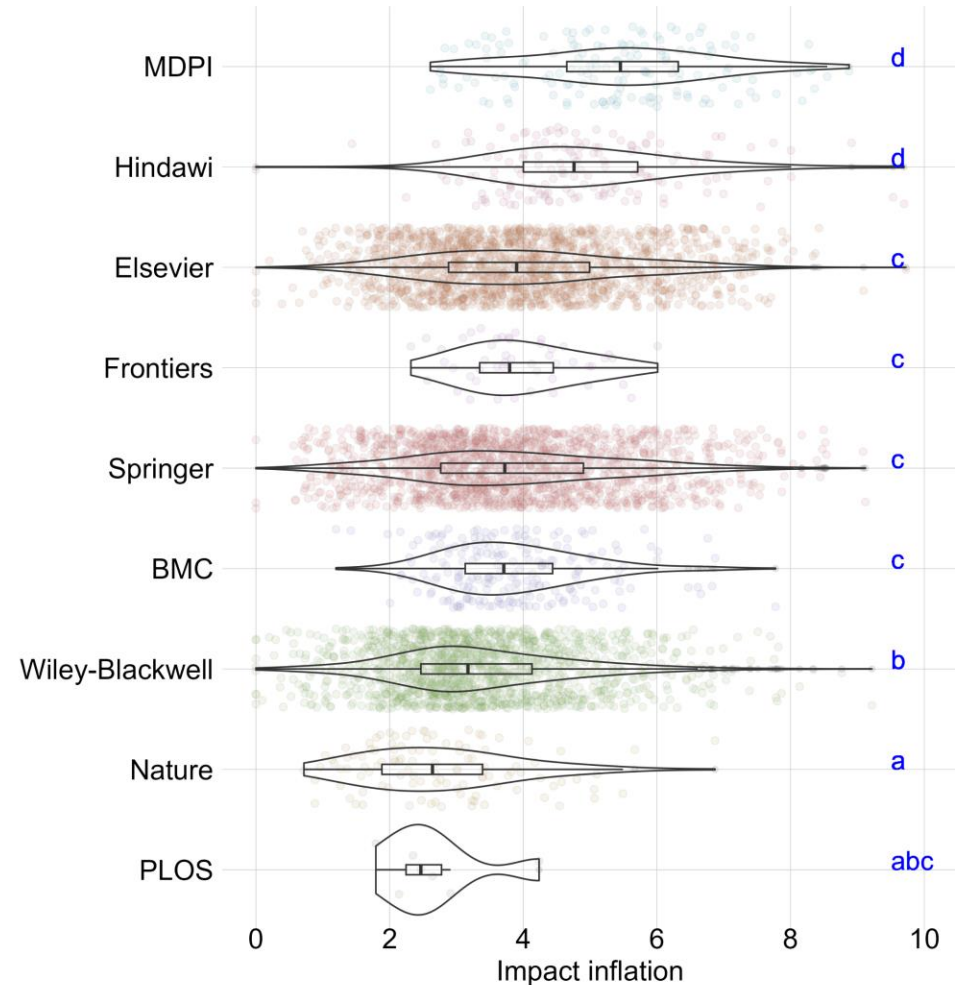
Goodhart’s law

Impact inflation

Impact inflation

- MDPI
- Hindawi

but not *Frontiers...*

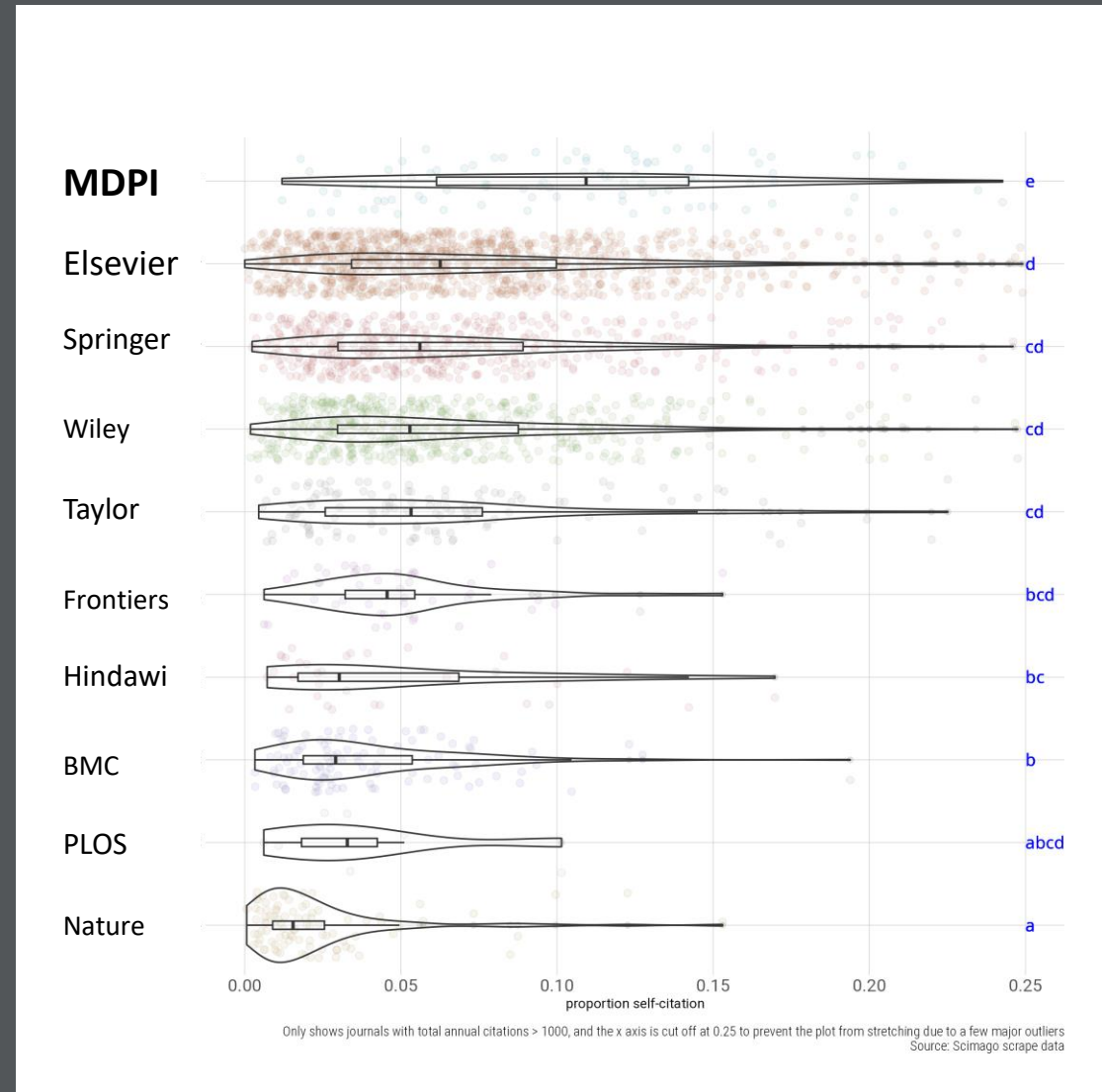


The x-axis is limited at 10 to prevent the plot from stretching to show just a few major outliers
Source: Scimago website data

Self-citations: MDPI

Self-cite rate

- MDPI



What's going on?

Trends

- Everyone's IF is going up
- One group (MDPI) self-cites a lot

Why?

- Lots of citations from a small pool of journals

Threats

- **Quality** signals distorted, threat to naïve scientists

What is going on?

Strain indicators at a glance: 2022 and evolution 2016-22

	2022					CHANGE 2016-22		
	TOTAL PAPERS	SHARE SPECIAL ISSUE	TURNAROUND TIME (DAYS)	REJECTION RATE	IMPACT INFLATION	NUMBER PAPERS	SHARE SPECIAL ISSUE	TURNAROUND TIME (DAYS)
Overall	2816k	38%	136	60%	3.3	+47%	+27pp	+2
Elsevier	498k	--	--	71%	4.0	+41%	--	--
MDPI	264k	88%	37	40%	5.4	+1080%	+14pp	-33
Springer	250k	3%	136	--	3.9	+52%	-1pp	+2
Wiley-Blackwell	231k	5%	149	--	3.3	+36%	-2pp	+4
Frontiers	114k	69%	71	48%	4.0	+675%	+20pp	-20
Taylor & Francis	105k	--	167	--	NA	+59%	--	+15
Nature	57k	11%	171	--	2.8	+32%	+6pp	+42
BMC	44k	10%	142	--	3.9	+73%	+1pp	+10
Hindawi	39k	62%	111	--	5.0	+139%	+36pp	+16
PLOS	19k	1%	170	59%	NA	-23%	-3pp	+21

pp = 'percentage points'. Source: data scraped on the publishers' website or publishers' own publications. Overall: SI, TAT, Rejection % based on publishers with available data only. N per Scimago dataset. Elsevier: rejection rate change starts from 2018. All publishers: Special Issues sometimes named differently.

What is going on: admittedly... **one group** stands out

Strain indicators at a glance: 2022 and evolution 2016-22

	2022					CHANGE 2016-22		
	TOTAL PAPERS	SHARE SPECIAL ISSUE	TURNAROUND TIME (DAYS)	REJECTION RATE	IMPACT INFLATION	NUMBER PAPERS	SHARE SPECIAL ISSUE	TURNAROUND TIME (DAYS)
Overall	2816k	38%	136	60%	3.3	+47%	+27pp	+2
Elsevier	498k	--	--	71%	4.0	+41%	--	--
1st: MDPI	264k	88%	37	40%	5.4	1080%	+14pp	-33
Springer	250k	3%	136	--	3.9	+52%	-1pp	+2
Wiley-Blackwell	231k	5%	149	--	3.3	+36%	-2pp	+4
Frontiers	114k	69%	71	48%	4.0	+675%	+20pp	-20
Taylor & Francis	105k	--	167	--	NA	+59%	--	+15
Nature	57k	11%	171	--	2.8	+32%	+6pp	+42
BMC	44k	10%	142	--	3.9	+73%	+1pp	+10
Hindawi	39k	62%	111	--	5.0	+139%	+36pp	+16
PLOS	19k	1%	170	59%	NA	-23%	-3pp	+21

pp = 'percentage points'. Source: data scraped on the publishers' website or publishers' own publications. Overall: SI, TAT, Rejection % based on publishers with available data only. N per Scimago dataset. Elsevier: rejection rate change starts from 2018. All publishers: Special Issues sometimes named differently.

What's going on?

Why?

- MDPI, and others, are (were) successful
- Authors must *publish or perish*
- Funders (incl. Universities) drive this pressure
- Business model: rent space in *reputable* journals

Threats

- Waste of time and money
- *Quality signals* distorted

Conclusion: Respond to strain

Exponential growth is unsustainable

Goodhart's law is in action

Open Access not the problem per se

Transparency of key metrics poor

Predatory presses? Or, after Braudel, a *system* that feeds predators

Funders, universities, have biggest role to regulate the system (Wellcome, 2020)

End

**A journal's Impact Factor
tells me how good
someone's research is.
*Right? Wrong!***

**Find out how to use
metrics responsibly.**



University of Exeter