

Do states make scientific fields?

Curation, legitimacy, and the CDC

Peter McMahan &
Gabriel Lévesque
McGill University





Science and the state

- Relationship between ***state legitimacy*** and ***scientific research*** is a long-time focus of STS
- Bureaucratic enterprises depend on science
(Castilla 2009; Lévesque 2022)
- Strong focus on “coercive” processes:
regulation — enforcement — funding
(E.g. Braun 1998; Baker 2017)

Case: *The CDC*

State agency as an active participant in scientific discourse

Does the CDC influence scientific fields?

- CDC is a privileged *institutional actor* in the *construction of scientific fields*
- **Production:** Employ scientists to produce original research (in any journal)
- **Curation:** Sponsor journals to showcase current research (from any scientists)



Jean-Marc Côté (18...–19...). *La Chasse aux Microbes* [The Hunt for Germs] (1900). From the series France En L'an 2000 France in the Year 2000 (20th century). Ink on paper card. Dimensions not specified. Public domain image.

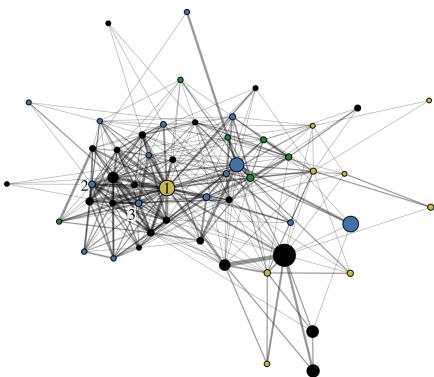
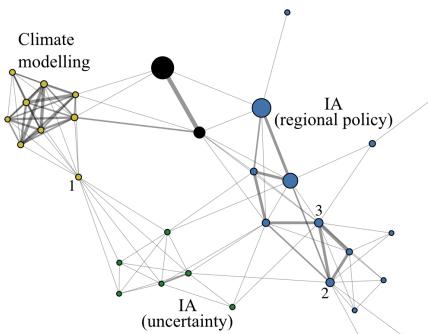
Scientific (sub)fields

- Maintained through *practice* — communication and publication of findings
- Site of *negotiation* around what counts as relevant **topics, methods, terminology, & explanatory frames**
(Hedgecoe 2002; Myers 1985; Hill & Carley 1999)
- Fields define and are defined by norms of research (standards)

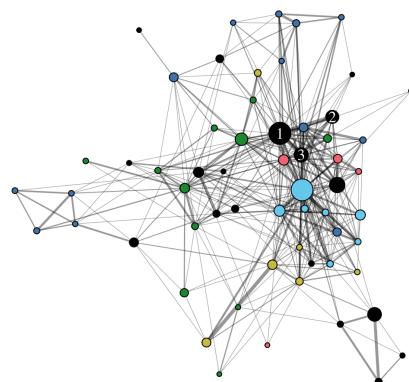
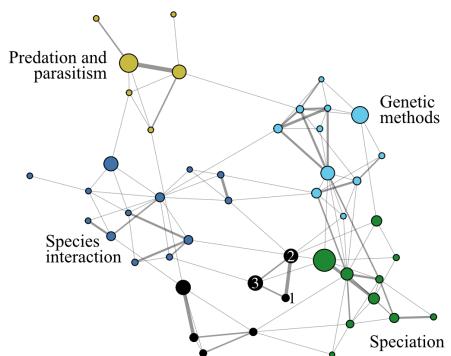
Construction of scientific knowledge

- Fields provide a shared frame, shaping the *translation* of research *output* into *findings*
(Knorr-Cetina 1999; Star & Griesemer 1989)
- Fields define the boundaries of *legitimate* science
(Latour 1983; Latour & Woolgar 1986)

Integrated assessment models of global climate change
(Parson and Fisher-Vanden 1997)



Geographic Structure of Insect Populations: Gene Flow, Phylogeography, and Their Uses
(Roderick 1996)



Unified language

- Scientific fields can help disparate researchers understand themselves as studying the “same” thing
- Shared explanatory frame or motivating narratives
(McMahan & McFarland 2021)
- Unifying terminology
(e.g. “AIDS” or “COVID-19”)

Fig. 8 from McMahan & McFarland (2021)

Analytic strategy

1 Measure linguistic convergence

Does scientific discourse on a topic converge around certain works after they are published?

- Ongoing negotiation concerning the legitimacy of frames/explanations/terminology/etc happens through scientific publication
- Are some publications stronger “attractors?”

2 Identify mechanisms of influence

Does the CDC’s privileged status as state agency lead to disproportionate influence?

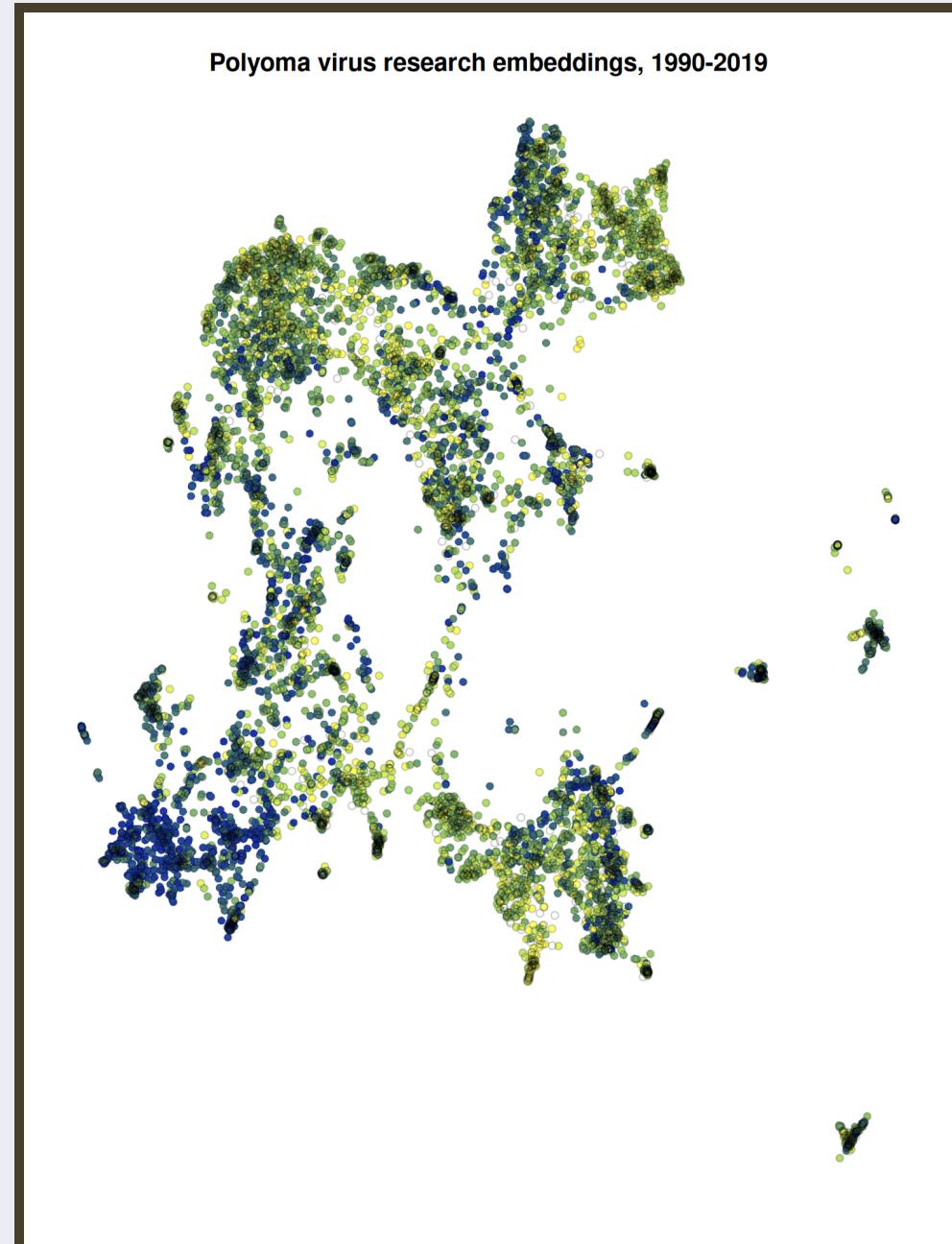
- **State curation:** does selection into a *CDC-published journal* affect articles’ influence?
- **State production:** does research carried out by *CDC scientists* have stronger influence?

Vector embeddings

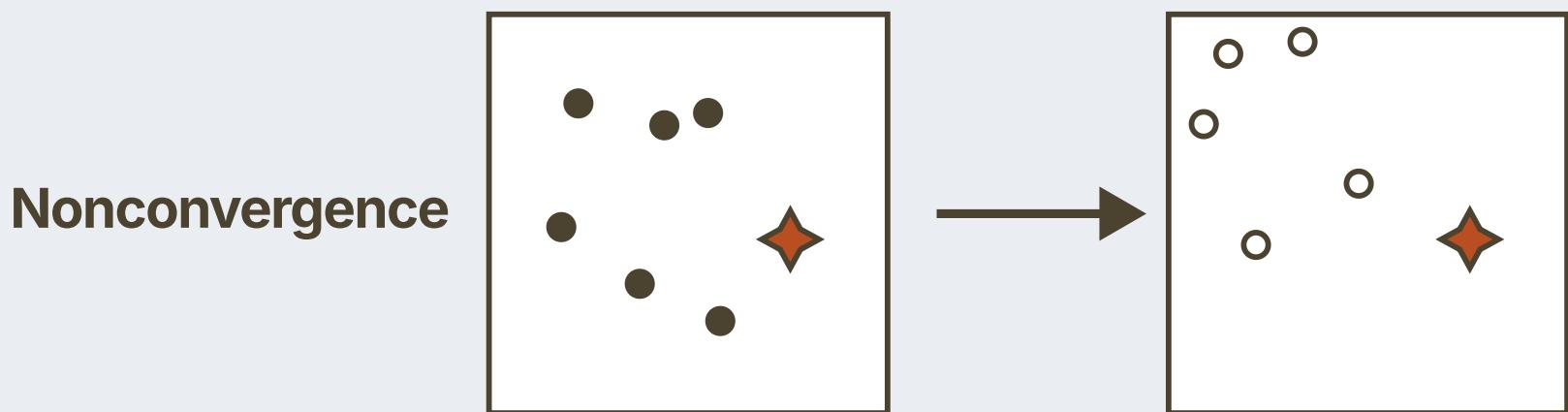
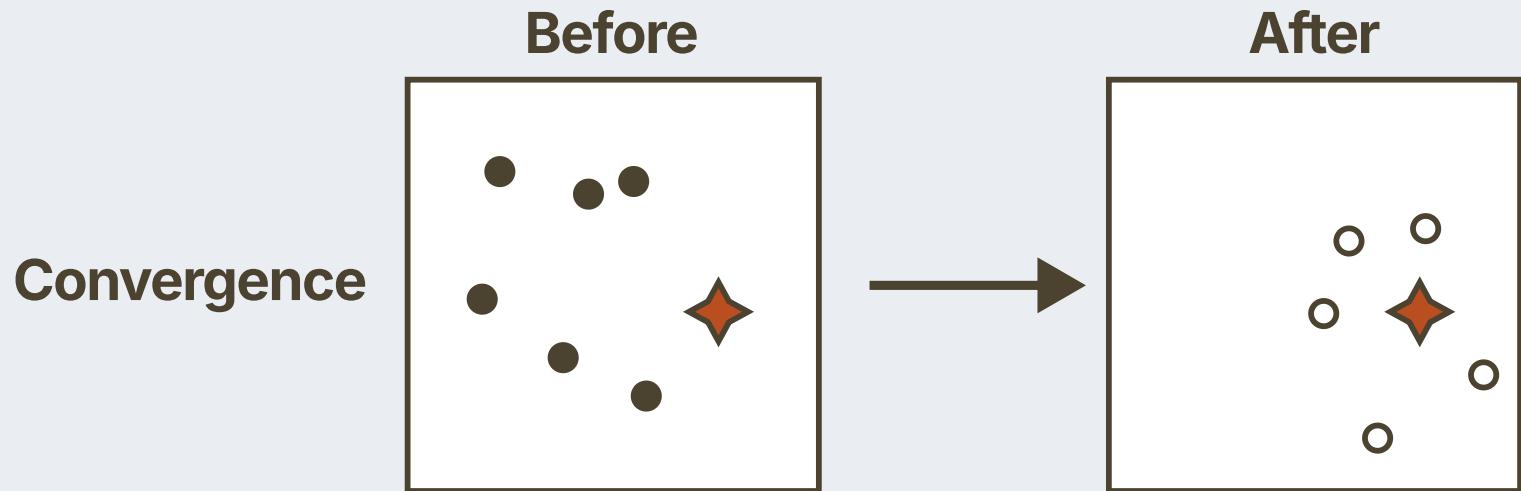
- Vector embeddings capture both the *topical* and the *rhetorical* features of text
- Embedding model:
 - Nomic Embed Text 1.5 (Nussbaum et al. 2024)
 - Open source, large context window, multimodal, & fast

Measuring convergence

- δ_i^b
average cosine similarity between article i and *all* articles published in the *previous 5 years*
- δ_i^a
average cosine similarity between article i and *all* articles published in the *following 5 years*
- $C_i = \delta_i^a - \delta_i^b$
convergence (increased similarity) of published work around article i



Two-dimensional projection (UMAP) of polyoma virus abstracts from 1990 (blue) to 2019 (yellow)



Data

- **OpenAlex**

Publications, journal details, author affiliations, abstracts, ...

(Priem, Piwowar, and Orr
2022)

Sample

- Publication year: 1990–2019
- 89 research topics in which CDC journals are active
 - *Emerging Infectious Diseases*
Preventing Chronic Disease
MMWR Morbidity and Mortality Weekly Report
 - Research topics automatically assigned by OpenAlex
- $N = 1,347,626$

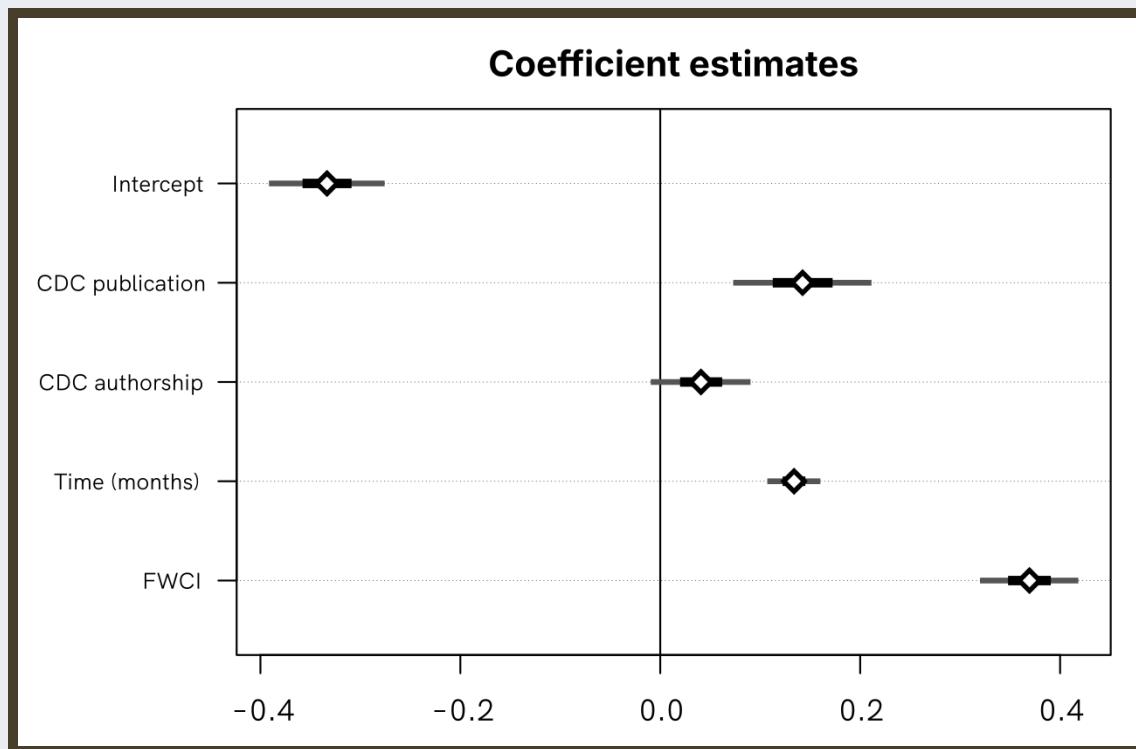
Matched sample

- Regression based on matched sample
 - **Treatment:**
Published in one of the three CDC journals **or**
At least 50% of authors are CDC affiliated
 - Matched on year, #authors, #institutions, #citations, & field-weighted citation index (FWCI) using nearest-neighbor (Mahalanobis)
- $N = 42,394$ (*treatment & control*)

Results (1)

Predicting standardized linguistic convergence with random intercepts and random slopes on *CDC publication* and *CDC authorship*

Coefficient estimates with 80% and 95% CI



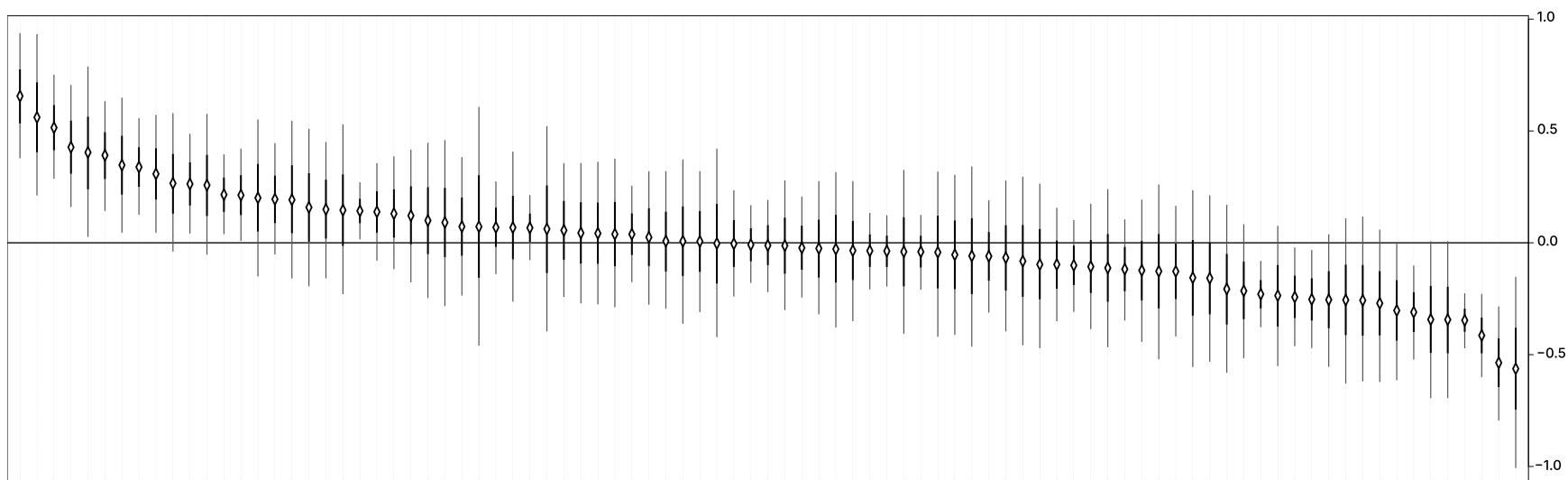
Model predictions

- Convergence far below average for non-CDC papers that are not cited
- The most highly cited papers (FWCI close to 1) can expect average levels of convergence
- Papers by CDC-affiliated authors are no different on average
- **Publications in CDC journals attract significantly increased linguistic convergence**

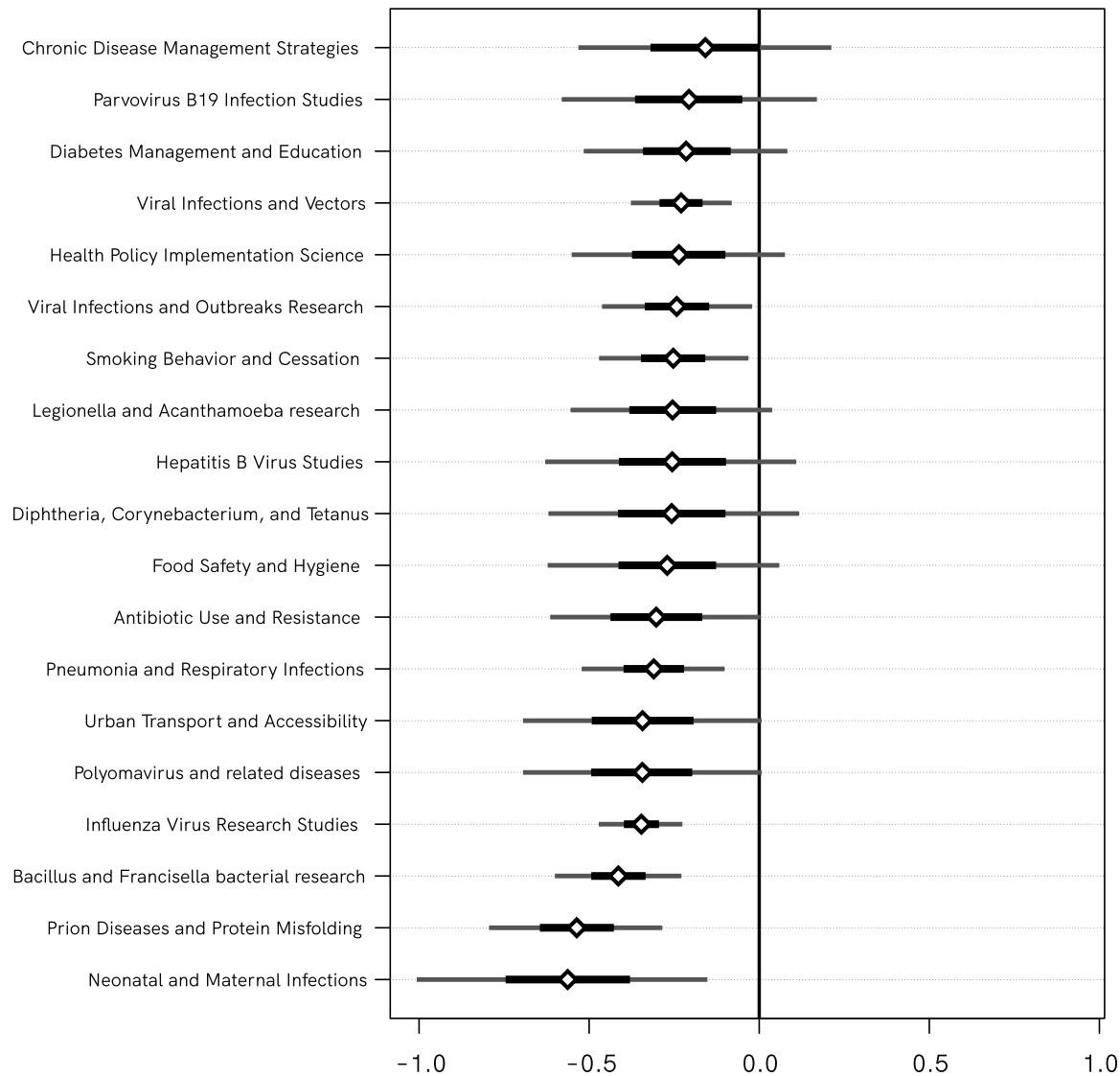
Results (2)

- Although the *average* effect of CDC publication is positive, the effect varies strongly by topic
- For some topics, CDC publication seems to drive future discourse *away*

Topic-level estimates of CDC publication convergence (random slopes with population effect)



Topic-level estimates of CDC publication convergence (lowest-magnitude estimates only)



State influence

- Population-level results suggest that field dynamics are driven disproportionately — *but by no means exclusively* — by the CDC as an institutional player

Mechanisms of influence

- CDC influence is predominantly realized through *curation* rather than *production* (on average)
- CDC's influence realized through *normative* rather than *mimetic isomorphism*?

(DiMaggio and Powell 1983; Reichmann 2011)

Variation by topic

- The “force” of the CDC is not uniform across domains
- Does CDC's influence depend on its *reputation* (perceived legitimacy) within a field?

(Carpenter 2010)

Thank you!

Slides available at

<https://github.com/mcmahanp/talks/>

Peter McMahan

(peter.mcmahan@mcgill.ca)

Gabriel Lévesque

(gabriel.levesque2@mail.mcgill.ca)



References

- Baker, Zeke. 2017. "Climate State: Science-State Struggles and the Formation of Climate Science in the US from the 1930s to 1960s." *Social Studies of Science* 47 (6): 861–87. <https://doi.org/10.1177/0306312717725205>.
- Braun, Dietmar. 1998. "The Role of Funding Agencies in the Cognitive Development of Science." *Research Policy* 27 (8): 807–21. [https://doi.org/10.1016/S0048-7333\(98\)00092-4](https://doi.org/10.1016/S0048-7333(98)00092-4).
- Carpenter, Daniel. 2010. *Reputation and Power: Organizational Image and Pharmaceutical Regulation at the FDA*. Princeton: Princeton University Press.
- Castilla, Emilio J. 2009. "The Institutional Production of National Science in the 20th Century." *International Sociology* 24 (6): 833–69. <https://doi.org/10.1177/0268580909343498>.
- DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48 (2): 147–60. <https://doi.org/10.2307/2095101>.
- Hedgecoe, Adam M. 2002. "Reinventing Diabetes: Classification, Division and the Geneticization of Disease." *New Genetics and Society* 21 (1): 7–27. <https://doi.org/10.1080/14636770220122746>.
- Hill, Vanessa, and Kathleen M. Carley. 1999. "An Approach to Identifying Consensus in a Subfield: The Case of Organizational Culture." *Poetics* 27 (1): 1–30. [https://doi.org/10.1016/S0304-422X\(99\)00004-2](https://doi.org/10.1016/S0304-422X(99)00004-2).
- Knorr-Cetina, K. 1999. *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge, Mass.: Harvard University Press. <http://swbplus.bsz-bw.de/bsz077639464inh.htm>.
- Latour, B., and S. Woolgar. 1986. *Laboratory Life: The Construction of Scientific Facts*. Princeton Univ Pr.
- Latour, Bruno. 1983. "Give Me a Laboratory and I Will Raise the World." In *Science Observed: Perspectives on the Social Study of Science*, edited by Karin D. Knorr-C and Michael Mulkay, 32.
- Lévesque, Gabriel. 2022. "From Right or Wrong to True or False: Moral and Epistemic Framing in Debates over Cannabis Policy Reformulation." *Politics & Policy* 50 (4): 773–94. <https://doi.org/10.1111/polp.12482>.
- McMahan, Peter, and Daniel A. McFarland. 2021. "Creative Destruction: The Structural Consequences of Scientific Curation." *American Sociological Review* 86 (2): 341–76. <https://doi.org/10.1177/0003122421996323>.
- Myers, Greg. 1985. "Texts as Knowledge Claims: The Social Construction of Two Biology Articles." *Social Studies of Science* 15 (4): 593–630. <https://doi.org/10.1177/030631285015004002>.
- Nussbaum, Zach, John X. Morris, Brandon Duderstadt, and Andriy Mulyar. 2024. "Nomic Embed: Training a Reproducible Long Context Text Embedder."
- Reichmann, Werner. 2011. "Institutionalizing Scientific Knowledge: The Social and Political Foundation of Empirical Economic Research." *Sociology Compass* 5 (7): 564–75. <https://doi.org/10.1111/j.1751-9020.2011.00384.x>.
- Star, Susan Leigh, and James R. Griesemer. 1989. "Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39." *Social Studies of Science* 19 (3): 387–420. <https://doi.org/10.1177/030631289019003001>.

All topics

- Salmonella and Campylobacter epidemiology
- Viral gastroenteritis research and epidemiology
- Smoking Behavior and Cessation
- Antibiotic Resistance in Bacteria
- Mosquito-borne diseases and control
- Opioid Use Disorder Treatment
- Food Security and Health in Diverse Populations
- Antimicrobial Resistance in Staphylococcus
- Influenza Virus Research Studies
- Zoonotic diseases and public health
- Brucella: diagnosis, epidemiology, treatment
- Escherichia coli research studies
- Viral Infections and Vectors
- Bacillus and Francisella bacterial research
- Tuberculosis Research and Epidemiology
- Legionella and Acanthamoeba research
- Chronic Disease Management Strategies
- Antibiotic Use and Resistance
- Parasitic infections in humans and animals
- Vaccine Coverage and Hesitancy
- Diabetes, Cardiovascular Risks, and Lipoproteins
- Cervical Cancer and HPV Research
- Vector-borne infectious diseases
- Clostridium difficile and Clostridium perfringens research
- Disaster Response and Management
- Antifungal resistance and susceptibility
- Pneumonia and Respiratory Infections
- Listeria monocytogenes in Food Safety
- Health disparities and outcomes
- Parasitic Infections and Diagnostics
- Food Safety and Hygiene
- Parasite Biology and Host Interactions
- Animal Disease Management and Epidemiology
- Streptococcal Infections and Treatments
- Virology and Viral Diseases
- Global Cancer Incidence and Screening
- Hepatitis Viruses Studies and Epidemiology
- Malaria Research and Control
- Toxoplasma gondii Research Studies
- Fungal Infections and Studies
- Animal Virus Infections Studies
- Bacterial Infections and Vaccines
- Prion Diseases and Protein Misfolding
- Herpesvirus Infections and Treatments
- Pneumocystis jirovecii pneumonia detection and treatment
- Diabetes Management and Education
- Poxvirus research and outbreaks
- Bartonella species infections research
- Mycobacterium research and diagnosis
- Respiratory viral infections research
- Obesity, Physical Activity, Diet
- Viral Infections and Outbreaks Research
- Research on Leishmaniasis Studies
- Health Policy Implementation Science
- Public Health Policies and Education
- Vector-Borne Animal Diseases
- Viral Infections and Immunology Research
- HIV/AIDS Research and Interventions
- Hepatitis B Virus Studies
- Data-Driven Disease Surveillance
- Parasitic Diseases Research and Treatment
- Physical Activity and Health
- Vibrio bacteria research studies
- Hepatitis C virus research
- Colorectal Cancer Screening and Detection
- Reproductive tract infections research
- Rabies epidemiology and control
- Syphilis Diagnosis and Treatment
- Blood Pressure and Hypertension Studies
- HIV Research and Treatment
- Leptospirosis research and findings
- Diphtheria, Corynebacterium, and Tetanus
- Burkholderia infections and melioidosis
- Yersinia bacterium, plague, ectoparasites research
- Parasites and Host Interactions
- Trypanosoma species research and implications
- Infections and bacterial resistance
- Virus-based gene therapy research
- Urban Transport and Accessibility
- Polyomavirus and related diseases
- Healthcare Policy and Management
- Travel-related health issues
- Primary Care and Health Outcomes
- Obesity and Health Practices
- Bacterial Identification and Susceptibility Testing
- Mollusks and Parasites Studies
- Neonatal and Maternal Infections
- Global Public Health Policies and Epidemiology
- Parvovirus B19 Infection Studies